Comparative Analysis of Different Machine Learning Classifiers for the Prediction of Chronic Diseases

Kirti Gupta, Pardeep Kumar, and Shuchita Upadhyaya

Kurukshetra University, Kurukshetra, kirtigupta811@gmail.com, mittalkuk@gmail.com, shuchita_bhasin@yahoo.com

Abstract.

Chronic Diseases are the most dangerous diseases for humans and have significant effects on human life. Chronic Diseases like heart disease & Diabetes are the main causes of death. Precise diagnosis of these diseases on time is very significant for maintaining a healthy life. A comparative study of different machine learning classifiers for chronic disease prediction viz Heart Disease & Diabetes Disease is done in this paper. This paper forms the basis of understanding the difficulty of the domain and the amount of efficiency achieved by the various methods recently.

Keywords. «Machine Learning (ML), Prediction, Classification, Support Vector Machine (SVM), Decision Tree (DT), Artificial Neural Network (ANN)».

1. INTRODUCTION

The key demand of healthcare organizations tends to provide quality treatment at an affordable rate. The exact diagnosis of patients is required on time for delivering quality services to them. Unwanted and insufficient outcomes may produce poor quality clinical diagnosis and care. Machine learning techniques help in fast decision-making & reducing cost by making use of historical clinical data. Healthcare helps in maintaining healthy life by making use of proper preventive treatment. Automated techniques can be used to optimize the cost, facility, speed, precision, and reliability of this decision-making process.

Chronic diseases are prime reasons for death everywhere. Heart disease adversely affects the health of an enormous population [1]. The principal Cause of Heart Disease is the unhealthy & fast life style of human beings.

Diabetes is a chronic disease that occurs because of awkwardness in the discharge of insulin. Due to this, sugar level of blood remains unsettled. Diabetes is a vital cardiovascular disorder that can adversely affect the whole-body system. Early detection of diabetes helps in maintaining a healthy life [2]. Patients can have weight loss, less vision, infection, frequent

urination, etc. There are 3 categories of diabetes viz. Type I, Type II & Gestational diabetes [3]. The State in which human body declines to deliver insulin is termed as type I. It mainly occurs during childhood or adolescence [3]. The state in which human body becomes resistant to insulin is termed as type II. It occurs when insulin is inside, but the human body is not using it correctly. It's more common in adults. Long-term complications that occur because of diabetes are kidney issues, coronary illness, stroke, and tumor. Pregnant women are distressed by Gestational diabetes. Diagnosis of Gestational Diabetes on time is crucial for the safety of both the mother and infant. The main aims of healthcare organizations are to minimize cost, increase the correctness of the result and be more patient-centric [4].

In this paper, some recent researches related to heart and diabetes disease are discussed in section II. Datasets used are discussed in section III. Experimental results are deliberated in section IV & Conclusion is delineated in section V.

2. LITERATURE REVIEW

Early detection & diagnosis of heart diseases help in maintaining a healthy life. Deepika & Seema [5] performed a comparative analysis on "Naive Bayes, Decision tree, Support Vector Machine (SVM) classifiers" to predict heart & diabetes diseases. It was observed for heart disease prediction that SVM gave the highest accuracy rate of 95.5%. For diabetes prediction, the accuracy of naive bayes was 73.5% which was the highest among all compared algorithms. Nikhar & Karandikar [6] compared Naive Bayes algorithm & decision tree algorithm to find out which of them is more suitable for heart disease prediction. From observation, they originate that the DT algorithm is better than NB. Feshki & Shinjani [7] used Particle Swarm Optimization & Neural Network Feed Forward BackPropagation method for anticipating affected & non-affected patients. Their main focus was on reducing the cost by using feature selection with variable ranking. They used 8 features namely Gender, Age, Blood Pressure, Cholesterol, FBS, Exercise Features (Old peak, Slope, Ex_ang) in their study. Thomas et al. [8] used ANN & ID3 algorithm in their study. Firstly, they applied KNN algorithm to classify age then they used ID3 algorithm for heart disease prediction. They found Smoking & history of heart disease as two important factors for heart disease prediction. Shah et al. [9] increased the accuracy of the SVM technique for heart disease prediction. They used RBF based SVM on three datasets cleveland, Switzerland, and Hungarian of UCI repository. The proposed technique gave 82.2%, 85.8%, and 91.3% accuracy for <<Cleveland, Hungarian, and Switzerland datasets respectively>>. Mohan et al. [10] proposed a hybrid method HRFLM for the prediction of cardiovascular disease. They joined the features of Random Forest & linear method in HRFLM. They used 13 features of UCI machine learning heart disease dataset. They performed their prediction on the R studio. The proposed model gave an 88.7% accuracy for heart disease prediction.

For diabetes prediction, Various researchers proposed various machine learning models in their study. Nirmala et al. [11] proposed an Amalgam KNN model- a combination of KNN and k-mean clustering for diabetes prediction. They used 10-fold cross-validation with different k values using WEKA Software tool. The proposed hybrid algorithm gave higher performance than simple KNN & k-mean clustering algorithms. Sanakal & Jayakumari [12] compared fuzzy C-Mean clustering with the SVM algorithm for diabetes prediction. The

accuracy of SVM & FCM was 59.5% & 94.3% respectively. They found fuzzy c-mean clustering as a better tool for the diabetes prediction. Vijayan et al. [13] used decision stump as a base classifier in the AdaBoost algorithm for diabetes prediction. They used a dataset of UCI repository as well as a dataset of Kerala in their study. The proposed algorithm gave an accuracy of 80.72% for the prediction of diabetes. The proposed algorithm gave better results than SVM, Naive Bayes & Decision Tree classifiers. Santhanam & Padmavathi [14] proposed a diabetes prediction model in which they used SVM as a base classifier. For reducing instances & for removing noisy data, they used K-mean clustering. For attribute selection, they used a Genetic algorithm. The proposed model took advantage of both supervised (SVM) & unsupervised algorithm (K-mean clustering). Anand et al. [15] developed a GUI-based diabetes prediction model based on the present lifestyle of people. For conducting the research, questionnaires were prepared with the help of doctors. The attributes included in their questionnaire were Eating Habit (Roadside Eating, Junk food), Sleeping Duration, Sugar Intake, Exercise Duration, Blood Pressure, BMI, Heredity Diabetes, Gender, Belly Size, etc.

3. DATASETS AND METHODOLOGIES

For Comparing different classifiers on heart disease, two heart disease datasets are used in this paper. One is available on Kaggle website consisting of 1190 records having 12 features of patients from US, UK, Switzerland and Hungary & other is available on “UCI machine learning repository"[16] consisting of 303 records having 14 attributes from Cleveland clinic foundation. For diabetic prediction, "Pima Indian diabetes dataset" from Kaggle website is used in this paper. The dataset contains 768 records of people. "For the diabetes disease prediction dataset of Tabriz, Iran is also used by researchers"[16].

Table 1: Datasets Description

Dataset	No. of Records	Attributes	Dataset Source
HD Dataset1	303	14	UCI ML repository
HD Dataset 2	1190	12	Kaggle website
Diabetes dataset	768	9	Kaggle website

Table 2: Results on Diabetes Dataset

Models	Accuracy	Classification Error	Recall	Precision	F1-score
Naive Bayes	0.75	0.25	0.79	0.81	0.80
K-Nearest Neighbors	0.74	0.26	0.82	0.79	0.80
Logistic Regression	0.71	0.29	0.83	0.75	0.79
Linear Kernel SVM	0.77	0.23	0.86	0.80	0.83
Polynomial Kernel SVM	0.78	0.22	0.87	0.80	0.83
RBF Kernel SVM	0.76	0.24	0.86	0.79	0.82
Decision Tree	0.73	0.27	0.82	0.78	0.80
Random Forest	0.75	0.25	0.79	0.81	0.80
Neural Network	0.75	0.25	0.88	0.77	0.82

Table 3: Results on Heart Disease Dataset 1

Models	Accuracy	Classification Error	Recall	Precision	F1-score
Naive Bayes	0.81	0.19	0.81	0.78	0.89
K-Nearest Neighbors	0.84	0.16	0.88	0.79	0.84
Logistic Regression	0.88	0.12	0.73	1.00	0.84
Linear Kernel SVM	0.86	0.14	0.73	0.95	0.83
Polynomial Kernel SVM	0.83	0.17	0.65	0.94	0.77
RBF Kernel SVM	0.86	0.14	0.77	0.91	0.83
Decision Tree	0.83	0.17	0.69	0.90	0.78
Neural Network	0.88	0.12	0.73	1.00	0.84
Random Forest	0.90	0.10	0.77	1.00	0.87

Table 4: Results on Heart Disease Dataset 2

Models	Accuracy	Classification Error	Recall	Precision	F1-score
Naive Bayes	0.84	0.16	0.85	0.82	0.83
K-Nearest Neighbors	0.87	0.13	0.90	0.83	0.87
Logistic Regression	0.83	0.17	0.81	0.83	0.82
Linear Kernel SVM	0.84	0.16	0.83	0.84	0.83
Polynomial Kernel SVM	0.86	0.14	0.87	0.84	0.85
RBF Kernel SVM	0.85	0.15	0.81	0.87	0.84
Decision Tree	0.84	0.16	0.79	0.86	0.83
Neural Network	0.84	0.16	0.80	0.84	0.82
Random Forest	0.91	0.09	0.91	0.89	0.90

4. EXPERIMENTAL RESULTS AND ANALYSIS

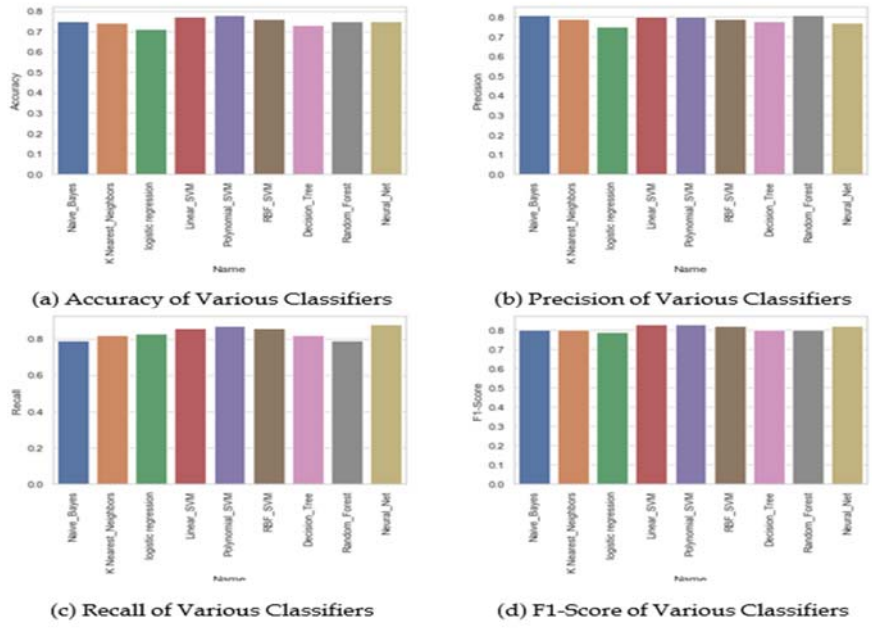
Seven different Classifiers namely "Naive Bayes, KNN, Logistic regression, SVM, Decision tree, Random Forest, Neural Network" are applied on the all mentioned datasets for prediction of diabetes & heart disease. For K-NN Classifier, different values of k are used

in order to find the best possible output of the model. Similarly, for Finding the best performance of SVM, different kernels like linear, polynomial & RBF are applied on the datasets. From Observations, it can be concluded that SVM is best among all other specified classifiers for the prediction of diabetes. From Observations, it can also be deduced that Random Forest is best among all other specified classifiers for heart disease prediction.

5. CONCLUSION & FUTURE SCOPE

Diagnoses & prediction of chronic disease are the toughest challenges in the pharmaceutical field. It is based on the thorough investigation of clinical data of victims. Because of advancement in machine learning and IT, Feature Selection becomes easy. This also decreases the complexity of the system & increases its accuracy. It is recommended that Feature Selection & Feature Ranking should be done based on Cost & Time so that one should get accurate & precise result in a short time & with minimum cost possible. Researchers have tried several different models and the one with the highest accuracy is picked for the Pre- diction of Chronic Diseases. It has been observed that not only the choice of machine learning technique but also the dataset selected for prediction affects the efficiency of the model.

As prediction deals with uncertainty, machine learning techniques can be combined with fuzzy logic to better deal with uncertainty. Both soft & statistical data can be used in combination to effectively deal with uncertainties in the medical field.



(a) Accuracy of Various Classifiers

(b) Precision of Various Classifiers

(c) Recall of Various Classifiers

(d) F1-Score of Various Classifiers

Figure 1. Performance of Various Classifiers on Diabetes Dataset

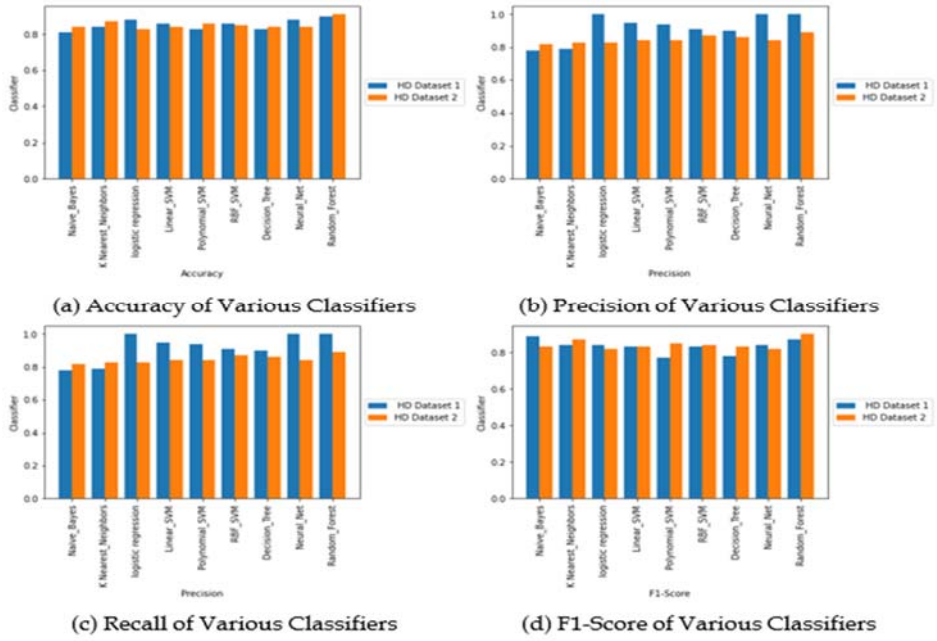


Figure 2. Performance of Various Classifiers on Heart Disease Datasets

6. REFERENCES

[1] T. Karayılan, O. Kılıç, 'Prediction of heart disease using neural network', In 2017 International Conference on Computer Science and Engineering (UBMK), pages 719–723. IEEE, 2017.

[2] H. Kaur, V. Kumari, 'Predictive modelling and analytics for diabetes using a machine learning approach.', Applied computing and informatics, 2020.

[3] C Kalaiselvi, G. Nasira, 'A new approach for diagnosis of diabetes and prediction of cancer using anfis', In 2014 World Congress on Computing and Communication Technologies, pages 188–190. IEEE, 2014.

[4] B Nithya. 'Study on predictive analytics practices in health care system', IJETTCS, 5:98–102, 2016.

[5] K. Deepika, S. Seema. 'Predictive analytics to prevent and control chronic diseases', In 2016 2nd International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), pages 381–386. IEEE, 2016.

[6] S. Nikhar, A. Karandikar. 'Prediction of heart disease using machine learning algorithms', International Journal of Advanced Engineering, Management and Science, 2(6):239484, 2016.

[7] M. Feshki, O. Sojoodi Shijani, 'Improving the heart disease diagnosis by evolutionary algorithm of pso and feed forward neural network', In 2016 Artificial Intellience and Robotics (IRANOPEN), pages 48–53. IEEE, 2016.

[8] J. Thomas, R. Theresa Princy, 'Human heart disease prediction system using data mining techniques', In 2016 international conference on circuit, power and computing technologies (ICCPCT), pages 1–5. IEEE, 2016.

[9] S. Muhammad et al., 'Feature extraction through parallel probabilistic principal component analysis for heart disease diagnosis', Physica A: Statistical Mechanics and its Applications, 482:796–807, 2017.

[10] S. Mohan et al. 'Effective heart disease prediction using hybrid machine learning techniques', IEEE Access, 7:81542–81554, 2019.

[11] M. NirmalaDevi et al., 'An amalgam knn to predict diabetes mellitus', In 2013 IEEE International Conference ON Emerging Trends in Computing, Communication and Nanotechnology (ICECCN), pages 691–695, 2013.

[12] R. Sanakal, T. Jayakumari, 'Prognosis of diabetes using data mining approach-fuzzy c means clustering and support vector machine', International Journal of Computer Trends and Technology, 11(2):94–98, 2014.

[13] V. Vijayan, C. Anjali, 'Prediction and diagnosis of diabetes mellitus—a machine learning approach', In 2015 IEEE Recent Advances in Intelligent Computational Systems (RAICS), pages 122–127. IEEE, 2015.

[14] T. Santhanam, M.S. Padmavathi, 'Application of k-means and genetic algorithms for dimension reduction by integrating svm for diabetes diagnosis', Procedia Computer Science, 47:76 – 83, 2015. Graph Algorithms, High Performance Implementations and Its Applications (ICGHIA 2014).

[15] A. Anand, D. Shakti, 'Prediction of diabetes based on personal lifestyle indicators', In 2015 1st International Conference on Next Generation Computing Technologies (NGCT), pages 673–676, 2015.

[16] Weifeng Xu et al., 'Risk prediction of type ii diabetes based on random forest model', In 2017 Third International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB), pages 382–386. IEEE, 2017.

Deep Learning Methods for Anomaly Detection

Rajeev Subedi[1, a)] **Anil Kumar** [2, b)] **Narendra Kumar** [3, c)]

Author Affiliations
[1, 2, 3] *School of Computing, DIT University Dehradun, Uttarakhand, India*
Author Emails
[a)] *Corresponding author: 1000014411@dit.edu.in*
[b)]*dahiyaanil@yahoo.com*

[c)]*narendra298@gmail.com*

Abstract.

The invention and upgrading of systems and instruments that collect data at each point are narrowing the world. As a result, data is growing at an exponential rate. We are unable to process such a vast volume of information in a timely manner. The data comprises numbers that change over time, as well as other factors and functions that cause it to fluctuate suddenly. Anomaly is a term used to describe an abrupt change in data from its standard. The purpose of this survey is to present the deep learning approaches and the deep learning architectures with which we can detect anomalies. The review also presents the summarized view of the various approaches, techniques, datasets used to detect the anomalies in the studies done over the years. The prime goal of this paper is to summarize the various deep learning approach, architecture, and datasets with which we can detect anomalies.

Keywords. Anomaly, Outliers, Neural network, Deep learning, Auto encoder.

1. INTRODUCTION

The Anomalies are rapid changes in data from normal deviation into clusters. Anomaly detection is the process of analysing the data provided by machines or sets of machines to find abnormalities. Anomaly detection has recently had a big influence in areas including hospital monitoring systems, security, banking, IoT and sensor networks, marketing, and natural disasters. As a consequence, anomaly detection has been a popular research topic for decades, owing to its intrinsic complexity and tumultuous nature[1]. While traditional techniques of identifying anomalies have existed for a long time, deep learning emerged to outwit numerous learning tasks by learning complicated data using neural networks, which are computer representations of brain neurons. Many research methodologies are employed to build the most effective model that can identify irregularities with a reduced signal to noise ratio. The goal of such a model is to detect as many outlier clusters as possible while reducing risk. The necessity for new models that can evaluate enormous datasets was essential since previous approaches could not handle such quantities to create outliers. Traditional approaches failed to optimise time-series, picture, and sequential data due to their complicated structures[2]. Whereas a standard model relies on human feature selection

from a dataset, a model using deep learning outlier detection may learn graded discriminative features[2].

2. DIFFERENT DEEP LEARNING APPROACHES FOR ANOMALY DETECTION

2.1. Supervised anomaly detection

Separating training and test datasets is essential for supervised anomaly detection. Despite better efficiency and outcomes, semi-supervised and unsupervised techniques of anomaly detection were more reliable due to greater unlabeled training samples in the created dataset [2][3].

2.2. Unsupervised Anomaly Detection

Unsupervised deep learning anomalydetection builds and trains a model using an unclassified and uncategorized dataset[1]. A model evaluates unscaled datasets and predicts hidden patterns. It is explained in the study publication [2] that an unsupervised anomaly detection system is able to learn from the data and find anomalies by separating normal from abnormal data points.

2.3. Semi-Supervised Anomaly Detection

It is using both scaled and unscaled datasets to train a model is called hybrid learning[4][5].The scaled to unscaled data ratio is decreasing, indicating that less scaled data is used for training and more unscaled data is used for better outcomes [6].

3. DIFFERENT ARCHITECTURES IN DEEP LEARNING FOR ANOMALY DETECTION

3.1. Multi-layer Perceptron neural network

Multilayer perceptron (MLP) is a multiple layered feed-forward neural network[7], [8].

These activation functions[9] are represented as

$$y(v_i) = \tanh(v_i) \quad (3.1)$$

$$y(v_i) = (1 + e^{-vi})^{-1} \quad (3.2)$$

Here, back-propagation, popular supervised learning technique is used for training the dataset[9].

The multilayer perceptron neural network as given:

Let us consider d_j as the targeted value and y_j as the output value calculated by the perceptron algorithm. Now the error in one neuron is generated by:

$$e_j(n) = d_j(n) - y_j(n) \quad (3.3)$$

with which our error function is calculated by:

$$\varepsilon(n) = \frac{1}{2} \Sigma e_j^2(n) \quad (3.4)$$

Now apply the gradient descent, the value while updating the weights in each cycle is calculated by:

$$\Delta w_{ij}(n) = -\eta (d\varepsilon / dv_j) y_i(n) \quad (3.5)$$

3.2. Convolution Neural Network(CNN)

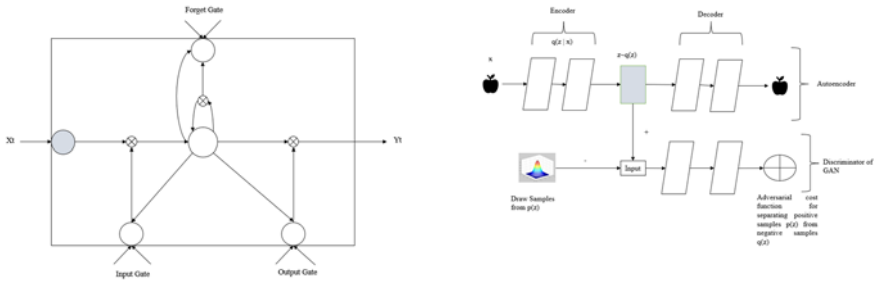
CNN prioritises the local dependencies in the input data gathered earlier in the process. Many engineers, academicians built CNN models using image datasets. A variety of methodologies and datasets have been used by other researchers to create the model [10]-[11].

3.3. Recurrent Neural Network

It is a type of advanced deep learning that operates by sending the hidden layer output back to the input layer, resulting in the prediction of that layer's output [14]. Unlike other neural networks, RNNs have an internal memory that allows them to perform the same function for each input.

3.4. Long Short-Term Memory

In Researchers developed the LSTM as a sophisticated RNN in the late 1990s and early 2000s[10]. The vanishing gradient issue in RNN happens when back-propagation is used to train a model [17]. LSTM is an efficient solution. Advanced recurrent neural network architecture that uses previously stored data to analyse and forecast abnormalities [15].



(a) LSTM network architecture[14] (b) An architecture of adversarial autoencoder[19]

Figure 3.1 An architecture of LSTM Network and adversarial autoencoder

3.5 Adversarial Auto Encoder

Adversarial Autoencoder (AAE) is a promising strategy in deep learning to identify anomalies that may transform an autoencoder model into a generative adversarial network with the primary goal of minimising AE reconstruction error and mapping a prior to the hidden code vector[13].

3.6 Restricted Boltzmann Machine

It is a stochastic neural network controlled by energy principles and distinguished by variables such as entropy, energy, and temperature[16]. With applications ranging from feature extraction to filtering to pre-training neural networks to picture reconstruction [33-42], networks are increasingly becoming unsupervised learning machines. RBM is an

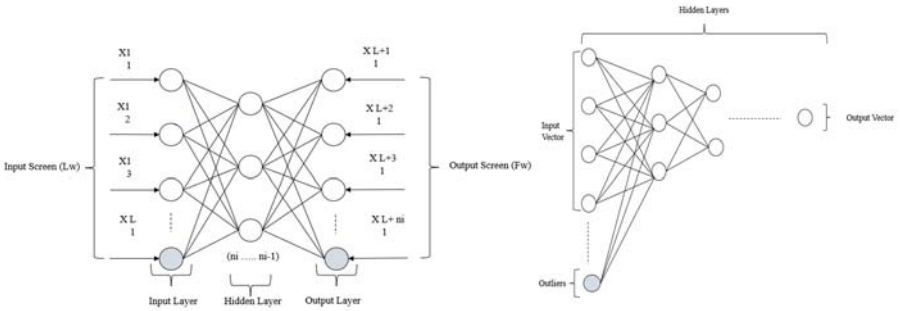
unsupervised learning technique with two layers: input and hidden. Because there is no output layer, the propagation entails returning the reconstruction to the input layer[15].

3.7 Generative Adversarial Networks

In order to simulate the large variety of distribution of complicated and multi-dimensional data, we need more sophisticated approaches, such as generative adversarial networks (GANs)[16],[17]. The primary flaw of GANs is that they are unstable to train and provide absurd outputs.

3.8 Deep Belief Networks

Deep Belief networks (DBNs)[18] are generative model for neural network that stack RBMs layer by layer and carry out the learning process.



(a) Multilayer perceptron network architecture[7] **(b)**Deep belief network architecture

Figure 3.3 Multilayer perceptron and Deep belief network architecture[21].

Above architecture of DBN shows us the combine structure of simple and hierarchical connected RBMs [19].

3.9 Autoencoder

The key objective of an autoencoder is to learn in an unsupervised manner interpretation of the dataset which can later be used to solve several real-life applications[20].Also, several variations of regularization techniques for autoencoders were studied in various case studies which includes Sparse Autoencoders[21]–[23],Denoising Autoencoders[24], [25] and Contractive Autoencoders[26].

3.10 Deep Neural Networks

Moreover, with these features DNNs is widely used for anomaly detection methods surpassing the effectiveness and accuracies of MLP, CNN, RNN, LSTM, RBM and LSTM and are being used as black boxes in a model with their dense non-linear structure[27], [28].

4. COMPARISON RESULTS

Supervised, semi-supervised and unsupervised approaches were presented in the paper reviewed throughout the research, but most importantly unsupervised learning approach is most desired for anomaly detection.Couple of more anomaly detection architecture based on LSTM[29], LSTM-SVM[30], CNN-LSTM[11], LSTM-RNN[32] and Stacked

LSTM[31] were observed where we could see various decision function and some methods are extremely good where maximum ratio of anomalies was detected whereas some were average.This research was conducted by making model to detect anomalies from normal patters using deep Auto-encoder and restricted Boltzmann machine which uses backpropagation by setting equal inputs, outputs and generate the results based on MSE, RMSE and AUC.Last but not the least, we observed Deep Auto-encoder (DAE) architecture [33]which elaborates the need of hybrid semi-supervised anomaly detection model which could minimize the complexity observed in dimensionality in high dimensional space for high-dimensional data.

5. CONCLUSION

This paper presents the Deep learning anomaly detection approach and methods along with the literature survey based on the approach, detection technique, architecture, datasets used, and result obtained. The prime goal of this paper was to summarize the various deep learning approach, architecture, and datasets with which we can detect anomalies. Another aspect of this research was to find the application areas within deep learning for which we can opt anomaly detection technique. Anomaly detection using deep learning technique is the major area of research and this paper could be useful for researchers in the field.

6. REFERENCES

1. G. Pang, C. Shen, L. Cao, and A. van den Hengel, ACM Computing Surveys 54, (2021).
2. R. Chalapathy and S. Chawla, (2019).
3. N. Görnitz NICOGOERNITZ, K. Rieck KONRADRIECK, and U. Brefeld, Toward Supervised Anomaly Detection Marius Kloft (2013).
4. B.R. Kiran, D.M. Thomas, and R. Parakkal, Journal of Imaging 4, (2018).
5. E. Min, J. Long, Q. Liu, J. Cui, Z. Cai, and J. Ma, in Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (Springer Verlag, 2018), pp. 322–334.
6. P. Perera and V.M. Patel, IEEE Transactions on Image Processing 28, 5450 (2019).
7. G. Raman MR, N. Somu, and A.P. Mathur, International Journal of Critical Infrastructure Protection 31, (2020).
8. R. Collobert and S. Bengio, Links between Perceptrons, MLPs and SVMs (2004).
9. 9 T. Teoh, P. Ng, G. Chiew, E.J. Franco, and de Y. Goh, Anomaly Detection in Cyber Security Attacks on Networks Using MLP Deep Learning (n.d.).
10. A. Krizhevsky, I. Sutskever, and G.E. Hinton, Communications of the ACM 60, 84 (2017).
11. Y. Heryadi and H.L.H.S. Warnars, in 2017 IEEE International Conference on Cybernetics and Computational Intelligence, CyberneticsCOM 2017 - Proceedings (Institute of Electrical and Electronics Engineers Inc., 2018), pp. 84–89.
12. A. Dimokranitou, G. Tsechpenakis, J. Yu Zheng, and M. Tuceryan, STATEMENT OF THESIS APPROVAL (2017).
13. G.H. de Rosa, M. Roder, D.F.S. Santos, and K.A.P. Costa, International Journal of Information Technology (Singapore) 13, 49 (2021).
14. A. Pumsirirat and L. Yan, Credit Card Fraud Detection Using Deep Learning Based on Auto-Encoder and Restricted Boltzmann Machine (2018).

15. I.J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, (2014).
16. A. Radford, L. Metz, and S. Chintala, (2015).
17. M.K. Sharma, D. Sheet, and P.K. Biswas, in ACM International Conference Proceeding Series (Association for Computing Machinery, 2016).
18. I. Kakanakova and S. Stoyanov, in ACM International Conference Proceeding Series (Association for Computing Machinery, 2017), pp. 73–79.
19. D. Bank, N. Koenigstein, and R. Giryes, (2020).
20. P. Malhotra, L. Vig, N. Gugulothu, and G. Shroff, Sparse Neural Networks for Anomaly Detection in High-Dimensional Time Series (2018).
21. M.R. Shahid, G. Blanc, Z. Zhang, and H. Debar, (2019).
22. J. Sun, X. Wang, N. Xiong, and J. Shao, IEEE Access 6, 33353 (2018).
23. J. Chen, J. Li, W. Chen, Y. Wang, and T. Jiang, Renewable Energy 147, 1469 (2020).
24. M.G. Narasimhan and S. Sowmya Kamath, Multimedia Tools and Applications 77, 13173 (2018).
25. S.F. Lokman, A.T. Othman, S. Musa, and M.H. Abu Bakar, in Advanced Structured Materials (Springer Verlag, 2019), pp. 195–205.
26. K. Amarasinghe, K. Kenney, and M. Manic, in Proceedings - 2018 11th International Conference on Human System Interaction, HSI 2018 (Institute of Electrical and Electronics Engineers Inc., 2018), pp. 311–317.
27. W. Samek, T. Wiegand, and K.-R. Müller, (2017).
28. G. Kim, H. Yi, J. Lee, Y. Paek, and S. Yoon, (2016).
29. T. Ergen and S.S. Kozat, IEEE Transactions on Neural Networks and Learning Systems 31, 3127 (2020).
30. P. Baldi, Autoencoders, Unsupervised Learning, and Deep Architectures (2012).
31. T.K. Dang, R. Wagner, J. Küng, N. Thoai, M. Takizawa, and E. Neuhold, editors , Future Data and Security Engineering (Springer International Publishing, Cham, 2016).
32. H. Song, Z. Jiang, A. Men, and B. Yang, Computational Intelligence and Neuroscience 2017, (2017).
33. N. Kumar, H. Shukla and R. Tripathi, International Journal Of Intelligent Engineering And Systems 10, (2017).
34. H. Shukla, N. Kumar and R. Tripathi, International Journal Of Computer Applications 95, (2014).
35. N. Kumar, A. Dahiya, K. Kumar and S. Tanwar, 2021 9Th International Conference On Reliability, Infocom Technologies And Optimization (Trends And Future Directions) (ICRITO) (2021).
36. K. Kumar, N. Kumar and R. Shah, International Journal Of Intelligent Networks 1, (2020).
37. N. Kumar, H. Shukla and R. Tripathi, International Journal Of Intelligent Engineering And Systems 10, (2017).
38. A. Sharma, N. Kumar, International Journal of Science and Research (IJSR) 3,(2014)
39. N. Kumar, H. Shukla, A. Tiwari and A. Dahiya, SSRN Electronic Journal (2019).
40. N. Kumar, H. Shukla and R. Tripathi, International Journal Of Intelligent Engineering And Systems 10, (2017).
41. N. Kumar, K. Kumar and A. Kumar, Sersc.Org (2020).
42. K. Kumar, R. Singh, P. Ranjan and N. Kumar, Algorithms For Intelligent Systems ((2021).

Biographies

Er. Rajeev Subedi received his bachelor's degree (B.Tech) in Computer Science & Engineering from Chandigarh University in 2019 and currently perusing M.Tech in Computer Science & Engineering from DIT University respectively. He is also working as an IT consultant in multiple schools and colleges in Nepalgunj, Nepal. His research areas include machine learning, deep learning, and cloud computing.

Dr. Anil Kumar is now employed at DIT University as a Professor of CSE,Head-Data Science Research Group and Accreditation Coordinator . Prof. Anil Kumar received his M.Tech. from the Delhi College of Engineering and his Ph.D. from the Manipal Group. He has over 24 years of teaching and working experience. He is an IEEE Senior Member and served on the Executive Committee of the IEEE Computer Society India Council in 2015 and 2016. His research interests include cryptography, image processing algorithms, artificial intelligence, neural systems, signal and system analysis, and genetic algorithms.

Dr. Narendra Kumar completed his M.Tech(Computer Science) from BIT Mesra Ranchi, and Ph.D.(Computer Science) from D.D.U Gorakhpur University. Narendra Kumar has more than 14 years' rich experience in Computer Science & Engineering field. He is working at DIT University as an Assistant Professor for the last 14 years, His research area includes image processing, Optimizations techniques and the internet of things and deep learning. Narendra is Editor of two books including publication like CRC Press, Taylor & Francis Group and Springer publication. He has published numerous research papers in international journals and conferences including Springer, IEEE, and Elsevier.

Multiple Fractal Keys Encryption for Audio

Deepak Negi
Associate Professor
Amrapali Group of Institutes, Haldwani, U.K. (India)
dr.deepaksinghnegi@gmail.com

Mahadev
Assistant Professor
Quantum University, Roorkee, U.K. (India)
mahadev.agra@gmail.com

Basudeo Singh Roohani
Assistant Professor
Quantum University, Roorkee, U.K. (India)
bsroohani2007@gmail.com

Ravindra Sharma
Assistant Professor
Swami Rama Himalayan University, Jolly Grant,
Dehradun, U.K. (India)
ravindrasharma97@gmail.com

Abstract.

The security of multimedia applications over communication networks is an important and challenging issue, as they can easily be intercepted. Providing privacy and preventing unauthorized access is a core function of multimedia encryption. The encryption of data is one way keeping content secret in the field of information security. In order to accomplish this objective, the content must be altered and made intelligible only to users of the secret content. This paper is unique in that it uses a dynamical systems and fractals approach to audio file encryption and decryption to achieve completely different results.

Keywords. Fractals, Mandelbrot set, Julia Set.

1. INTRODUCTION

Internet transactions are reliant on the security of information. Information can be shared across a network using the web, which is a widely popular and interactive medium [9]. Virtually, users are in touch with each other thanks to their multimedia gadgets, which have a positive impact on human life [12]. Initially, the security of electronic networks has been a concern for keeping information secret between the parties. Data integrity, entity reputation, and data authentication are required in order for two parties to communicate securely using an unsecure channel [11]. Participants remain in the dark about the data because the confidentiality guarantees their privacy. The integrity of data refers to the fact that there has been no alteration to it. Cryptography is the science of converting data into unreadable formats called cipher text, which may consist of symbols or a blend of alphabetical characters with symbols [7, 14]. Thus, one can prevent cybercriminals from

gaining access to important data. Algorithms for encrypting data comprise a combination of public keys and function structures. We can infer that securely communicating on the internet, engaging in online transactions, and using secured instant messaging require secure communication [3, 5]. Computing systems rely on cryptography to protect information. Security issues are extremely complex and important for networks. Additionally, not all users engage in legal activities, making crime at networks a widespread issue. Most are linked to financial crimes, and these cases have risen repeatedly in recent years [2, 6]. There is a major issue with wireless devices regarding the loss of confidentiality, content privacy, and location privacy. Moreover, to secure the information on the user, all the identifying information has to be encapsulated in cryptography to prevent it from being accessed by others. With hybrid encryption, we are solving the security issue associated with audio and multi-media formats such as video, text and images [6, 9 and 13].

2. PRELIMINARIES

1.1. Dynamical systems and fractals

The concept of fractals is a geometrical structure which has two major physical properties, namely self-similarity and dimension. In a dynamic system, changes take place over time. The dynamics of some systems are predictable whereas others are uncertain, i.e. chaotic [1, 4 and 7].

1.2. Mandelbrot set

The Mandelbrot set define by symbol M to denote the quadratic $Q_c(z) = z^n + c$ is describe as the group of all $c \in C$ where the orbit is bounded at the point 0, such as , $M = \left\{ c \in C : \left\{ Q_c^n(0) \right\}; \quad n = 0,1,2,3 \ldots is\, bounded \right\}$ [3, 12].

$M = \left\{ c \in C \left\{ Q_c^{\,n}(0) \, does\, not\, tends\, to \infty \, as\, n \to \infty \right\} \right\}$ we choose the initial point 0, as 0 is the only critical point of Q_c [2, 4].

1.3. Julia set

The Julia set is related to the set of points whose orbit is described by the function $Q_c(z)$. When the starting point is chosen as 0, since 0 is the most important critical point in iteration function $Q_c(z)$ [3, 5 and 8].

3. PROPOSED ALGORITHM FOR AUDIO FILES

Our discussion in this section centres on the encryption and decryption processes in sine wave form and their respective histograms. The histogram represents graphical information about data. The proposed audio encryption and decryption method implements fractal keys that are beneficial in keeping data secure [1, 14]. We used the common

"WAV" format file to test the scheme, since they are widely used as digital audio file formats on computers. In these formats, more information is provided about the data and higher quality audio is provided, as the waveform audio format has the sine wave, which is the simplest waveform, with only one frequency associated with it. Plain text of the audio file containing numeric data with the keys, superimposed on the matrix, yields the cipher text in mash format [3]. A picture of the original file in sine format can be seen in fig 1. As shown in fig 2, the encrypted sine file requires modifications after encryption, such that it is different from its original sine file format. The encrypted sine file after decryption yields its original sine file format, see fig 3.

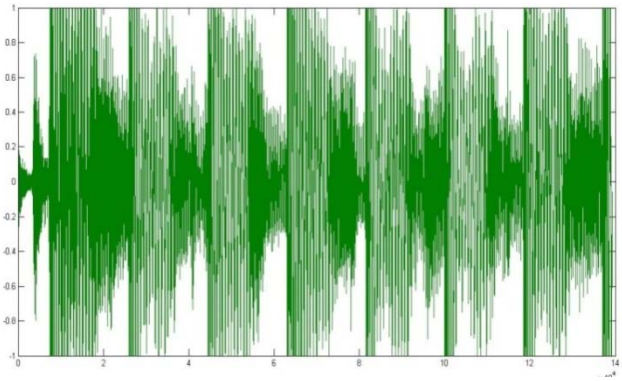
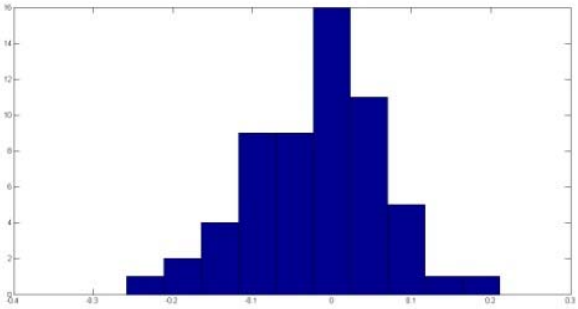


Fig1. Original audio file



Histogram original audio file

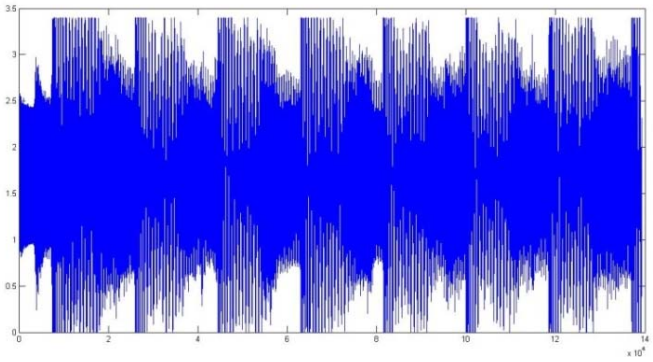
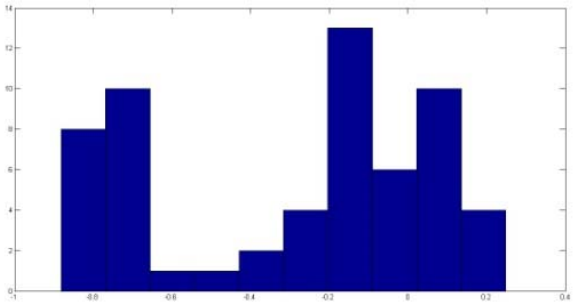


Fig2. After encryption



Histogram after encryption audio file

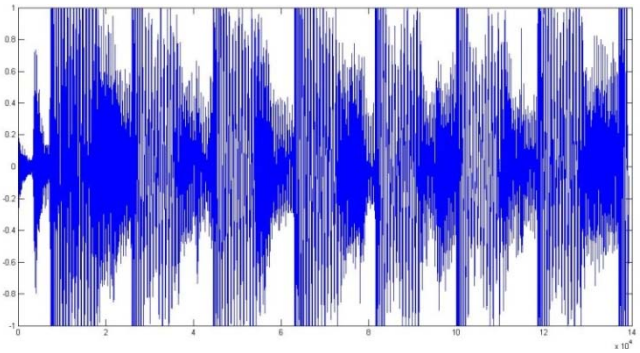
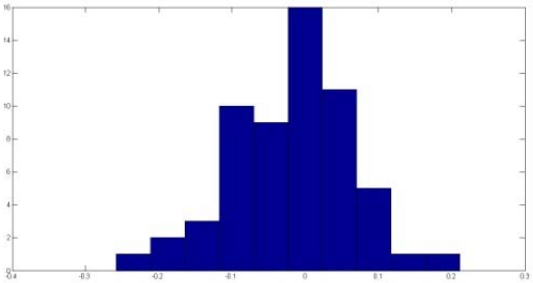


Fig3. Decrypted



Histogram decrypted audio file

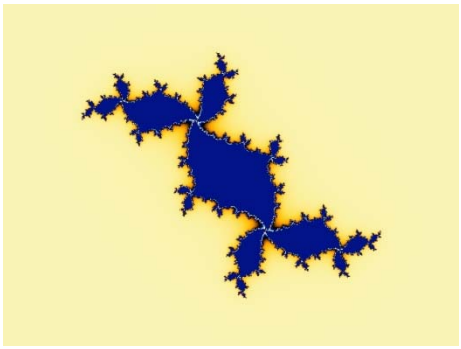


Fig4. Connected Julia set for(-0.16875 , 0.775i)

AT TRANSMITTER SIDE

First Step: From the connected Julia set transmitter chooses one public key c .

Second Step: To generate the three private keys d, d', d'' , a public key c is used along with a function $f(n)$.

Third Step: Read original audio ***wav*** file.

Fourth Step: Audio ***wav*** file is superimposed with private keys.

Fifth Step: Encrypted audio ***wav*** file is sent to the recipient.

AT RECIPIENT SIDE

First Step: An encrypted audio ***wav*** file is received by the receiver.

Second Step: To generate the three private keys e, e', e'', a public key c is used along with a function $f(n)$.

Third Step: A private keys e, e', e'' is used to decrypt an encrypted audio ***wav*** file

4. RESULTS AND DISCUSSIONS

4.1. Encryption and Decryption Process

There are connections between all Julia sets in the Mandelbrot set [1, 7 and 14], the proposed method begins by choosing a public key, c, and a connecting parameter over the connected Julia set corresponding to $f(n)$. A random Julia set is chosen. A fixed point along the iteration function $f(n)$ for each of the three private keys is used to initialize them; Finally, the audio wav file is encrypted by these three private keys and throw over the Network channel. Reversing the decryption process at the receiver end is necessary for the decryption to occur. We can generate the private keys by applying the public key and function to the received encrypted message and by applying these keys to the original audio data we can obtain the original audio data. A detailed description of the system works is given in Table 1.

Description of Process	Transmitter Side	Recipient Side
Select a complex public key c	-0.16875 , 0.775i	-0.16875 , 0.775i
Generate private keys d, d', d'' **along with function** $f(n)$	0.1002 - 0.0098i -0.1588 + 0.7730i -0.7411 + 0.5295i	0.1002 - 0.0098i -0.1588 + 0.7730i -0.7411 + 0.5295i e, e', e''

Superimpose audio message along with function	$f(n) = f_e(msg)$ 0.1002 - 0.0098i -0.1588 + 0.7730i -0.7411 + 0.5295i d, d', d''	$f(n) = f_e(msg)$ 0.1002 - 0.0098i -0.1588 + 0.7730i -0.7411 + 0.5295i e, e', e''
Encrypted message	0.2331 -0.8661 -0.1077 -0.3541 -0.7894 0.1861 -0.1002 -0.8271 -0.0794 -0.2916 -0.7765 0.0455 -0.3653 -1.0380 -0.0883 -0.1666 -0.6951 0.0689-0.5416 -1.0467 -0.0873 -0.3772 -1.0927 -0.3486 -0.6666 -1.3204 -0.5092 -0.6673 -1.0380 -0.0080 -0.2291 -0.9251 -0.2123 -0.4713 -0.9208 0.0406-0.2604 -1.0306 -0.3451 -0.5891 -1.0380 -0.1116 -0.4244 -1.1099 -0.2514-0.3676 -0.8896 -0.1800 -0.5729 -1.0888 -0.0717 -0.2538 -0.8817 -0.0862	0.2331 -0.8661 -0.1077 -0.3541 -0.7894 0.1861 -0.1002 -0.8271 -0.0794 -0.2916 -0.7765 0.0455 -0.3653 -1.0380 -0.0883 -0.1666 -0.6951 0.0689-0.5416 -1.0467 -0.0873 -0.3772 -1.0927 -0.3486 -0.6666 -1.3204 -0.5092 -0.6673 -1.0380 -0.0080 -0.2291 -0.9251 -0.2123 -0.4713 -0.9208 0.0406-0.2604 -1.0306 -0.3451 -0.5891 -1.0380 -0.1116 -0.4244 -1.1099 -0.2514-0.3676 -0.8896 -0.1800 -0.5729 -1.0888 -0.0717 -0.2538 -0.8817 -0.0862
Type of message	**Encrypted audio Message** $f(n) = f_e(msg)$	**Decrypted audio Message**

Table 1: Protocol for exchanging keys based on Fractals

5. CONCLUSION

In modern communication technologies, security is a crucial issue. Therefore, we must utilize a technique called cryptography to ensure the security of digital communications. As the Mandelbrot set contains infinite number of Julia sets, this paper provides a method that is efficient for encrypting and decrypting the audio "WAV" files. It is this property that allows researchers to use complex numbers and fixed point iterations to design efficient encryption/decryption codes. The Mandelbrot and Julia fractal sets are intrinsically connected to each other, so selecting a random key for each of these sets is possible. The level of security is, therefore, extremely high

6. REFERENCES

[1] D. Gulick, "CHAOS AND FRACTALS."

[2] D. Negi and A. Negi, "A behavior of Tricorns and Multicorns in N-Orbit," *International Journal of Applied Engineering Research,* vol. 11, pp. 675-680, 2016.

[3] D. Negi, A. Negi, and S. Agarwal, "The complex key cryptosystem," *International Journal of Applied Engineering Research, ISSN,* pp. 0973-4562, 2016.

[4] G. A. Edgar, Classics on fractals: CRC Press, 2019.

[5] G. Emmanuel, G. G. Hungilo, and Pranowo, "Numba acceleration of image steganography using Mendelbrot set fractals," in *AIP Conference Proceedings*, 2020, p. 030009.

[6] I. El Hanouti and H. El Fadili, "Security analysis of an audio data encryption scheme based on key chaining and DNA encoding," *Multimedia Tools and Applications,* vol. 80, pp. 12077-12099, 2021.

[7] M. Akhmet, M. O. Fen, and E. M. Alejaily, "Dynamics with fractals," *Discontinuity, Nonlinearity, and Complexity,* vol. 10, pp. 173-184, 2021.

[8] M. Khan, F. Masood, and A. Alghafis, "Secure image encryption scheme based on fractals key with Fibonacci series and discrete dynamical system," *Neural Computing and Applications,* vol. 32, pp. 11837-11857, 2020.

[9] M. T. GENÇOĞLU, "Enhancing The Data Security by using Audio Steganography with Taylor Series Cryptosystem," *Turkish Journal of Science and Technology,* vol. 16, pp. 47-64, 2021.

[10] S. Agarwal, "A fractal based image cipher using Knuth shuffle method and dynamic diffusion," *IJCNC,* vol. 11, pp. 81-100, 2019.

[11] S. Agarwal, "Preserving Information Security Using Fractal-Based Cryptosystem," in *Handbook of Research on Cyber Crime and Information Privacy*, ed: IGI Global, 2021, pp. 539-566.

[12] S. Banerjee, M. K. Hassan, S. Mukherjee, and A. Gowrisankar, *Fractal Patterns in Nonlinear Dynamics and Applications*: CRC Press, 2020.

[13] T. Mythili, A. Sofiabanu, R. Mohanasundari, and K. Sreekanth, "Data Hiding with Image and Audio Steganography Cryptosystem in Network," *International Journal of Recent Trends in Engineering & Research,* vol. 5, pp. 5-10, 2019.

[14] W. R. Smith, *Chaos, Fractals, and Dynamics*: CRC Press, 2020.

Biographies

Dr. Deepak Negi received his Bachelor of Science degree from Kumaun University Nainital, India in 2006. He received Masters in Computer Science degree from Uttarakhand Technical University, Dehradun India in 2009 and Doctorate in Computer Science and Engineering degree from Uttarakhand Technical University, Dehradun India in 2018. My area of research is application of fractal and dynamical systems, Network security, databases and data privacy.

Dr. Mahadev received his PhD degree in Computer Science from Gurukula Kangri Vishwavidyalaya, Haridwar in 2019. He received MCA degree from Dr. B. R Ambedkar University, Agra in 2001. His area of specialization is in web Technology and Machine Learning

Mr. Basudeo Singh Roohani received his bachelor degree in Computer Science. & Engineering from Bundelkhand Institute of Engineering & Technology (BIET) in 1999, Jhansi, Uttar Pradesh, India. He received a Master's degree in Computer Science and Engineering from Uttar Pradesh Technical University in 2006 Lucknow, Uttar Pradesh, India.. He is currently pursuing. Doctorate in Computer Science and Engineering degree from Dr. A.P.J. Abdul Kalam Technical University Lucknow,Uttar Pradesh, India. His research areas include Artificial Intelligence and Machine Learning.

Automotive Imbalance Dataset Analysis and Solution using Deep Learning Algorithm

M S Sunitha Patel[1], Srinath S[2]

[1]*Assistant Professor, Dept. of CSE, ATME College of Engineering, Mysuru, Karnataka, India, mssunithapatel@gmail.com*
[2]*Assosciate Professor, Dept. of CSE, SJCE, Mysuru, Karnataka, India srinath@sjce.ac.in*

Abstract.

A very basic stage and a very significant part of achieving effective results for the specified challenge is to train the machine learning algorithms with data. The primary skill in data science engineering is gathering raw data and building datasets, pre-processing and annotating for the given challenge. In machine learning-based problem-solving strategies, creating the correct dataset for automotive image processing takes around half the time. The varied weather conditions, road conditions, driving conditions, and other factors make automotive image processing more difficult. The collection of imbalanced datasets during image dataset construction is a common difficulty in deep learning models for multiclass analysis. The purpose of this study is to demonstrate the process of creating, testing, training, validating, and evaluating a dataset for three distinct vehicle classes: bus, car, and truck. The Convolutional Neural Networks model is used to evaluate performance and represented through Receiver Operating Characteristics (ROC) for balanced and unbalanced datasets (CNN). The ability of image augmentation to increase training performance on imbalanced datasets has been proved experimentally.

Keywords. Imbalance dataset, image data science, deep learning, CNN, Image Augmentation.

1. INTRODUCTION

The quality of data sets is determined by the problem statement, the product being utilized, the development time, and the development cost. To produce a fully autonomous car, both the dataset quality and the number of training images necessary must be exceedingly high. . Thus, for the given problem, it is vital to understand the type of data that will be utilised for training in order to avoid incurring extra development costs, lead times, and calculation time. Numerous automotive image processing systems require a fast response time [1] in order to drive the vehicle within milliseconds. Active safety systems such as Adaptive Cruise Control (ACC) [2], Lane Keeping Assistance (LKA) [3], and Collision Assistance (CA) [4], among others, require a very fast response time to avert catastrophic incidents. Another important consideration is that due to the necessity to research and train the Region Of Interest, the datasets required to train ACC, LKA, and CA may differ (ROI). The principal use of image processing in this situation may be to estimate the distance between two vehicles because the fundamental function of ACC is to manage the vehicle's speed and the spacing between vehicles ahead. If an unintended lane change occurs in LKA, the vehicle control system should inform the driver. So, in LKA, image data must be gathered that suits unintentional lane change circumstances such as sudden lane mark changes, lane mark fading, low visibility of lane markers owing to fog conditions, and so on. Several automotive image processing applications, such as Parking Assistance (PA) and Night Vision Camera Assistance (NVCA) [5], require a response time of less than one second. Due to the fact that the output of NVCA and PA systems is used by drives to perform the next control action, meeting the response time for people should be achievable. Because human beings govern the next action, the image's clarity is critical. NVCA places a greater emphasis on images recorded at night to aid drivers in the event of a predicted risk. It is self-evident that automotive image processing has more complex requirements in order to satisfy the diverse needs of end users. Automotive firms have a lot of opportunity to create cost-effective and trustworthy solutions with deep learning. The paper is organised into four pieces. The first

portion describes the rationale for the current work, followed by difficulties and solutions in automotive image datasets, as well as literature studies. The third piece covers the approach used to solve the problem, the fourth section explains the simulation results for different dataset sizes, and the last section explains the conclusion and possible future study.

2. BACKGROUND

The number of cameras in modern vehicles has expanded in order to improve driver safety and comfort. In sophisticated feature automobiles, the number of cameras used can be more than ten [6] [7], and modern high safety vehicles contain numerous safety and driver comfort features. As the number of cameras used grows, image processing and data analysis tools will face additional obstacles. Because cameras are mounted on a vehicle that is subjected to dynamic conditions such as vibration, dust, and rain [8,] the issues begin with picture acquisition. The acquired images will include a lot of noise and unnecessary data due to the dynamic situations. Following image acquisition, image enhancement [9] is used to improve image quality such as resolution, colour rendition, and so on, depending on the noise level in the collected images. After the image has been improved, image segmentation is a crucial stage in the dataset production process. The data science engineer must capture the right objects or features in the datasets to train the machine learning model. Image segmentation mainly deals with the required 'object detection and classification,' so the data science engineer must capture the right objects or features in the datasets to train the machine learning model. To have a safe driving car, for example, sophisticated vehicle driver assistance systems such as drive less vehicles must detect which area in the acquired image or video is safe to drive the vehicle by analysing environments [10]. The final step in image processing is'representation and description,' which is the required section of the image that can be used for further processing after object recognition and calcification. However, for automobile safety applications, a high level of representation and description is essential. Pixel is the lowest level of feature that provides required information in representation and description.

Vehicle type categorization is a classic example of automotive image processing that is required in several of the vehicle applications listed above. The first step in applying deep learning to classify vehicles is to gather the appropriate images and develop datasets. Datasets can be built from scratch or acquired from available sources. There are research papers available for vehicle classification using CNN [11],[12] in which the researchers created their own dataset; in a few cases [13],[14], the researchers created the dataset for the vehicle's rear part; and in another few cases [15], VeRi-776[16], the researchers used vehicle datasets available in open source PKU-VD [15], VeRi-776[16]. To gain access to open source datasets, as indicated previously, researchers must obtain permission from the owners. However, in order to obtain consent from the proprietors, certain time delays or even no response may be seen. Additionally, other vehicle image datasets are freely accessible to the public. The primary difficulty in creating automobile picture databases is obtaining the necessary number of photographs for various applications. For example, photos or videos of various sorts of road lanes, vehicle types, and environmental conditions (rain, snow, dusty, etc.) are required to train the deep learning algorithms for LKS. The major issues in the automotive image datasets are:

1. Availability of image datasets for all the required features to train the model.
2. Availability of required image quantity to train the model.
3. Availability of appropriate images for real-time vehicle driving scenarios
4. Availability of appropriate images for different environmental conditions

All of these factors combine to create a significant common challenge: handling imabalance datasets. Imbalance data set occurs in classification problems during performance analysis of multiple classes with the uneven quantity of datasets. To overcome issues seen during imbalanced datasets, image augmentation can be utilized. In this paper, multiclass vehicle classification like bus, car, and truck has been analyzed in three different scenarios of image

datasets. The main reason to take the 3 mentioned classes is that getting the required quantities of images for buses and trucks in open access is quite difficult.

3. METHODOLOGY

CNN architecture has been used to analyze the impact on different image dataset sizes and their performance on the machine learning algorithm. The methodology of dataset modeling, training, and validation for CNN architecture has been explained as follows.

3.1 MODELING DATASETS

Once the challenges inherent in automotive image processing are comprehended, image data collection for the specified task must begin. The primary issue with automobile image collections is a lack of required images in open-source. As previously stated, there are a few open sources accessible. However, for the chosen topic, which involves three distinct types of vehicles, while there are several vehicle datasets accessible, photos for bus and truck datasets were not available in sufficient quantities. As mentioned in table-1, car image datasets are gathered with 5000 images where has for bus and trucks managed to gather 2900 and 3400 images respectively from open source. To increase the image dataset size, a data augmentation approach has been utilized by using existing images. There are various data augmentation techniques are available like padding, cropping, flipping in different directions, etc. Depending on the problem and available images one can create additional images for the data sets. By using the image augmentation technique additional images for bus and truck has been created for 2100 and 1600 images respectively. Two examples of image augmentation have been shown in Figure 3.1 and Figure 3.2.

Tabel-1: Dataset Scenarios for Analysis

Dataset Scenarios	Number of Bus images	Number of Car images	Number of Truck images
1.Minimal Images	2900	2900	2900
2.Imbalance Images	2900	5000	3400
3. Added Augmented Images to make 5000 equal images	2100	0	1600

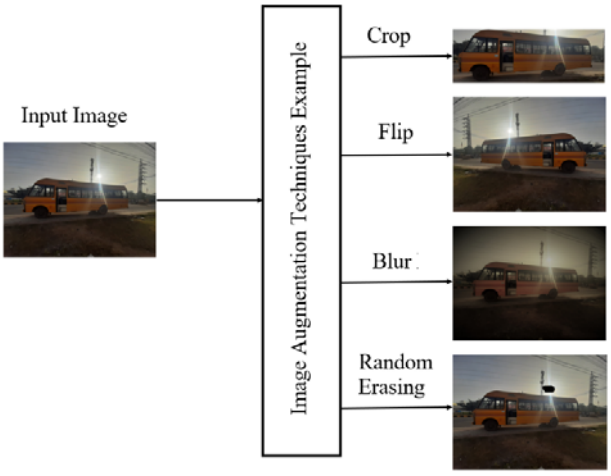


Figure 3.1 Image augmentation examples for bus image

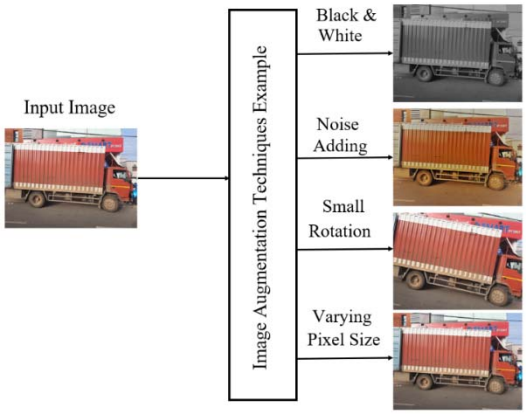


Figure 3.2 Image augmentation example for truck image

3.2 TRAINING, TESTING, AND VALIDATION DATASETS

As mentioned in Figure 3.3, total images in the datasets are segregated as training, testing, and validation. However, in the current scenario, only training and testing datasets in the ratio of 80% and 20% have been implemented. Sometimes validation datasets can also be used to fine-tune the learning algorithms.

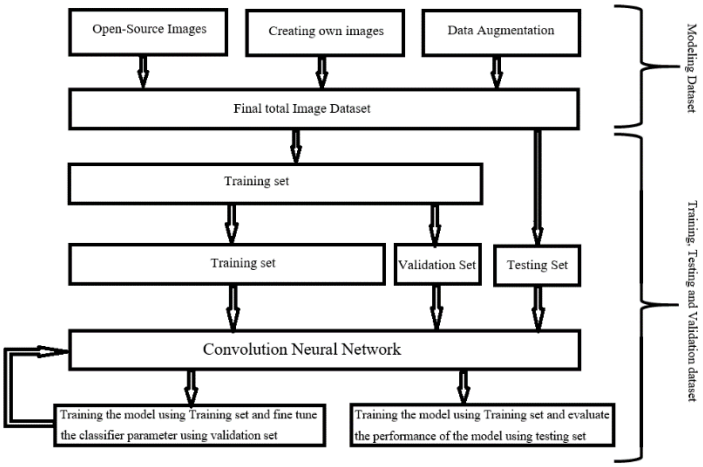


Figure 3.3 Testing and training for CNN

3.3 CNN ARCHITECTURE

A three-layer convolution neural network has been implemented as shown in Figure 3.4. An experiment has been carried out using Python 3.7.9, with IDE PyCharm and OpenCV Library. Three different datasets as mentioned in the table-1 have been created separately and three different inputs have been provided for the CNN model. Image size of 64 x 64 has been normalized and feature extraction including filtering, image segmentation, and image enhancement has been taken care of from the CNN model. CNN models consist of convolution layer, pooling layer, fully connected layer, and output layer. The first convolution layer consists of 32 filters and followed by 64 filters in the second and third layers. Padding has been considered to maintain input and output image size. Max pooling and Average pooling are widely used in the research article however max-pooling has been considered because of its benefits in capturing maximum value in a feature map. The output of max pooling with window size 2 and stride size 2 is fed to the size 2 of the fully connected layer. The fully connected layer is the deep layer that is used for classification. Finally, the fully connected layer passes its input to softmax with 2 layers for the final image classified output. In the current work, three different datasets are passed with fifty epochs for each scenario to train the learning algorithm. The output of each scenario is executed in three different phases and results have been captured and explained in the next section.

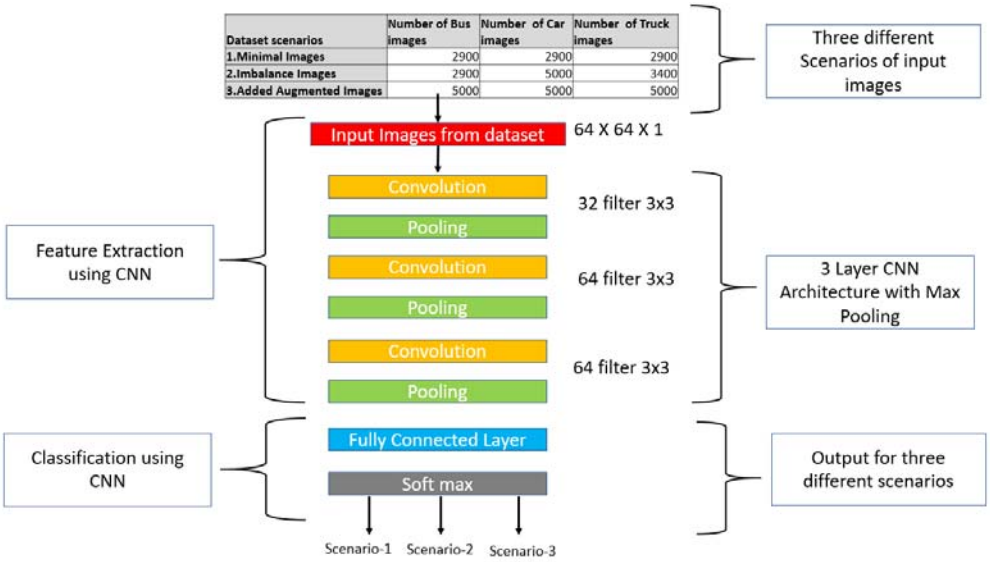


Figure 3.4 CNN Architecture

4. RESULT AND DISCUSSION

Performance evaluation of different dataset scenarios can be analyzed using accuracy as mentioned in equation 1. The result of the ROC curve is as shown in Figure 4.1, it is quite evident that an imbalanced dataset gives low performance compared to the balanced dataset. The main reason the performance of the imbalance dataset is deteriorated because of false negative and false positive will increase for bus and truck class. Also, it can be observed that after image augmentation the output ROC is increased from 88% to 98% scenario-1 to scenario-3.

$$Accuracy = \frac{(\text{True Positive} + \text{True Negative})}{(\text{True Positive} + \text{True Negative} + \text{False Positive} + \text{False Negative})} \quad (1)$$

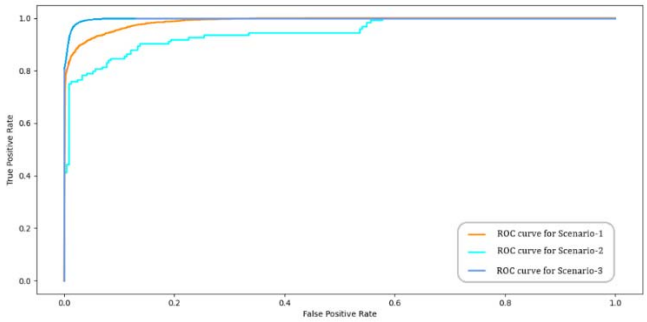


Figure 4.1 ROC curve for all three Scenarios

From Figure 4.2 as well it is quite evident that for the imbalance dataset error rate is more compared to the balanced dataset. After image augmentation, the error rate has reduced by nearly 10% compared to scenario-1 and scenario-3.

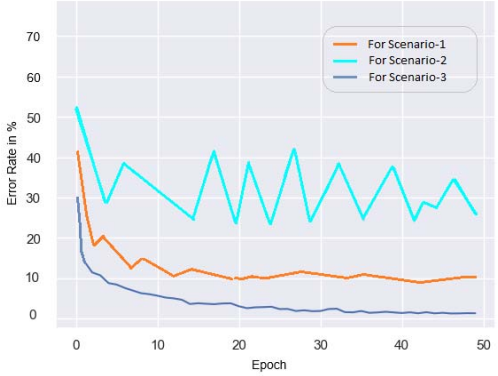


Figure 4.2 Performance graph for different scenario

5. FUTURE WORK

Adding images by utilizing the image augmentation technique has improved the performance of the CNN training model. The next part of our research work is to have a deep study on the different image augmentation techniques and their impact on performance. So future work will be concentrated on a detailed study of padding, cropping, flipping, rotation, etc, and its performance evaluation for different automotive functional conditions LKS, ACC, CA, etc.

6. REFERENCES

[1] U. Handmann, T. Kalinke, C. Tzomakas, M. Werner, and W.v. Seelen, “An Image Processing System for Driver Assistance,” in Proceedings of the IEEE Conference on Intelligent Vehicles, Stuttgart, Germany, 1999, pp. 481–486, IEEE

[2] M. M. Trivedi, T. Gandhi and J. McCall, "Looking-In and Looking-Out of a Vehicle: Computer-Vision-Based Enhanced Vehicle Safety," in IEEE Transactions on Intelligent Transportation Systems, vol. 8, no. 1, pp. 108-120, March 2007, DOI: 10.1109/TITS.2006.889442.

[3] Fei Han, Weiwen Deng, Sumin Zhang, Bei Ren and Ying Wang, AVision-based Forward Collision Warning System Developed under virtual Environment, SAE International Journal of TransportationSafety, Volume 4, Issue 2, July 2014, DOI: 10.4271/2014-01-0754.

[4] Shih-Shinh Huang, Chung-Jen Chen, Pei-Yung Hsiao and L. -C. Fu, "On-board vision system for lane recognition and front-vehicle detection to enhance driver's awareness," IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004, pp. 2456-2461 Vol.3, DOI: 10.1109/ROBOT.2004.1307429.

[5] J. McCall and M. M. Trivedi, “Human behavior based predictive brake assistance,” in Proc. IEEE Intelligent Vehicles Symphonium, Jun 2006.

[6] Rafael C. Gonzalez, Richard E Woods, Digital Image Processing, Third Edition, Dorling Kindersley (India) Pvt. Ltd,2014.

[7] M S Sunitha Patel and Dr.Srinath S, Soft Computing Approaches for Automotive Image Processing: Opportunities & Challenges, Advances in Automation, Signal Processing, Instrumentation, and Control, Scopus Indexed Springer Proceedings – Lecture Notes in Electrical Engineering 2020.

[8] Fei Han, Weiwen Deng, Sumin Zhang, Bei Ren and Ying Wang, AVision-based Forward Collision Warning System Developed under virtual Environment, SAE International Journal of TransportationSafety, Volume 2, Issue 2, July 2014

[9] Yunbo Rao, Leiting Chen, A Survey of Video Enhancement Techniques, Journal of Information Hiding and Multimedia Signal ProcessingVolume 3, Number 1, January 2012.

[10] J. McCall and M. M. Trivedi, Video based lane estimation and tracking for driver assistance: Survey, algorithms, and evaluation, IEEE Trans.Intell. Transp. Syst., vol. 7, no. 1, pp. 20–37, Mar. 2006.

[11] M. A. Hedeya, A. H. Eid, and R. F. Abdel-Kader, “A super learner ensemble of deep networks for vehicle-type classification,” IEEE Access, 2020.

[12] Abhishek N, S Gopalswamy, S Rathinam, “Vision based Techniques for Identifying Emergency Vehicles, SAE Technical Paper, 2019-01-0889, April 2019.

[13] Y. Lou, Y. Bai, J. Liu, S. Wang, and L. Duan, ``VERI-wild: A large dataset and a new method for vehicle re-identi_cation in the wild," inProc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2019, pp. 3235_3243.

[14] M. Naphade, S. Wang, D. C. Anastasiu, Z. Tang, M.-C. Chang, X. Yang, Y. Yao, L. Zheng, P. Chakraborty, C. E. Lopez, and. Sharma, ``The 5th AI city challenge," in Proc. IEEE/CVF Conf. Comput.Vis. Pattern Recognit. Workshops, Jun. 2021.

[15] Xinchen Liu, Wu Liu, Tao Mei, and Huadong Ma. A deep learning-based approach to progressive vehicle reidentification for urban surveillance. In Proc., ECCV, pages 869–884, 2016.

[16] Hongye Liu, Yonghong Tian, Yaowei Yang, Lu Pang, and Tiejun Huang. Deep relative distance learning: Tell the difference between similar vehicles. In Proc. CVPR, pages 2167–2175, 2016.

[17] https://www.kaggle.com/brsdincer/vehicle-detection-set

[18] https://data.mendeley.com/datasets/pwyyg8zmk5/2

[19] http://ai.stanford.edu/~jkrause/cars/car_dataset.html

[20] https://github.com/datacluster-labs/Indian-Vehicle-Image-Dataset

[21] https://www.gti.ssr.upm.es/data/Vehicle_database.html

[22] https://data.mendeley.com/datasets/pwyyg8zmk5/2

Biographies

M S Sunitha Patel received a Bachelor's Degree in computer engineering from Visvesvaraya Technological University, Belgaum in 2005, a Master's Degree from Visvesvaraya Technological University, Belgaum in 2010. She is currently working as an Assistant Professor at the Department of Computer Engineering, ATME College of Engineering, Mysuru, Karnataka. Her research areas include image processing, machine learning, data mining and deep learning.

Dr. Srinath.S is working as an Associate Professor, Department of Computer Science and Engineering, Sri Jayachamarajendra College of Engineering, Mysuru. He obtained his B.E Degree in Computer Science and Engineering from Siddaganga Institute of Technology in 1995. He later obtained his Masters in Software Engineering in the year 2002 and topped the university by securing first rank and gold medal from the esteemed Visvesvaraya Technological University, Belgaum. He obtained his Ph. D Degree in Computer Science for his thesis “Investigation and Development of Robust algorithms for recognizing and analyzing printed and hand embossed Kannada Braille characters: Novel approaches” from the University of Mysore in the year 2015. He has 25 years of rich experience in the field of teaching.

AN EFFICIENT APPROACH TO DETECT DEPRESSION THROUGH PREDICTIVE ANALYSIS

Sakshi Rastogi[1], Gaurav Kumar Srivastava[2], Sunil Kumar Vishwakarma[3]
Student[1], Assistant Professor [2, 3]
Department of Computer Science & Engineering[1-3]
Babu Banarasi Das University, Lucknow[1-3]
E-Mail ID- sakshirastogi.2607@gmail.com[1], gaurav18hit@bbdu.ac.in[2], sunilvishwakarma83@gmail.com[3]

Abstract.

Nowadays in this world when Machine is learning by algorithms and performing according to the inputs Python has its unique importance. Depression is said to be a feeling in which a person feels having a low mood and a state of strongly not liking someone/something. It has many effects on the body of the person. The symptom which is recognized as the core of depression is not feeling interested in the works or not feeling pleasure in the things that give joy to them earlier. The primary intention of this research is to carry out a comparison amongst the different Algorithms of Machine Learning based on accuracy, precision, sensitivity, F1 score, and Confusion Matrix to find which algorithm gives the best performance on depression data. The final aim of this research paper is to provide a model that will predict depression in the human body.

Keywords. Random Forest Classifier, Extra Trees Classifier, Multi-Layer Perceptron, Support Vector Classifier, F1 Score, Confusion Matrix, Predictive Analytics

1. INTRODUCTION

Depression is said to be a feeling in which a person feels having a low mood and a state of strongly not liking someone/something. The symptom which is recognized as the core of depression is not feeling interested in the works that give joy to them earlier. This can result in the person having a state of sadness, thinking difficulty, problems in paying attention. It can also lead to an increase or decrease in the diet and sleeping time of the person. In this condition person also experience the feeling of dejection, hopelessness and suicidal thoughts.

1.1 Symptoms of Depression

- Problem in Sleeping
- Loss of Interest
- Increment in Fatigue
- Emotions that are Uncontrollable
- Appetite of a Person Changes
- Weight of a Person Changes

2. LITERATURE SURVEY

This section is concerned with all the previous works that are done in this field. It presents a study that gives the comparison between different Algorithms of Machine Learning. They have used algorithms as- Logistic Regression, Random Forest, XG Boost, Support Vector Machine, Ada Boost, K-NN and Decision Tree. They have performed their research on the prediction of liver disease at an early stage. They have compared the above Machine Learning Algorithms on the basis of Accuracy, Precision, Recall, F1 Score, an area under curve and Specificity. They have collected their datasets from UCI Machine Learning Repository. There result in states that Random Forest Algorithm performs the best in terms of accuracy with 83.70%. Random Forest also performs well in the terms of other parameters too. So, they concluded Random Forest as the best algorithm that can be used in predicting Liver Disease. It describes about the machine learning techniques principles and he also described the use of them in the domains of real-world applications. They further describe the challenges and potential they need to perform in their research. On the basis of their goal, they shortly discussed how the methods of machine learning are being used in providing an appropriate way in solving the problems of the real world. The conclusion was that machine learning is built upon the data that is provided to the algorithms for learning purposes and the performance provided by them. It has the algorithms of Machine Learning for doing the predictions on anxiety, depression and stress in their paper. They have gathered their data by making the questionnaire related to their topic. This consists of the data of several cultures and communities which are employed and unemployed. They realize in their research that classes they made were imbalanced at the time when they start making confusion matrix. So they measure f1 score to identify the best accuracy model. They find that Random Forest Classifier as the best model. The conclusion was that the f1 score is the important aspect in finding the best accuracy model. It describes the use of various kinds of machine learning. It also merges the results of the analysis that comes from all the algorithms that were used for performing their research. Their main purpose was to increase the awareness of Machine Learning among the persons. Their conclusion presents that it is necessary for the Machine Learning model to continuously grasp from the past doing that come from countries that are developed, set up algorithms of machine learning mostly for the making enterprises in domestic areas and providing help of the economy in developing industry. It presents about the survey on how machine learning can be used for providing investigation on depression. The methods which they use in their systems are based on the method of detection via posts on social media, syntax and semantic analysis of the person's emotion in order to predict the depression levels of different age groups. Some have performed comparative research on four Algorithms of Machine Learning. For the purpose of reducing attributes, they used CFSSubsetEval. They have collected their datasets from OASIS-Brains.org. They finally concluded as J48 is the best algorithm for the purpose of detecting Dementia.Some have conducted their research on various Algorithms of Machine Learning with the aim of finding the effectiveness. The datasets that were used in this research were from different types of clinics. In these datasets are small, medium and large that can be accessed publicly. The comparison between algorithms was done on the basis of the requirement of accuracy and time in training and testing algorithms. The result implies that K-Nearest Neighbor performed well amongst all the algorithms used.It also presents that social network data gives the opportunity to work on the user's moods and attitudes when they convey messages with the use of social media. The data for the analysis was on the

Facebook data that they collected from an online public source. They gave their analysis on 7146 Comments on Facebook. They got the conclusion as 54.77% depressive person who conveys between mid-night to mid-day & 45.22% depressive person who conveys between mid-days to mid-night. Some have done comparative research amongst some popular Algorithms of Machine Learning. They have used two datasets in order to provide the best efficiency. They have collected all the information including datasets from the UCI Machine Learning storehouse. There first dataset contains 6500 rows & 13 columns & second dataset contains 1055 rows & 13 columns. There result shows that Support Vector Machine performs the best accuracy of 99.38%.

3. PROPOSED METHODOLOGY

The model that we have proposed for this research is described in the Figure 1

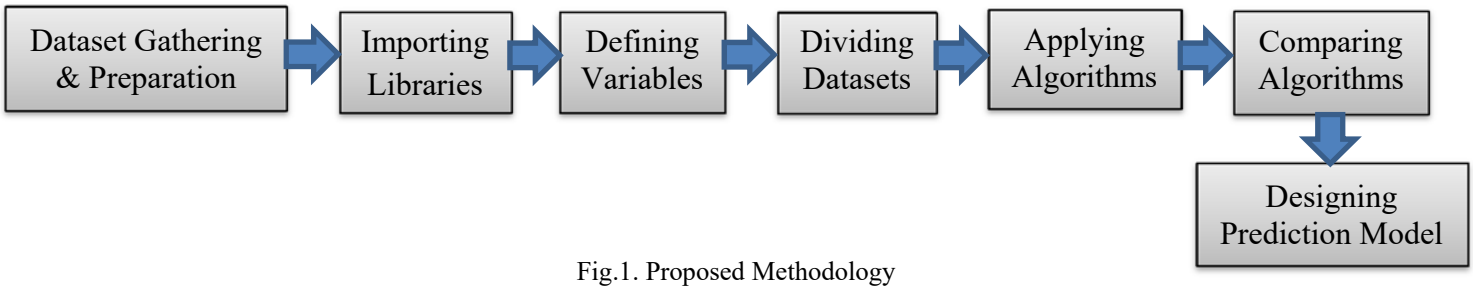


Fig.1. Proposed Methodology

3.1. Dataset Gathering & Preparation

In developing a Machine Learning Model or performing any type of research in Machine Learning the first and most important step is to gather the datasets. We have taken the Depression datasets from an online portal. There were 1290 rows and 27 columns in the original dataset. Table 1 describes the original Depression Dataset.

Table 1. The Original Dataset

1.	Timestamp	Age		obs_ consequence	Comment
2.	8/27/2014 11:29	37	-----	No	NA
3.	8/27/2014 11:29	44	-----	No	NA
⋮	⋮	⋮		⋮	⋮
1259.	11/30/2015 21:25	46	-----	No	NA
1260.	2/1/2016 23:04	25	-----	No	NA

Table 2 describes the final dataset that we have used for performing this research. In the column 'Gen', 0 represents Female and 1 represents Male and in all the other columns 0 stands for No and 1 stands for Yes.

Table 2. The Final Dataset

1.	Age	Gen		phys_health_ consequence	Target
2.	37	0	--------	0	1
3.	44	1	--------	0	1
⋮	⋮	⋮		⋮	⋮
1248.	32	1	--------	0	0
1249.	36	1	--------	0	0
⋮	⋮	⋮		⋮	⋮
1259.	46	0	--------	0	1
1260.	25	1	--------	0	1

3.2. Importing Libraries

Python has plenty of libraries that can be used by the developers for performing different kinds of research like on Machine Learning Models, Robotics etc. All the libraries are predefined and are easily available we just have to import them and use them according to our requirement. We have done the same we have imported the libraries like pandas, matplotlib, sklearn etc. We have used the following libraries-

1. Pandas- We used this library for reading the datasets.
2. Matplotlib, Seaborn- We have used the seaborn library for plotting heat map and matplotlib library for plotting the labels, giving the titles etc. around the heat map.
3. Numpy- We have used this library for designing the prediction model.
4. Sklearn- We have used many modules of this library in this research paper. We have used the modules like model_selection, ensemble, neural_network etc.

3.3. Defining Variables

Defining the variables that are included in the creation of the Model is also the most important step as the Model fits only on variables. For performing this we have to follow this pseudo code-

Step 1. Define the variable X by dropping target column and storing all other columns in it.
Step 2. Put the axis=1 for dropping target column.
Step 3. Define the variable y by storing target variable in it.

3.4. Dividing Datasets

It is necessary to divide the datasets into Training Module and Testing Module so that we can perform the works on Training Module and make predictions on Testing Module. We have divided the whole datasets into Training and Testing Module. We have divided the dataset into 80-20 ratio that means our Training Module contains 80% and Testing Module contains 20% respectively. For performing this we have to certain pseudo code the code is as follows-

Step 1. Define the ratio in which you want to divide the dataset.
Step 2. Define the Test Size according to the desired ratio.
Step 3. Apply the train_test_split module.

3.5. Applying Algorithms

The main step in this research is to apply algorithms. We have applied all Machine Learning Algorithms like Random Forest Classifier, Extra Trees Classifier, Ada Boost Classifier, Decision Tree Classifier and Multi-Layer Perceptron. For performing this we have defined the pseudo-code that is given as follows-

Step 1. Store each algorithm into different variables.
Step 2. Fit each model on the training dataset using the fit method.
Step 3. Use predict method for doing the predictions.

3.6. Comparing Algorithms

The pre-final step in this research is to compare the above-mentioned models. We have compared all the algorithms on the basis of accuracy, precision, sensitivity, F1 score and Confusion Matrix to find the best algorithm amongst all the models. For doing the work the pseudo code is as follows-

Step 1. Use accuracy score for finding accuracy.
Step 2. Use precision score for finding precision.
Step 3. Use recall score for finding sensitivity.
Step 4. Use f1 score for finding F1 Score.
Step 5. Print the results of comparison.

The another parameter that we are giving for comparison is the Confusion Matrix. Confusion Matrix is divided into two rows and two columns. It is plotted using matplotlib and seaborn libraries. We have done this in Three Steps. We have plot the Confusion Matrix of 2 Algorithms at a time and followed by the other Algorithms. Pseudo that is applicable for doing this is as follows-

Step 1. Use confusion matrix for plotting confusion matrix.
Step 2. Use the heatmap for doing the same.
Step 3. Print thr results.

3.7. Designing Prediction Model

The final step in the study is to evaluate a model that tries to predict depression in the human body. This prediction model works on the dataset that is provided to the model. This model has been made with the help of library named as numpy. The input contains the data in the same manner that is defined in the dataset. It implies that the 1st value is the Age of the Person, 2nd is the Gender in the form of 0 and 1, 3rd defines that if the person is having any Family History of Depression?, 4th is asking that if the person is working in any tech company?, 5th is asking that if the person is going through any treatment?, 6th is asking if the person is working remotely? 7th is asking if the person has mental health consequence? 8th is asking if the person has physical health consequence? and finally 9th is the target variable that predicts the depression in human body. The pseudo code for this is shown as follows-

Step 1. Use numpy for designing model.
Step 2. Input the values in the form of array.
Step 3. Reshape the array which was given as input.
Step 4. Use the results of comparing algorithms.
Step 5. Print the results either in the form of 0 or 1 i.e., 0 means Patient doesn't have Depression and 1 implies Patient have Depression.

4. RESULTS & DISCUSSIONS

The application of all above mentioned machine learning methods with all parameters- accuracy, precision, sensitivity, F1 score & Confusion Matrix is described in Table 3. In the case of Confusion Matrix, the first value represents False Negative value, second

represents False Positive value, third represents True Negative value & fourth represents True Positive value respectively. From the below table, it has been clearly shown that Ada Boost Classifier and Multi-Layer Perceptron both have shown the good performance in some parameters but Support Vector Classifier has shown the best performance in all the parameters. In the case of Confusion Matrix, Multi-Layer Perceptron has performed good as it predicts 78 True Positives but Support Vector Classifier has again performed the best as it has given the total 88 True Positives. So, it has been cleared that Support Vector Classifier is the best algorithm amongst all the presented algorithms so we have used the same algorithm for performing the final step of research i.e., designing of the Prediction Model. According to dataset which was given the prediction model performs effectively as it can clearly predict depression in human body.

Table 3. Result Analysis

Sr. No.	Model	Accuracy	Precision	Sensitivity	F1 Score
0	Random Forest	0.480159	0.488722	0.507812	0.498084
1	Multi-Layer Perceptron	0.503968	0.506224	0.953125	0.661247
2	Extra Trees Classifier	0.476190	0.482143	0.421875	0.450000
3	Ada Boost Classifier	0.523810	0.533333	0.500000	0.516129
4	Decision Tree Classifier	0.476190	0.482456	0.429688	0.454545
5	Support Vector Classifier	0.551587	0.549669	0.648438	0.594982

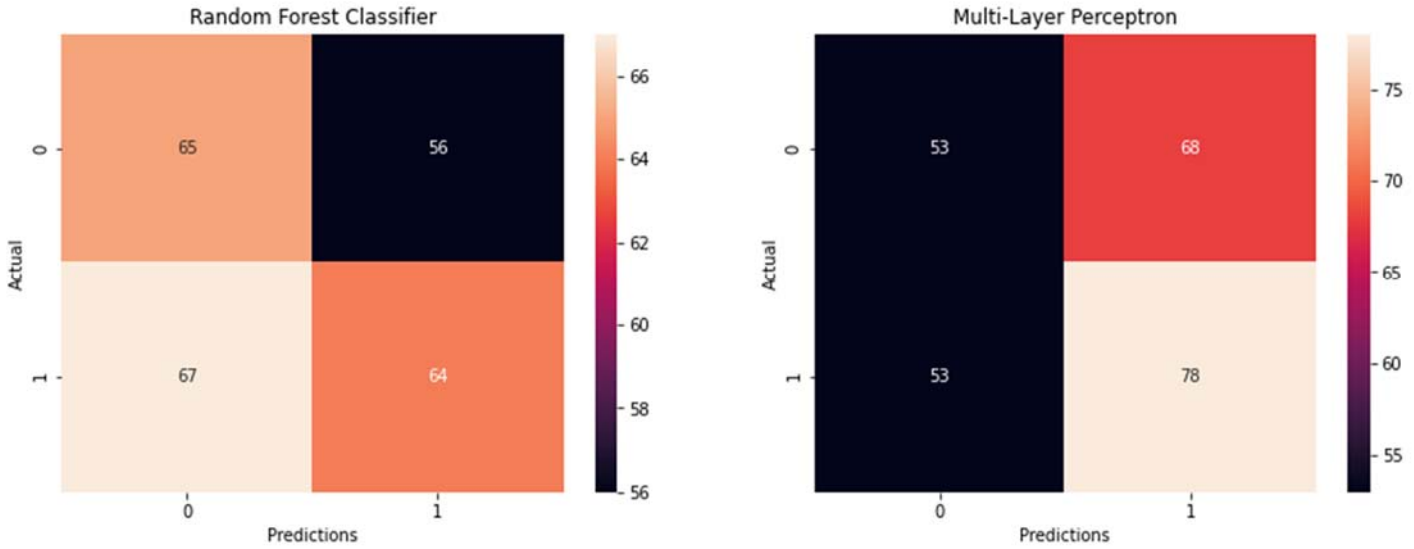


Fig.3. Confusion Matrix of Random Forest Classifier and Multi-Layer Perceptron

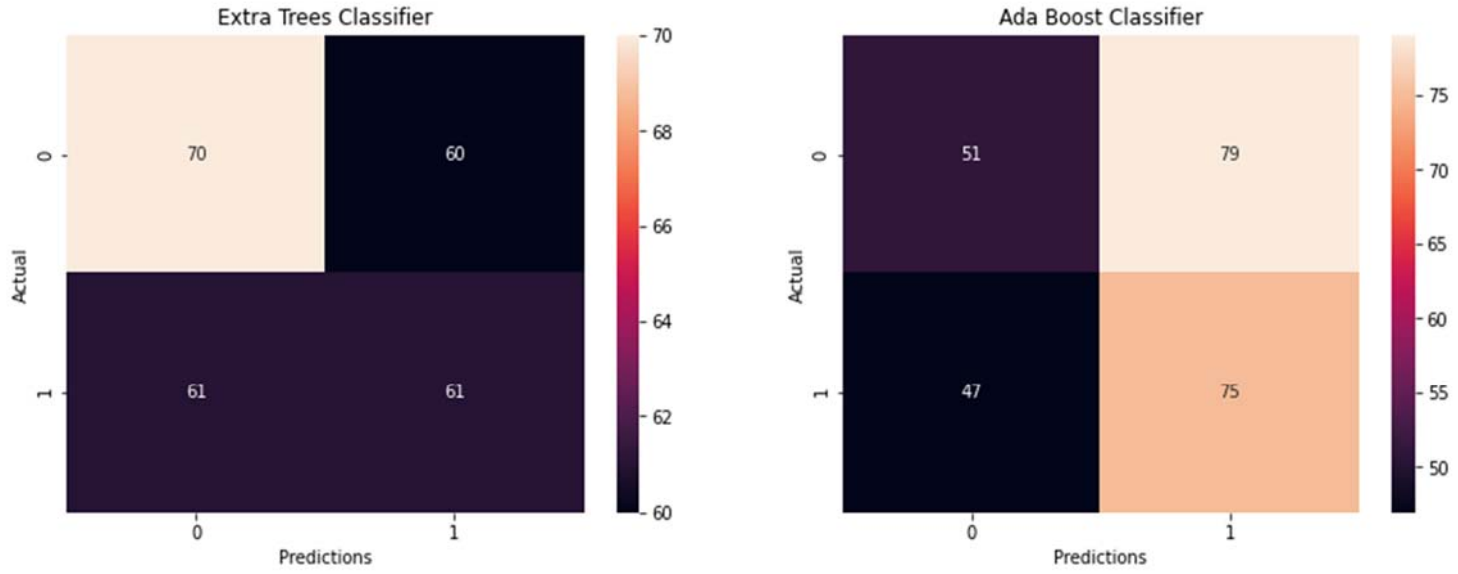


Fig.4. Confusion Matrix of Extra Trees Classifier and Ada Boost Classifier

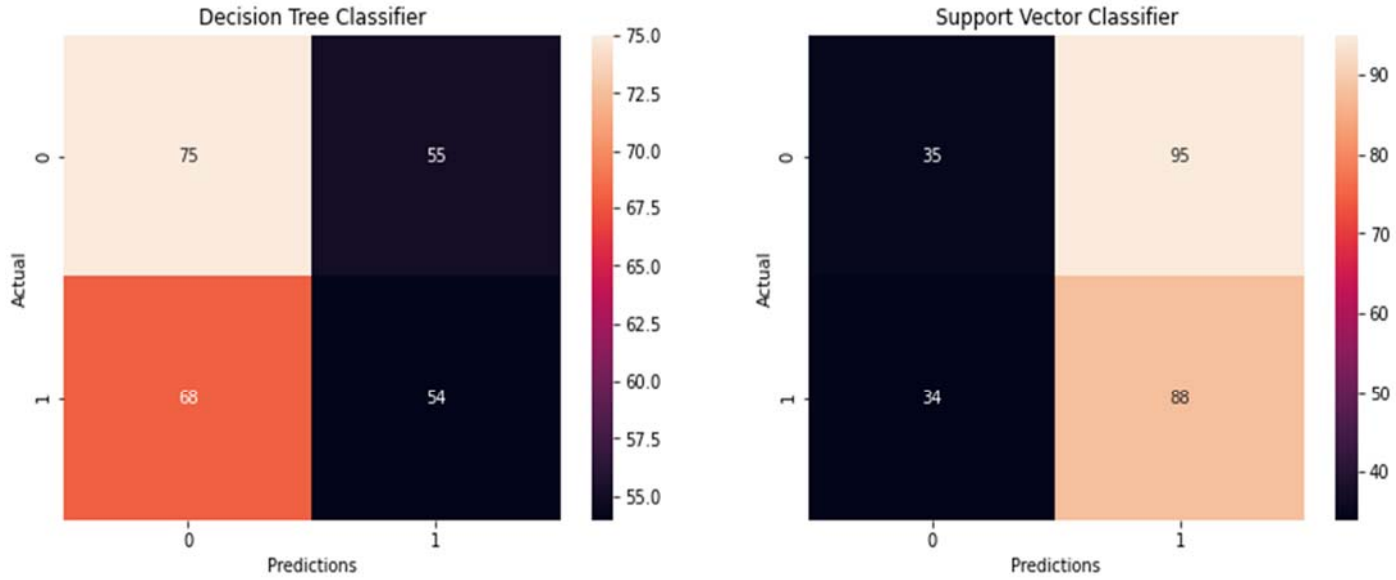


Fig.5. Confusion Matrix of Decision Tree Classifier and Support Vector Classifier

5. CONCLUSION

This paper has shown a comparative study between multiple Machine Learning Algorithms very well. The dataset used in this research is completely authenticated and unique as the proper analysis and editing is done on the dataset. According to the results provided by this research the Support Vector Classifier has performed really well in all the parameters. So, this algorithm can be used for future usage.

This paper has also shown a Model that will predict Depression in the human body according to the given dataset. The model predicts in the form of 0 and 1 i.e., 0 implies that the person is not having Depression and 1 implies that the person is having Depression. In this research we have used the results of the comparative study that is done as the primary objective of this research.

6. FUTURE SCOPE

In future this research will be helpful in the following aspects which are given as follows-

- It will be helpful for researchers to perform study in Predictive Analytics
- It will be helpful in selecting the best Machine Learning Model if the aim is to determine Depression at an initial stage
- It will be helpful in designing the Prediction Model using Machine Learning
- It will be helpful in predicting Depression from the Human Body

7. REFERENCES

[1] M. Ghosh *et al.*, "A comparative analysis of machine learning algorithms to predict liver disease," *Intelligent Automation and Soft Computing*, vol. 30, no. 3, pp. 917–928, 2021, doi: 10.32604/iasc.2021.017989.

[2] I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," *SN Computer Science*, vol. 2, no. 3, May 2021, doi: 10.1007/s42979-021-00592-x.

[3] A. Priya, S. Garg, and N. P. Tigga, "Predicting Anxiety, Depression and Stress in Modern Life using Machine Learning Algorithms," in *Procedia Computer Science*, 2020, vol. 167, pp. 1258–1267. doi: 10.1016/j.procs.2020.03.442.

[4] W. Jin, "Research on Machine Learning and Its Algorithms and Development," in *Journal of Physics: Conference Series*, Jun. 2020, vol. 1544, no. 1. doi: 10.1088/1742-6596/1544/1/012003.

[5] G. Srivastava, S. Kumar, H. Pandey, G. Kumar Srivastava, S. Kumar, and H. Pandey, "Modelling of an offline and online software for normalization of microarray data of gene expression by Perl, Bioperl and PerlTk and Perl-CGI," 2019.

[6] Vishwakarma Sunil Kumar, Sharma Birendra Kumar, and Abbas Syed Qamar, "Digital Watermarking for Image Authentication using Spatial-Scale Domain based Techniques," *International Journal of Recent Technology and Engineering*, vol. 8, no. 4, pp. 2334–2341, Nov. 2019, doi: 10.35940/ijrte.d8215.118419.

[7] D. Bansal, R. Chhikara, K. Khanna, and P. Gupta, "Comparative Analysis of Various Machine Learning Algorithms for Detecting Dementia," in *Procedia Computer Science*, 2018, vol. 132, pp. 1497–1502. doi: 10.1016/j.procs.2018.05.102.

[8] M. Diwakar, P. Singh, and A. Shankar, "Multi-modal medical image fusion framework using co-occurrence filter and local extrema in NSST domain," *Biomedical Signal Processing and Control*, vol. 68, Jul. 2021, doi: 10.1016/j.bspc.2021.102788.

[9] Md. R. Islam, M. A. Kabir, A. Ahmed, A. R. M. Kamal, H. Wang, and A. Ulhaq, "Depression detection from social network data using machine learning techniques," *Health Information Science and Systems*, vol. 6, no. 1, Dec. 2018, doi: 10.1007/s13755-018-0046-0.

[10] K. Sethi, A. Gupta, G. Gupta, and V. Jaiswal, “Comparative Analysis of Machine Learning Algorithms on Different Datasets,” 2019. [Online]. Available: www.ccsarchive.org

[11] A. Chakraborty, M. Jindal, M. R. Khosravi, P. Singh, A. Shankar, and M. Diwakar, “A Secure IoT-Based Cloud Platform Selection Using Entropy Distance Approach and Fuzzy Set Theory,” *Wireless Communications and Mobile Computing*, vol. 2021, 2021, doi: 10.1155/2021/6697467.

[12] P. Singh, M. Diwakar, X. Cheng, and A. Shankar, “A new wavelet-based multi-focus image fusion technique using method noise and anisotropic diffusion for real-time surveillance application,” *Journal of Real-Time Image Processing*, vol. 18, no. 4, pp. 1051–1068, Aug. 2021, doi: 10.1007/s11554-021-01125-8.

Biographies

Sakshi Rastogi, currently pursuing Master of Technology from Babu Banarasi Das University, Lucknow. I have done Bachelor of Technology from Invertis University, Bareilly. My research areas are Software Testing and Machine Learning.

Dr. Gaurav Kumar Srivastava, currently working as Assistant Professor in the Department of Computer Science and Engineering at Babu Banarasi Das University, Lucknow, 226028, India. He has completed his Bachelor of Technology in Computer Science and Engineering from Dr. A.P.J. Abdul Kalam Technical University, Lucknow, completed his Master of Technology in Computer Science and Engineering from Babu Banarasi Das University, Lucknow and completed Ph.D. in Computer Science and Engineering from Maharishi University of Information Technology, Lucknow. He has published various research articles in International Peer reviewed Journals/ Conferences.

Sunil Kumar Vishwakarma, currently working as Assistant Professor in the Department of Computer Science and Engineering at Babu Banarasi Das University, Lucknow, 226028, India. His research areas are Computer Vision and Image Processing. He has completed his Master of Technology from IET Lucknow and Bachelor of Technology from BBDIET & RC Bulandshahr.

Artificial Intelligence in MOOCs:

A Bibliometric Perspective

Ishteyaaq Ahmad[1]

Sonal Sharma[2]

Gulsun Kurubacak[3]

Ajay Kumar[4]

Rajiv Kumar[5]

Sikha Ahmad[6]

[1, 2, 5] *Uttaranchal Institute of Management, Uttaranchal University, Dehradun, India, ishteyaaq@gmail.com*[1], *sonal_horizon@rediffmail.com*[2], *rajiv.gill1@gmail.com*[5]

[3]*Anadolu University College of Open Education, Turkey,gkurubac@anadolu.edu.tr*

[4]*Bharati Vidyapeeth (Deemed to be University) Institute of Management and Research,New Delhi, ajay.kumar@bharatividyapeeth.edu*

[6]*Doon University, Dehradun, India, sikhaahmad@gmail.com*

Abstract.

Artificial intelligence has come a long way in recent years, and it currently represents an emerging technology that will change the way people live. Despite the fact that many instructors are unaware of its scope and, more significantly, what it implies, the use of AI in education is growing rapidly, with the aim of improving the quality of teaching and learning. There are many different approaches to this, but they all involve the use of data analytics and machine learning to make improvements in education. The purpose of this research was to look at the scientific output on artificial intelligence in MOOCs that were indexed in Scopus databases between 2012 and 2021. For this a customised approach for bibliometric investigations was applied on the SCOPUS databases. A total of 476 articles were included in the sample. The findings revealed that the issue has a lot of interest, but that the literature on it is still in its early stages.

Keywords. Intelligent systems; artificial intelligence; machine learning; MOOCs; bibliometric study

1. INTRODUCTION

From time immemorial human beings have been learning and adapting to the new environment. This movement has continued for millions of years in all dimensions of human life but picked up incredible momentum in the modern era. The advent of information technology amplified this learning and adapting process by leaps and bounds [1], [2], [3], [4]. The world in which we are living is changing at such a fast and unpredictable rate and it is difficult to visualize how different things were just a few years back. Artificial Intelligence (AI) has gone a great way in recent years, and it is now regarded as a disruptive technology that will affect people's lives in the not-too-distant future. Various studies have already examined the primary applications of artificial intelligence in education, as well as the fundamental design difficulties and methodologies used in the system's adaptation to the user [5]. Traditionally, intelligent educational systems (IESs) were considered in the context of their functional components. Domain knowledge, a representation of the learner's current state, and instructional knowledge are the three mandatory subsystems [6]. There are many different ways in which AI can be used in education. One way is through the use of data analytics to improve teaching and learning practices. Here systems can involve machine learning, where algorithms can be trained on large sets of data to learn from them and then make predictions or recommendations.

Massive open online courses (MOOCs), which have been popular for a few years, are one of the fields of education where a large number of learners are enrolled. These courses can be taken at any time, on any device and are a cheaper and more flexible way to learn. The best thing about MOOCs is that people can choose from a wide range of courses and study at their own pace, allowing them to complete the course faster or take it slower. As artificial intelligence is also growing rapidly in education where its primary goal is to improve the quality of teaching and learning by making it more accessible, personalized, adaptive, and interactive. Because of enormous MOOCs enrolments, AI is likely to play a significant role in analysing the large data set of MOOCs. Data analytics may be used by AI to improve teaching and learning techniques. The machine learning algorithms can be trained on large datasets of MOOCs to learn from them and then make predictions or recommendations about how to teach better or learn something new.

A search of the scientific literature on artificial intelligence in MOOCs can be useful in this scenario to see how many research papers are available. As a result, this bibliometric analysis looks at the scholarly output on artificial intelligence in MOOCs that was indexed in Scopus between 2012 and 2021, and would address the following research questions:

RQ1: What is the status of paper publication in the last ten years (2012 – 2021)?

RQ2: What are the major organisations and countries that produce the most artificial intelligence research within MOOCs?

RQ3: What are the most-cited publications and popular keywords for research in artificial intelligence output in MOOCs?

2. METHODOLOGY

In this research paper, we employed a bibliometric study approach that adhered to the PRISMA pronouncement's requirements. The metadata of scientific work on the Scopus database in the last ten years (2012–2021) was analysed for this. For this purpose, the search was broadened to include the popular subsets of artificial intelligence viz., "machine learning" and "intelligent systems". The Scopus database was searched using the keywords ("Artificial Intelligence" OR "Machine Learning" OR "Intelligent System") AND "MOOC". The results were limited to publication years (2012–2021). The data analysis was carried out on 476 research papers that were indexed in Scopus. Software, such as Microsoft Excel, VOSviewer version 1.6.18 were used for investigations.

3. RESULTS AND DISCUSSIONS

3.1. Status of paper publication in the last ten years

Taking into account the "year of production" variable, the findings revealed that the largest production peaks in the Scopus database in the year 2020 with 91 publications, while the years 2021 and 2019 each had 89 and 85 publications, respectively. Figure 3.1 depicts the year-by-year release of materials in MOOCs on the theme of artificial intelligence from 2012 to 2021. The findings indicate that there is a growing interest in the subject, since the number of publications has increased in the last ten years.

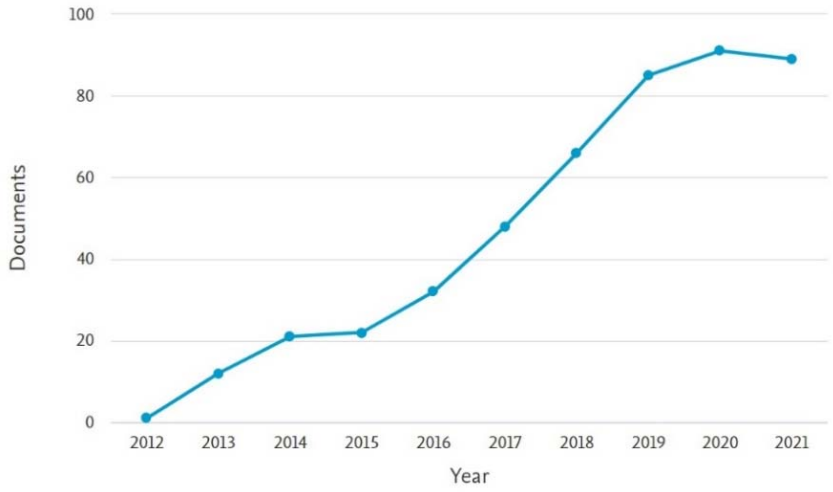


Figure 3.1. Year of production of published documents

In terms of "document type" conference papers accounted for 71.7 percent of overall output in the Scopus database, followed by journal articles, which accounted for 20.3 percent of total production. Table 1 depicts the document type of research publication in MOOCs on the theme of artificial intelligence from 2012 to 2021.

Comparing the documents by source for the counts of at least 10, the series “Lecture Notes in Computer Science” has the highest number of publications in the last ten years with 72

documents, followed by "ACM International Conference Proceeding Series" with 31 documents. The complete result is shown in Figure 3.2.

Table1. Documents by type

Document type	Documents	Percentage
Conference Paper	335	71.7
Journal Article	95	20.3
Conference Review	25	5.4
Review	5	1.1
Book Chapter	4	0.9
Book	1	0.2
Erratum	1	0.2
Short Survey	1	0.2

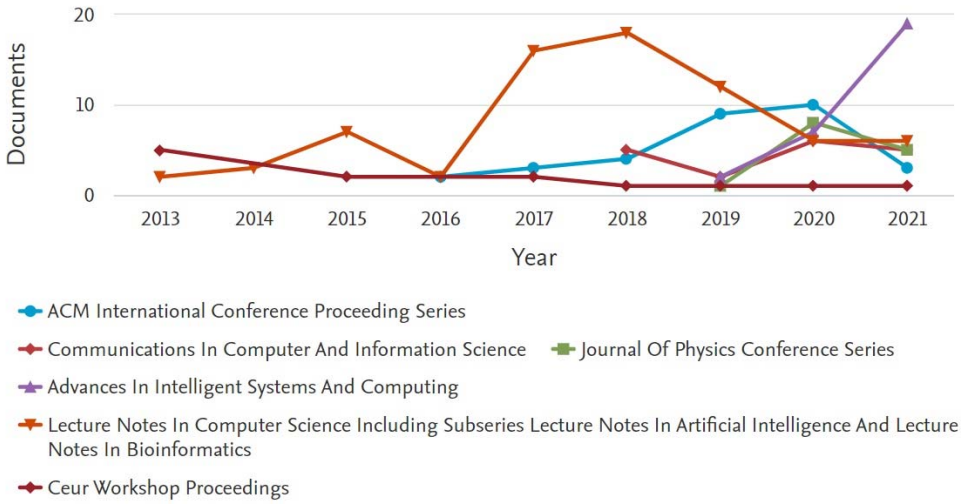


Figure 3.2. Published documents by source

Analysing the documents by subject area, the research found that 44.3% of the documents were related to computer science, 14.2% were related to mathematics, and 13.7% were related to social sciences. The complete subject analysis is shown in Table 2.

3.2. Major organisations and countries that produce the most research

According to the Scopus database, the top institution with research publications on artificial intelligence in MOOCs, in the previous ten years, from 2012 to 2021, was Carnegie Mellon University (USA), followed by Universidad Autónoma de Madrid

(Spain), and Norges Teknisk-Naturvitenskapelige Universitet (Norway). Table 3 summarises the top 10 results for this world-wide search.

Table 2. Documents by Subject Area

Subject area	Documents	Percentage
Computer Science	391	44.3
Mathematics	125	14.2
Social Sciences	121	13.7
Engineering	115	13.0
Decision Sciences	37	4.2
Physics and Astronomy	21	2.4
Business, Management and Accounting	15	1.7
Materials Science	9	1.0
Psychology	9	1.0
Medicine	8	0.9
Others	32	3.6

Table 3. Documents by Affiliation

Organizations	n	%
Carnegie Mellon University (USA)	11	2.31
Universidad Autónoma de Madrid (Spain)	8	1.68
Norges Teknisk-Naturvitenskapelige Universitet (Norway)	7	1.52
Universidad de Valladolid (Spain)	7	1.52
Universidad Carlos III de Madrid (Spain)	6	1.26
Huazhong Normal University (China)	6	1.26
Ecole Mohammadia d'Ingenieurs, Mohammed V University (Morocco)	6	1.26
Stanford University (USA)	6	1.26
Durham University (England)	6	1.26
Massachusetts Institute of Technology (USA)	6	1.26

In terms of "countries", according to the Scopus database, China has most publications on artificial intelligence research on MOOCs, with 110, followed by the United States with 83, India with 37, and Spain with 34 research articles. Figure 3.3 depicts the top ten countries of the entire search outcome.

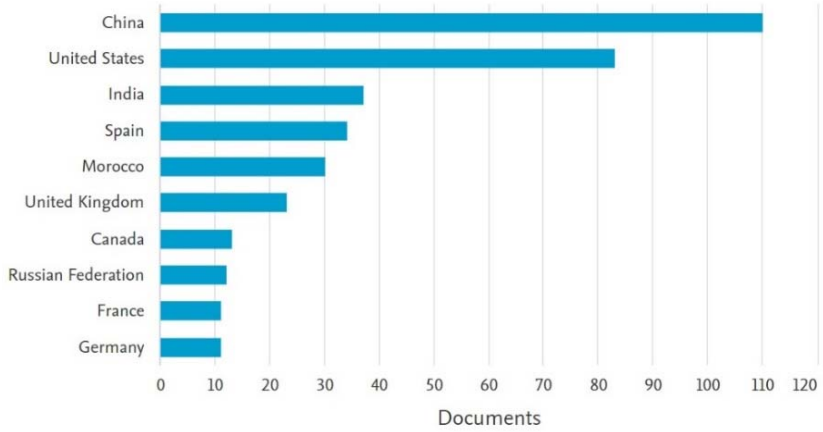


Figure 3.3. Published documents by country

3.3. Most-cited publications and popular keywords for research in artificial intelligence output in MOOCs

Finally, data on the variable "the most-cited publications" was gathered. This was picked as a criterion for highlighting the most significant research. Table 4 summarises the top five most referenced publications in Scopus database.

Table 4. The most cited articles in Scopus on “artificial intelligence in MOOCs”

Title	Authors	Journal	Year	Citations
Rebooting MOOCresearch[7]	Reich, J.	Science	2015	187
MOOCs: So many learners, so much potential [8]	Kay, J., Reimann, P., Diebold, E., Kummerfeld, B.	IEEE Intelligent Systems,	2013	157
Temporal Models for Predicting Student Dropout in Massive Open Online Courses [9]	Fei, M., Yeung, D.-Y.	Proceedings - 15th IEEE International Conference on Data Mining Workshop, ICDMW	2015	135
Characteristics of massive open online courses (MOOCS): A research review, 2009-2012 [10]	Kennedy, J.	Journal of Interactive Online Learning,	2014	112
Combination of machine learning algorithms for recommendation of courses in E-Learning System based on historical data [11]	Aher, S.B., Lobo, L.M.R.J.	Knowledge-Based Systems,	2013	107

The Scopus database was analysed for keywords using VOSviewer version 1.6.18. The minimum threshold of 25 appearances was required for a word to be included in the research. Only 24 keywords out of 2792 terms satisfied this criterion. Figure 3.4 depicts a co-occurrence network map, generated by VOSviewer using 24 keywords relying on a relevance score with three final clusters depicted in different colours (red, green, and blue). With phrases like "forecasting", “machine learning”, “learning systems”, “learning algorithms”, “decision trees”, “learning analytics” and “data mining” the red cluster appears to focus on the predictive aspects of MOOC. The green cluster appears to be focused on many elements that may influence MOOC acceptance, with phrases like “online education”, “distance education”, “online learning”, “big data", “intelligent systems”, “students”, “education computing” and “curricula” included. Finally, with the keywords like "education", “computer aided instruction”, "engineering education", "education", “artificial intelligence” and "teaching" the blue cluster appears to be focused on the implementation of the MOOCs.

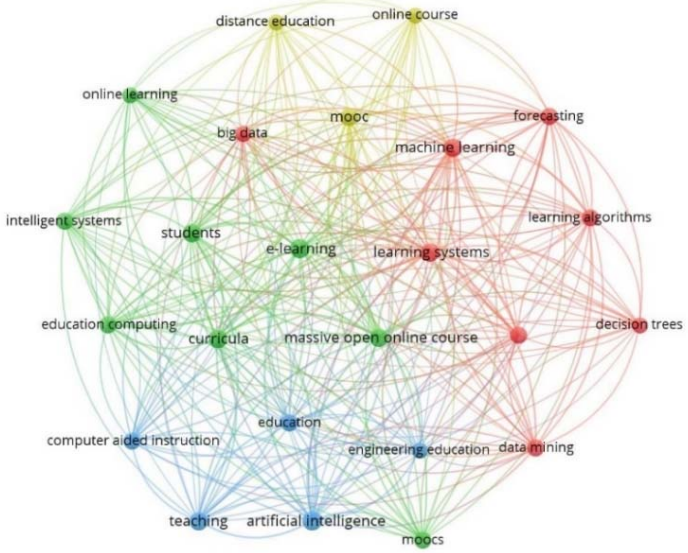


Figure 3.4. Keyword co-occurrence network map in the published documents

4. DISCUSSION AND CONCLUSIONS

Scopus database reveals that the number of documents published on use of artificial intelligence in MOOCs has increased in recent years. The majority of the publications were proceedings papers, signifying a high level of interest in the subject. Although, there exists a significant gap between conference papers (71.7%) and journal articles (20.3%). The literature suggests that MOOCs integrated with artificial intelligence are now a possibility. The study looked at a wide range of published works on artificial intelligence in MOOCs, demonstrating the researchers' heterogeneity on the subject. Several authors have discussed the MOOC dropout prediction studies [9], [12], [13][14]. While significant

scholars laid emphasis on numerous other topics some of these were viz., clickstream analysis of learners' [15], [16], satisfaction of the learners [17], [18], Time-based metrics related to assessments and student interactions [19], use of K-means algorithm MOOC dataset [20], machine learning algorithms to filter and classify the MOOC learners [21], emotional tendencies of learners [22], learning behaviours in MOOC [23], intelligent analysis [24], Virtual and Augmented Reality [25] etc.

MOOCs (Massive Open Online Courses) are slowly conquering the education business. 2017 Higher Education Edition of the Horizon Report predicted that by 2022, Artificial intelligence will be applied in higher education [26]. Moreover, the technologies like IOT, blockchain and AR / VR can be incorporated in the future. Although pandemic has disrupted the last two years, AI is set to become a widespread feature in MOOCs within few years. While artificial intelligence and MOOCs are both real, we acknowledge that artificial intelligence research in MOOCs is still in its early phases.

5. REFERENCES

[1] S. Al-Qatawneh, N. Alsalhi, A. Al Rawashdeh, T. Ismail, and K. Aljarrah, "To E-textbook or not to E-textbook? A quantitative analysis of the extent of the use of E-textbooks at Ajman University from students' perspectives," *Educ. Inf. Technol.*, vol. 24, no. 5, pp. 2997–3019, 2019, doi: 10.1007/s10639-019-09912-4.

[2] I. Ahmad, S. Jasola, and A. Anupriya, "Supplementing higher education with MOOCs: A case study," in *2017 International Conference on Emerging Trends in Computing and Communication Technologies, ICETCCT 2017*, 2018, vol. 2018-Janua, doi: 10.1109/ICETCCT.2017.8280346.

[3] I. Ahmad, S. Jasola, S. Ahmad, and S. Sharma, "Technology Enabled Learning in Higher Education: A Case Study," *Xplore Xavier's Res. J.*, vol. 11, no. 1, pp. 68–73, 2020, [Online]. Available: https://9dec1e70-3d6c-406f-960b-920f18cdab25.filesusr.com/ugd/03b37e_0a4fe5f7f9764c23a239c95f05ae24fc.pdf.

[4] I. Ahmad, S. Ahmad, A. K. Pant, and R. Kumar, "Intelligent Decision Support System for Tourism," *Int. J. Latest Res. Sci. Technol.*, vol. 8, no. 5, pp. 1–4, 2019, [Online]. Available: https://www.mnkjournals.com/journal/ijlrst/pdf/Volume_8_5_2019/10978.pdf.

[5] C. Soledad, "Intelligent Systems in Education: A review of current research lines," *RELIEVE*, vol. 10, pp. 3–22, 2004.

[6] G. Cumming and J. Self, "Intelligent educational systems: identifying and decoupling the conversational levels," *Instr. Sci. 1990 191*, vol. 19, no. 1, pp. 11–27, Jan. 1990, doi: 10.1007/BF00377983.

[7] J. Reich, "Rebooting MOOC Research," *Science (80-.).*, vol. 347, no. 6217, pp. 34–35, Jan. 2015, doi: 10.1126/SCIENCE.1261627.

[8] J. Kay, P. Reimann, E. Diebold, B. Kummerfeld, and B. (2013). Kay, J., Reimann, P., Diebold, E., & Kummerfeld, "MOOCs: So Many Learners, So Much Potential ... IEEE Intelligent Systems," *IEEE Intell. Syst.*, vol. 28, no. 3, pp. 70–77, 2013.

[9] M. Fei and D. Y. Yeung, "Temporal Models for Predicting Student Dropout in Massive Open Online Courses," *Proc. - 15th IEEE Int. Conf. Data Min. Work. ICDMW 2015*, pp. 256–263, 2016, doi: 10.1109/ICDMW.2015.174.

[10] J. Kennedy, "Characteristics of massive open online courses (MOOCS): A research review, 2009-2012," *J. Interact. Online Learn.*, vol. 13, no. 1, pp. 1–16, 2014, [Online]. Available: http://www.ncolr.org/jiol/issues/pdf/13.1.1.pdf.

[11] S. B. Aher and L. M. R. J. Lobo, “Combination of machine learning algorithms for recommendation of courses in E-Learning System based on historical data,” *Knowledge-Based Syst.*, vol. 51, pp. 1–14, 2013, doi: 10.1016/j.knosys.2013.04.015.

[12] J. Liang, C. Li, and L. Zheng, “Machine learning application in MOOCs: Dropout prediction,” *ICCSE 2016 - 11th Int. Conf. Comput. Sci. Educ.*, no. Iccse, pp. 52–57, 2016, doi: 10.1109/ICCSE.2016.7581554.

[13] D. Sun, Y. Mao, J. Du, P. Xu, Q. Zheng, and H. Sun, “Deep learning for dropout prediction in MOOCs,” *Proc. - 2019 8th Int. Conf. Educ. Innov. through Technol. EITT 2019*, pp. 87–90, 2019, doi: 10.1109/EITT.2019.00025.

[14] F. Dalipi, A. S. Imran, and Z. Kastrati, “MOOC dropout prediction using machine learning techniques: Review and research challenges,” in *IEEE Global Engineering Education Conference, EDUCON*, May 2018, vol. 2018-April, pp. 1007–1014, doi: 10.1109/EDUCON.2018.8363340.

[15] A. Alamri, Z. Sun, A. I. Cristea, G. Senthilnathan, L. Shi, and C. Stewart, “Is MOOC learning different for dropouts? A visually-driven, multi-granularity explanatory ML approach,” 2020, doi: 10.1007/978-3-030-49663-0_42.

[16] M. M. Al-Rifaie, M. Yee-King, and M. D’Inverno, “Boolean prediction of final grades based on weekly and cumulative activities,” *2017 Intell. Syst. Conf. IntelliSys 2017*, vol. 2018-Janua, no. September, pp. 462–469, 2018, doi: 10.1109/IntelliSys.2017.8324334.

[17] B. Hmedna, A. El Mezouary, O. Baz, and D. Mammass, “A machine learning approach to identify and track learning styles in MOOCs,” *Int. Conf. Multimed. Comput. Syst. -Proceedings*, vol. 0, pp. 212–216, 2017, doi: 10.1109/ICMCS.2016.7905606.

[18] K. F. Hew, X. Hu, C. Qiao, and Y. Tang, “What predicts student satisfaction with MOOCs: A gradient boosting trees supervised machine learning and sentiment analysis approach,” *Comput. Educ.*, vol. 145, p. 103724, 2020, doi: 10.1016/j.compedu.2019.103724.

[19] N. I. Jha, I. Ghergulescu, and A.-N. Moldovan, “OULAD MOOC Dropout and Result Prediction using Ensemble, Deep Learning and Regression Techniques,” doi: 10.5220/0007767901540164.

[20] S. Shrestha and M. Pokharel, “Machine Learning algorithm in educational data,” Nov. 2019, doi: 10.1109/AITB48515.2019.8947443.

[21] Y. Mourdi, M. Sadgal, W. B. Fathi, and H. El Kabtane, “A machine learning based approach to enhance MOOC users’ classification,” *Turkish Online J. Distance Educ.*, vol. 21, no. 2, pp. 54–68, 2020, doi: 10.17718/TOJDE.727976.

[22] L. Wang, G. Hu, and T. Zhou, “Semantic analysis of learners’ emotional tendencies on online MOOC education,” *Sustain.*, vol. 10, no. 6, 2018, doi: 10.3390/su10061921.

[23] W. Feng, J. Tang, and T. X. Liu, “Understanding dropouts in MOOCs,” *33rd AAAI Conf. Artif. Intell. AAAI 2019, 31st Innov. Appl. Artif. Intell. Conf. IAAI 2019 9th AAAI Symp. Educ. Adv. Artif. Intell. EAAI 2019*, pp. 517–524, 2019, doi: 10.1609/aaai.v33i01.3301517.

[24] C. Li and H. Zhou, “Enhancing the efficiency of massive online learning by integrating intelligent analysis into MOOCs with an Application to Education of Sustainability,” *Sustain.*, vol. 10, no. 2, 2018, doi: 10.3390/su10020468.

[25] F. M. Nuraliev, U. E. Giyosov, and Y. Okada, “Enhancing Teaching Approach with 3D Primitives in Virtual and Augmented Reality,” in *World Conference*

Intelligent System for Industrial Automation, Springer, Cham., 2021, pp. 155–163.
[26] S. A. Becker *et al.*, *NMC Horizon Report: 2017 Higher Education Edition*. 2017.

Biographies

Ishteyaaq Ahmad is an academician in Uttaranchal Institute of Management, Uttaranchal University, Dehradun. His areas of interest include Online education, Software Engineering and Artificial Intelligence. He has to his credit a MOOC developed for SWAYAM platform and two e-Learning courses developed for Commonwealth of Learning (COL). He is also a recipient of Young Scientist and Professional research grant funded by Department of Science & Technology, Government of India.

Sonal Sharma is Professor & head-Department of Computer Applications, Uttaranchal University with an experience of more 21 years. She possesses strong leadership and administrative skills. Her research interest is in Data Warehousing, Data Mining, Machine Learning, Big Data Analytics and Data Sciences. She has presented/published 45 research papers in national and international conferences/seminars/journals, out of which 11 are Scopus Indexed. She has presided over as a speaker in IEEE conference at Chengdu, China.

Gulsun Kurubacak is a senior professor in Distance Education at the College of Open Education of Anadolu University.Dr. Kurubacak undertook graduate studies at Anadolu University, Turkey (MA. Educational Technology) and the University of Cincinnati, USA (Ed.D. Curriculum & Instruction), and also has worked a post-doctoral fellow at the College of Education at New Mexico State University, USA. She has abundance experience with R&D as well as Accreditation and Quality in open and distance learning at the College of Open Education.

Ajay Kumar working as Assistant Professor in Bharati Vidyapeeth (Deemed to be university) Institute of Management &Research New Delhi. He has more than 10 years of experience in teaching and his area of interest is Network Security and Machine Learning. He is Pursuing PhD. from Mewar University.

Rajiv Kumar is Associate Professor in Uttaranchal Institute of Management, Uttaranchal University, Dehradun. He is B.Tech, M.Tech in Computer Science and Engineering, PhD in computer science and Engineering. Dr. Rajiv Kumar has published several research papers in national and international journals of repute. His research interest includes Genetic Algorithm, Memetic Algorithm.

Sikha Ahmad is an academician in School of Social Sciences, Doon University, Dehradun. She is gold medalist in Economics from Gauhati University. She specializes in Mathematical Economics and Econometrics. Her research interests include Online learning, Behavioural Economics, Neuroeconomicsand Tourism. She is also the Course Coordinator of a popular MOOC on SWAYAM.

RoTFace: A Framework for Robust and Time Efficient Face Recognition System

Shilpa Garg[1], Dr. Sumit Mittal[2], Dr. Pardeep Kumar[3]

[1]MMICT & BM, MM(DU), Mullana, Ambala, Haryana, India, E-mail: shilpa111987@gmail.com

[2]MMICT & BM, MM(DU), Mullana, Ambala, Haryana, India, E-mail: sumit.mittal@mmumullana.org

[3]Department of Computer Science & Applications, Kurukshetra University, Kurukshetra, Haryana, India, E-mail: mittalkuk@gmail.com

ABSTRACT

Now a days face recognition is commonly used applications which need security. Many authentication systems have been developed like fingerprint, palm, iris, face and many more but face recognition system is widely used as authentication system to verify the person's identity as the face is most natural way to identify the person, its uniqueness property and as it requires no touch to the screen that's why its use is rising day by day but due to spoofing attack like photo attack, video replay attack, facial mask attack and many more, face recognition requires more attention. Also due to large date set, it takes a long time to search the client id. This paper proposed a framework RotFace for robust and time efficient face recognition system. At first, deep featured are extracted of the face images using deep network ResNeT50 then classification is divided into two parts. In first part, real attack predictor is applied by using Gaussian Naïve Bayes classifier. In second part, a client id predictor is applied by using KNN. Experiment is executed on Replay Attack Dataset and achieves very good results in terms of accuracy and time. This paper also analyses the proposed framework with the existing techniques of face recognition system and performs better than the existing techniques.

Keywords – Face Recognition, Liveness Detection, Deep Network, Gaussian NaïveBayes, KNN

1. INTRODUCTION

For the security purpose, many authentication systems have been developed but face recognition system is widely used as authentication system to verify the person's identity [1] as the face is most natural way to identify the person [2], its uniqueness property [3] and as it requires no touch to the screen that's why its use is rising day by day. But due to spoofing attack like photo, 2D & 3D mask, video replay and many more [4], it is misused by the attacker.Attackers used the recorded video of the authorized person in front of sensor device to authenticate the user and have gain access to the others device. So,for robust face recognition, it is necessary to check liveness of the authentic person [5].

A robust and time efficient framework for face recognition is proposed in this paper. Robustness of the face recognition system is to check whether the person is live or not.

For the experiment Replay Attack dataset of the face videos is used. This dataset contains both real and fake videos of the persons. At first, videos of the face are converted to the frames. Then the class balancing is done to remove the imbalance problem of the real and fake frames so that it will not affect the accuracy results. After balancing real and fake image, feature extraction technique is applied to extract deep features of the image and then the system is trained using real attack predictor.

If the image is detected as fake, it will skip the image for further processing which will decrease the run time and if the image is detected as real then the image is forwarded to the next step i.e., to the client id predictor which will recognize the face and gives the client id.

As seen in the Figure 1, after checking liveness of the image, image is discarded if the image is fake which will reduce the run time. Run time is vary from system to system. This

experiment is done in python 3.7 installed on i7 8th generation, 240 SSD, 2TB HDD, 8gb ram, 4gb Graphics system.

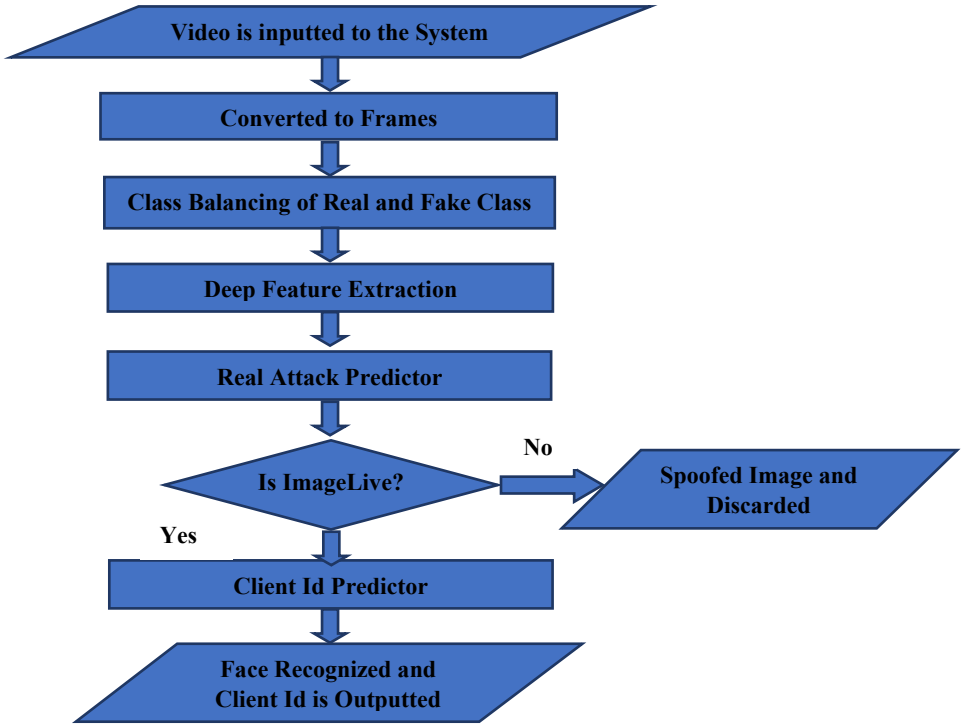


Figure 1: Proposed Framework of RotFace

2. LITERATURE SURVEY

Face recognition is widely used for the authenticity of a person these days and that is the reason, face recognition is a popular for research purpose. Many deep learning techniques for face recognition have been implemented.Mugalu et. al. proposed a web-based face recognition system in which MTCNN is used for detection and LinearSVC is used for classification and achieved 95% accuracy [6]. Sajjad et. al. suggested a hybrid approach for anti-spoofed face recognition system. This hybrid approach is two tier approaches. In tier-I,palm prints, face and finger print are integrated and in second tier, CNN is used to detect spoofing attack. This approach is applied on five dataset and achieved very promising results [7].

Linn et. al. proposed an anti-spoofing face recognition by detecting movement of eye CNN is used to extract features and classification. Experiment is done on three different dataset OWN Replay, NUAAand Replay Attack and achieved 96.5%, 98% and 95.8% accuracy respectively which is better than the exiting techniques [8]. Increase in the use of computing device and mobile needs checking of authentic user and many deep learning approached have been developed for face recognition [9]. Many deep feature extraction techniques like LBP, SIFT, DoG, HoG and SURF [10,11] etc. have been developed for feature extraction of the images similarly different deep learning classifier have been developed for detection of spoof attack.

3. CLASS IMBALANCE

For the experiment REPLAY ATTACK dataset is used from which video clips of real user and hand attacked video clips are used means video of real client in mobile or tablet is placed in front of sensor to access the device. These video clips are first converted to the frames and number of frames obtained from real and hand attacked video clips are different in number means class imbalance in real and hand attacked classes and class imbalance lead to deteriorate the accuracy. Different class imbalance ration gives different accuracy means if the ratio of real images are more than the spoofed images, it will lead to more accuracy but if the ratio real images are less than the spoofed images then the accuracy is less so it is compulsory to balance the class before classification [12].

4. FEATURE EXTRACTION

For deep feature extraction of the images five different deep features extraction techniques are used- ReSNet50, VGG16, VGG19, InceptionV3 and DenseNet121.These deep networks are the pretrained weighted networks. ReSNeT is referred to Residual Neural Network given by Kaiming et. al. in 2015 and perform better than VGG for image processing [13].

VGG16 is type of CNN model given by Simonyan et. al. VGG16 isbetter than AlexNet suggested by the authors and gives better accuracy for image processing[14].VGG19 is improvement over VGG16 and comprises of some more layers than VGG 16 and deeper than VGG16 [15]. InceptionV3 is also a deep CNN architecture used for image processing given by Google and developed as a challenge for imagenet recognition [16]. DenseNet is proposed in 2018 by Huang et al., suggested that this network is also an improvement over some other state of the art and worked in feed forward direction [17].

5. REAL ATTACK PREDICTOR

Real attack predictor is trained and used to predict the liveness of the person by using the deep features extracted from different techniques and machine learning algorithm - Gauss Naïve Bayes and Random Forest. Gauss Naïve Bayes is commonly used for image classification [18] and Random Forest performs very well to predict liveness of recognition system [19]. This trained Real Attack Predictor is used to predict the imputed image as real or spoofed during testing phase.

6. CLIENT ID PREDICTOR

Client id predictor is trained using deep feature of the face images and two machine learning algorithm – k-Nearest Neighbors (KNN) and Decision Tree. k-Nearest Neighbor and Decision Tree are good classifier for image processing [20]. After detecting liveness of the image, only lived person images is send to the client id predictor for face recognition which predict the client id in testing phase.

7. METHODOLOGY

Methodology discussed the dataset used, performance evaluation metrics and experimental setup and results.

7.1. DATASET

REPLAY-ATTACK dataset [21] contains recorded video clips of 50 persons in two varied conditions - adverse and controlled. In controlled condition, background is illuminated with fluorescent lamp light and uniform whereas in adverse condition, background is non uniform and illuminated with day light. This dataset contains two type of attack frame - Fixed and Hand. In fixed attack, a hard copy of the image of a person is placed at fixed position in front of camera whereas in Hand attack, videoclips from Mobile and tablet is hold by the operator's hand.

Figure 2: Sample Images of the Replay Attack Dataset. Top Row shows the Controlled Condition and bottom row shows the adverse condition. From left to right- Real Image, mobile phone and Tablet Attack

For this experiment, Real and Hand attack recorded video clips of 20 persons are used and converted to image frames. From these frames, 50 real frames of each client are chosen and 50 hand attacked frames of each client are chosen so total 2000 images are used for experimental work of proposed framework in which 1000 images are real images and 1000 images are attacked images of 20 clients.

7.2. PERFORMANCE EVALUATION METRICS

For measure and compare RoTFace performance, two metrics Accuracy and F1 score are used [22].Accuracy is used to calculate true prediction and represented by equation 1.F1 Score is also an accuracy measure and defined as weighted mean of recall and precision represented in equation 2.

$$Acc = \frac{t_{pv}+t_{nv}}{t_{pv}+t_{nv}+f_{pv}+f_{nv}} \quad (1)$$

$$F1 = \frac{2t_{pv}}{2t_{pv}+f_{pv}+f_{nv}} \quad (2)$$

Where t_{pv} predicts real images as real, t_{nv}predicts attacked images as attacked, f_{pv}predicts attacked images as real, f_{nv} predicts real images as attacked.

7.3. EXPERIMENT SETUP AND RESULTS

For experiment, 60:40 ration is followed for training and testing. From 2000 images of the total images, 600 real and 600 attacked i.e., total1200 images are used for training the predictor and 400 real and 400 attacked i.e., total 800 images are used for testing.
In experiment, five different feature extraction techniques VGG16, VGG19, inceptionV3, DenseNet121, ReSNet50 are used to extract deep features of the image. Then Gauss Naïve Bayes classifier is used as real attack predictor and then K Nearest Neighbour(KNN) is used as client id predictor similarly after deep feature extraction, Random Forest is used as real attack predictor and Decision Tree is used as client id predictor. These approaches are represented as VGG16_GK, VGG16_RFDT, VGG19_GK, VGG19_RFDT, InceptionV3_GK, InceptionV3_RFDT,DenseNet_GK, DenseNet_RFDT, ReSNeT_RFDT and RoTFace (ReSNet_GK) and gives 97.64%, 94.132%, 98.58%, 96.25%, 94.63%, 90.201%, 98.37%, 91.939%, 90.523%, 99.7% accuracy respectively.
Table 1 shows the accuracy, F1 score and Run time excluding training time (in sec) results of different approached used in experiment.
As shown in table, run time of proposed approach is too less than the other except InceptionV3_GK. Run time of inceptionV3_GK and RoTFace is 1.11 seconds but accuracy of proposed approach is much better than the InceptionV3_GK so the proposed approach is performing better than the other existing techniques.

Table 1:comparison results showing the accuracy, F1 score and Run time

Approach Used	Accuracy (in %age)	F1 score (in %age)	Run Time (in Seconds)
VGG16_GK	97.64	96.37	1.353
VGG16_RFDT	94.132	93.986	5.874
VGG19_GK	98.58	98.38	1.356
VGG19_RFDT	96.25	96.138	5.878
InceptionV3_GK	94.63	93.97	1.11
InceptionV3_RFDT	90.201	90.276	6.152
DenseNet_GK	98.37	96.81	3.28
DenseNet_RFDT	91.939	91.956	16.293
ReSNet50_RFDT	90.523	90.976	17.581
RoTFace (ReSNet50_GK) Proposed	**99.7**	**99.16**	**1.11**

Figure 3 Shows the graphical representation of accuracy and F1 score and clearly shows that proposed framework is performed better than the other approach used. Figure 4 shows graphical representation of the run time comparison of proposed approach with other

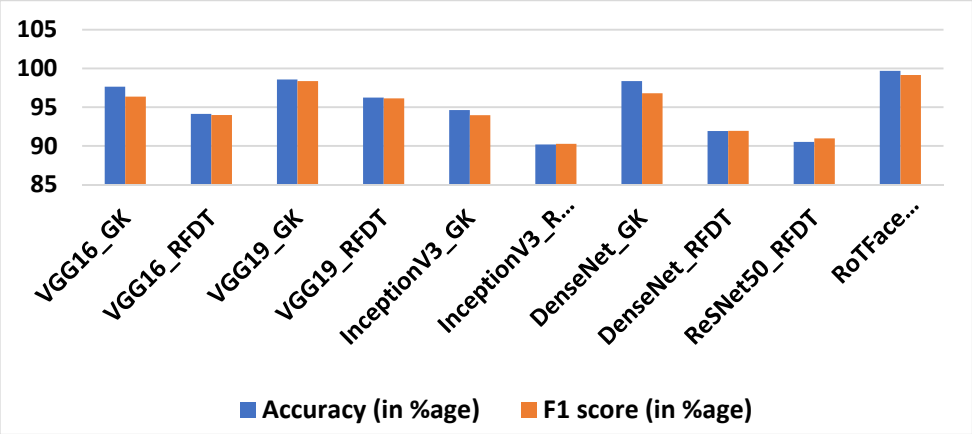


Figure 4:Accuracy and F1 Score comparison Graph

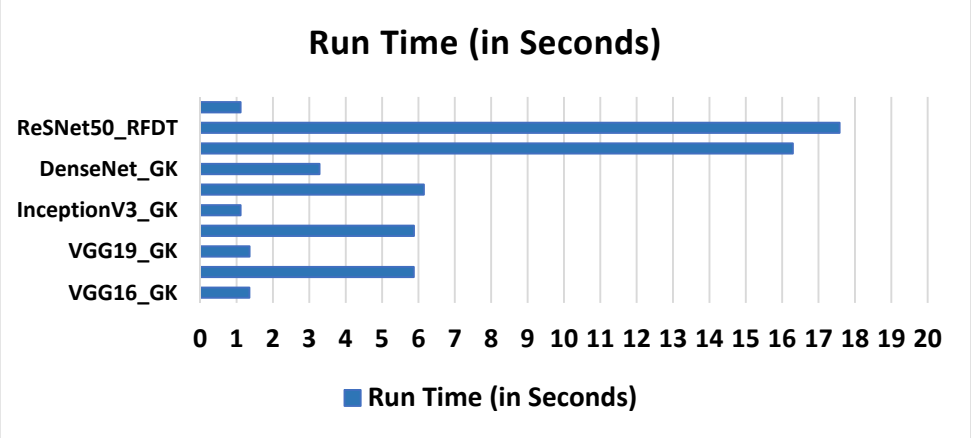


Figure 3:Graph shows the Run time comparison

approach.

8. CONCLUSION

This paper presents a robust and time efficient framework for face recognition system. The proposed approach RoTFace extracts the deep feature using ReSNet50 and Gauss Naïve Bayes and KNN classifiers are used to predict the spoofing attack and client id respectively. Also, to analyses the result with existing techniques, four other different deep feature technique is applied i.e., VGG16, VGG19, InceptionV3, DenseNet121 with Gauss Naïve Bayes and Random Forest as Real Attack predictor and KNN and Decision Tree as Client Id predictor. RoTFace gives 99.7% accuracy, 99.16% F1 score and runs in 1.11 seconds only and performed better than the existing techniques in terms of both run time and accuracy.As a future work try to work on real time face recognition and will try to reduce the run time of real time face recognition system.

REFERENCE

1. Thepade, S., Jagdale, P., Bhingurde, A. and Erandole, S., 2020, February. Novel Face Liveness Detection Using Fusion of Features and Machine Learning Classifiers. In *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)* (pp. 141-145). IEEE.

2. Garg, S., Mittal, S. and Kumar, P., 2019. Performance Analysis of Face Recognition Techniques for Feature Extraction. Journal of Computational and Theoretical Nanoscience, 16(9), pp.3830-3834.
3. Garg, S., Mittal, S., Kumar, P., &Athavale, V. A. (2020, February). DeBNet: Multilayer Deep Network for Liveness Detection in Face Recognition System. In 2020 7th International Conference on Signal Processing and Integrated Networks (SPIN) (pp. 1136-1141). IEEE.
4. Sajjad, M., Khan, S., Hussain, T., Muhammad, K., Sangaiah, A.K., Castiglione, A., Esposito, C. and Baik, S.W., 2019. CNN-based anti-spoofing two-tier multi-factor authentication system. *Pattern Recognition Letters*, *126*, pp.123-131.
5. Bharadwaj, S., Dhamecha, T.I., Vatsa, M. and Singh, R., 2013. Computationally efficient face spoofing detection with motion magnification. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* (pp. 105-110).
6. Mugalu, B.W., Wamala, R.C., Serugunda, J. and Katumba, A., 2021. Face Recognition as a Method of Authentication in a Web-Based System. *arXiv preprint arXiv:2103.15144*.
7. Sajjad, M., Khan, S., Hussain, T., Muhammad, K., Sangaiah, A.K., Castiglione, A., Esposito, C. and Baik, S.W., 2019. CNN-based anti-spoofing two-tier multi-factor authentication system. *Pattern Recognition Letters*, *126*, pp.123-131.
8. Linn, P.P.P. and Htoon, E.C., 2019, November. Face Anti-spoofing using Eyes Movement and CNN-based Liveness Detection. In *2019 International Conference on Advanced Information Technologies (ICAIT)* (pp. 149-154). IEEE.
9. Saha, S., Xu, W., Kanakis, M., Georgoulis, S., Chen, Y., Paudel, D.P. and Van Gool, L., 2020. Domain agnostic feature learning for image and video-based face anti-spoofing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (pp. 802-803).
10. Sun, W., Song, Y., Chen, C., Huang, J. and Kot, A.C., 2020. Face spoofing detection based on local ternary label supervision in fully convolutional networks. *IEEE Transactions on Information Forensics and Security*, *15*, pp.3181-3196.
11. Patel, K., Han, H., Jain, A.K. and Ott, G., 2015, May. Live face video vs. spoof face video: Use of moiré patterns to detect replay video attacks. In *2015 International Conference on Biometrics (ICB)* (pp. 98-105). IEEE.
12. Vo, T., Nguyen, T. and Le, C.T., 2019. A hybrid framework for smile detection in class imbalance scenarios. *Neural Computing and Applications*, *31*(12), pp.8583-8592.
13. Sharma, N., Jain, V. and Mishra, A., 2018. An analysis of convolutional neural networks for image classification. *Procedia computer science*, *132*, pp.377-384.
14. Fu, Y. and Aldrich, C., 2019. Flotation froth image recognition with convolutional neural networks. *Minerals Engineering*, *132*, pp.183-190.
15. Shaha, M. and Pawar, M., 2018, March. Transfer learning for image classification. In *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)* (pp. 656-660). IEEE.
16. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. and Rabinovich, A., 2015. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1-9).
17. Huang, G., Liu, Z., Van Der Maaten, L. and Weinberger, K.Q., 2017. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700-4708).
18. Rafique, A.A., Jalal, A. and Ahmed, A., 2019, August. Scene Understanding and Recognition: Statistical Segmented Model using Geometrical Features and Gaussian Naïve Bayes. In *IEEE conference on International Conference on Applied and Engineering Mathematics* (Vol. 57).
19. Thepade, S.D., Chaudhari, P., Dindorkar, M., Bang, S. and Bangar, R., Improved Face Spoofing Detection Using Random Forest Classifier with Fusion of Luminance Chroma.
20. Dino, H.I. and Abdulrazzaq, M.B., 2019, April. Facial expression classification based on SVM, KNN and MLP classifiers. In *2019 International Conference on Advanced Science and Engineering (ICOASE)* (pp. 70-75). IEEE.

21. Chingovska, I., Anjos, A. and Marcel, S., 2012, September. On the effectiveness of local binary patterns in face anti-spoofing. In *2012 BIOSIG-proceedings of the international conference of biometrics special interest group (BIOSIG)* (pp. 1-7). IEEE.
22. Powers, D. M. (2020). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *arXiv preprint arXiv:2010.16061*.

A Real Trust Behaviour Based ACO Routing Algorithm for MANET

Rashmi Mishra[1], Sakshi Koli[2], Pragya Kamal[3], Ramnarayan[4], Preeti Raturi[5], Emmanuel Ododo[6]
Assistant Professor Tula's Institute Dehradun[1,2,5,], ACM Member[3], Assistant Professor UIT, Uttaranchal University Dehradun[4], Graduate Assistant-Robotics Education, University of Uyo, Nigeria[6],
rashmi.mar30@gmail.com[1], kolisakshi84@gmail.com[2], pragya.kamal.official@gmail.com[3], ram6801@gmail.com[4], raturi.nidhi4@gmail.com[5], emmanuel.ododo15@gmail.com[6]

Abstract- The Ant Colony Optimization technique is used in this research to locate routes in a MANET, which is a novel approach. Pheromone values are updated in this protocol using a real-life trust behavior technique. In an ACO-based protocol, the start and updating of pheromone values are critical to the system's success. We used various methods to initialize the pheromone values and then changed them based on real-world trust behavior. Simulation tests over NS2 for 20-50 nodes in a 800x800 region are used to compare the AODV protocol and our proposed protocol.We show that the suggested protocol can exceed AODV in average delay and delivery ratio parameters. In terms of jitter and packet drop also, the proposed approach performs better. Because many disjoint routing paths give robustness to mobility, our investigation reveals that the proposed protocol outperforms existing pertinent techniques.

***Keywords*:** MANET, Trust Based algorithm, ACO, Routing Algorithm, AODV.

1. INTRODUCTION

Wireless networking gives electronic accessibility of services and information regardless geographic position of the user. Wireless network are basically of two types. This property of self-configuration and self-organization makes these networks unique and of higher applicability. Mobile Ad-hoc networks are self-configuring network [1]. They are infrastructure less networks and communication is done over wireless medium. These networks are self configuring in a sense that for connection establishment, routing or for link maintenance, no designated devices are used. Mobile node using radios communicate through each other and are responsible for finding and maintaining optimal route for transmission of data packets. Self-configuring nature of these networks, make them highly desirable in some fields like disaster recovery, police and military networks, mining operations, safety and rescue operations and in providing efficient, secure and comfort driving (VANET). Due to dynamic network connections routing is a tough task in MANET. Wireless link and frequently changing topology makes the route discovery and maintenance a tough task. A reliable path found at any instance of time can become highly unreliable after sometimes. Discovery of stable path is important in these networks due to energy limitation of mobile nodes [2].The size of MANET is required to be of larger size than that of the radio range of the wireless antennas. The size of the MANETs makes it necessary to do routing in a multi-hop fashion to provide communication ability to two nodes placed far apart in the network[3].MANETs contrast infrastructure network on the basis of multi-hop routing also. In infrastructure networks only one hop communication between mobile nodes and base stations is allowed. But in case of MANET all nodes can work as router and can establish multi-hop routes. An infrastructure network and MANET is depicted in Figure 1 and in Figure 2 respectively.

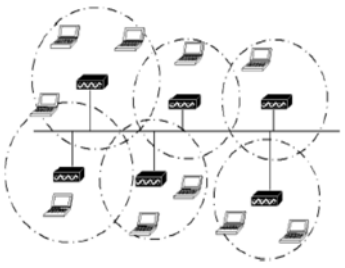


Figure 1 Network with fixed infrastructure

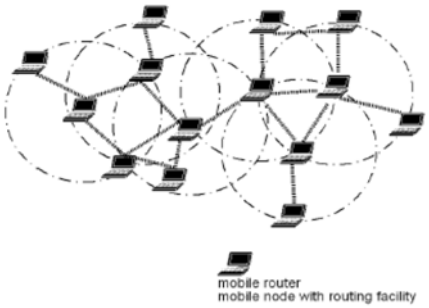


Figure 2. Mobile Ad hoc Network

- Proactive Protocols: These protocols use tables for routing information. Each node keeps track of all routes to known destinations in a routing table.[6] Routing data is available at all times, but storage costs are considerable.
- Reactive Protocols: These are on-demand protocol, i.e. the route is discovered when it is demanded by any node. Storage overhead is low as no unnecessary route information stored.
- Hybrid Protocols: They combine the positive factors of both previously defined protocols. By mixing the features of two protocols better results are obtained.

Figure 3 depicts the routing protocols classification.

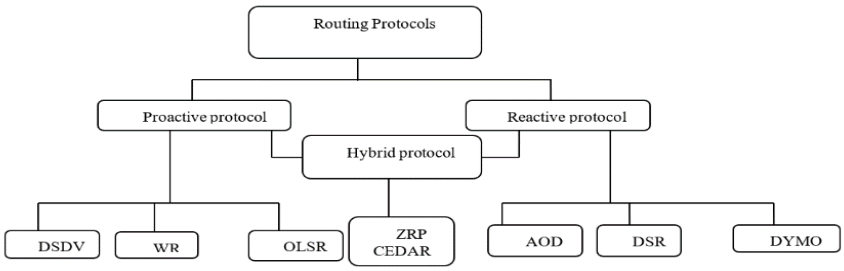


Figure 3. classification of MANET routing protocols

Before we precede toward our proposed algorithm it is necessary to understand ACO based routing and its suitability for routing in MANET.

The Ant Colony Optimization Approach: ACO represents a colony of the artificial ants which find global solutions to the challenging optimization problems through cooperation [8]. Artificial ants are the mobile agents that are somewhere different from the real ants. These mobile agents inherit the some traits from their natural counterpart but with that they have some additional properties that make them more relevant to be used with an engineering approach to solve some difficult problems and to provide efficient solution to them. Below are some traits that compare natural ants with the artificial one.

In [12], a large scale network is considered for routing. This protocol works over the concept of "logical routers" and "logical link". Logical routers are formed by the grouping of nodes which are geographically close.[10] So we can say that logical routers are collection of nodes and logical links are the paths between them. This protocol is also inspired by social behavior of insects. This paper represents a novel hybrid ACO based protocol called Ant Hoc Net[13]. MANET are dynamic in nature due to which routing is very tough in such network. In the demand based route discovery, forward mobile agents are used that collect the network information and on reaching the destination converted into backward ants to update the route information at each node. This algorithm provides multiple path setup facility. [11]The path maintenance phase exhibits the proactive behavior because along with the data packets some proactive ants are also sent that maintain the existing path and also try to find new ones. In[14], Lianggui Liu et al. recognized the problem of having single route selection which lead to load on shortest path. The paper represents a new routing theory named as AMQR. This protocol combines link disjoint multi-path routing and swarm intelligence approach. This protocol is very adaptive in nature and has better QoS support. In[15], author gave a unique approach of routing in MANET in which route discovery (reactive) and route maintenance (proactive) is done in hybrid manner. Ant like mobile agents is used in this ACO routing technique. Results show that this algorithm works efficiently than AODV in different scenario. In[16]an ACO based algorithm called ARAMA is proposed which performs resource management in MANET routing. Simulation of this work is done on GlomoSim for the comparison between HOPNET with AODV.

To achieve the efficient routing an ACO based scheme for multipath routing is proposed which is based on real life trust behavior. The proposed approach consider the value of pheromones as a key factor and gave a unique protocol.Pheromone disposed on the path defines the goodness and utility factor of that path. In this way initializing value of that pheromone and how these values are getting updated plays very important role in the performance of the algorithm. In our proposed algorithm we have especially worked over these two factors.

2. Proposed Approach

The proposed protocol exploits ACO for the construction of optimum routes based on AODV approach. The main objective of this research is to propose an ACO based protocol in which pheromone update process is inspired by a practical trust behaviour. Pheromone disposed on the path defines the goodness and utility factor of that path. In this way initializing value of that pheromone and how these values are getting updated plays an important role in the algorithm performance. In our proposed algorithm we have especially considered over these two factors. In ACO based routing algorithm pheromone values tells that how beneficial it would be to take that path to reach a particular destination, so in a way we can say that pheromone represents the trust toward that path. We merge this concept with real world trust assessment. Finally the analysis of performance is done on the basis of different matrices. Simulation output shows that our algorithm outperformed AODV under different parameter.

In ACO pheromone values are required to be initialized at the starting part of procedure. Conventionally for ACO, all ant tracks are initialized with zero (0) since it is assumed that initially pheromones are not available on the tracks. But in our proposed algorithm, we initialize the pheromone value in somewhat different way. We initialize it with one (1). It is because of the fact that the pheromone intensity also represents the connection in between two residing routers. A path having zero (0) intensity represents that there is no connection in between two routers. Here intensity of pheromone represents the values of pheromone.

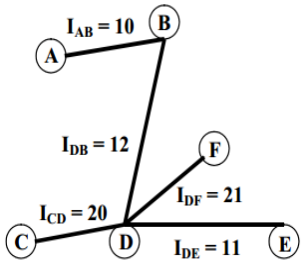


Figure 4.A network topology with a stable pheromone intensity

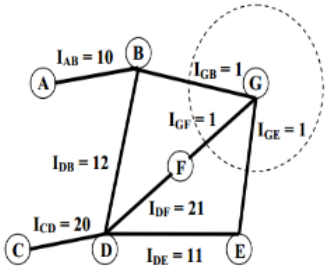


Figure 5. Changed topology in case ofaddition of a new router G.

We modified this approach and initialize the pheromone value by 1 which enhances the performance of algorithm. This modification is done in order to show that there is a path between B & G, F & G and E & G and Figure 5 clearly shows these modifications. One more benefit of using 1 as an initializing value is that if the source found that destination is in its routing table and has intensity value 1 i.e. if the source has direct link with the destination, then source would avoid sending any kind of ant packets in spite it would send a data packet directly.

Before discussing the core idea of pheromone updation behind our proposed approach, we will first take a real world example of trustworthiness evaluation as well as its changing behavior with changing circumstances and then explain our approach which will make it easy to understand the real benefits of applying this idea to this ACO based multipath routing algorithm for MANETs.

Let us consider a scenario of trust building process between initially two unknown people like Person A and Person B. Here, an assumption is taken that these two people have taken help from each other at 10 occasions and given a feedback in a binary manner, i.e. either they were happy with response or not.

Case1: If Person A continuously betrayed the trust of Person B for first three to four times, then there is a drastic decrease in trust level of Person B towards person A is noticed.

By considering this case we proposed a change in pheromones evaporation. Pheromones evaporation can be done exponentially because after each unsuccessful attempt of message/packet delivery, the trustworthiness of the route will degrade very drastically.

$$I_{AB}(new) = e^{(q-1)} I_{AB}(old)$$

Where; I_{AB} (new) = Updated pheromone value
I_{AB} (old) = Old pheromone value
$q \in (0, 1]$

Case2: If Person A continuously kept the trust of Person B for first three to four times, then there is a multiplicative increase in trust level of Person B towards person A will be noticed.

The same approach is applied in our practical trust base approach. In our proposed algorithm, pheromone value are updated in a quadratic manner because after each successful attempt of delivering whatever message or data we wants to deliver, the trustworthiness of the route increases multiplicatively.

$$I_{AB}(new) = I_{AB}(old) + (I_{AB}(old) * (q-1))^2$$

Where; I_{AB} (new) = Updated pheromone value
I_{AB} (old) = Old pheromone value
$q \in (0, 1]$

METHODOLOGY

This section defines all the basic function of the proposed routing algorithm.

Initialization: In our algorithm the pheromone values are initialized as one (1) in case there is any direct link in between two nodes. The concept of initialization is briefly explained before.

Route establishment: Route formation is a compulsory task in each routing technique. It is the basic job that router is allocated to perform. This job is performed as following:
When a new node enters in the web or at commencing of web, at early this router first finds its bordering or neighbour router and initializes the intensity of these paths as one (1). In addition, primarily it considers that across every single path, it is probable to grasp each router of the network. A kind of vector table is maintained such as distance vector routing in which intensity of their paths is stored. The table is coordinated in the standard form of "from A to B through C". But in spite of distance, it holds the intensity of given link. With bypassing of period, most excellent outgoing link for particular hosts will get stable in intensity table.
Route discovery: The proposed routing method doesn't dispatch RREQ packet as AODV or any other action as DSR. Yet it maintains its reactive nature. It sends ant packets for route discovery that calculate probability of each path. Before dispatching any packet to particular destination, it searches and selects the path on the basis of probabilistic selection explained below. Paths that are used in past by the most of the packets gets higher priority in selection. The given path invention is the adaptive method established on Ant Dominion Optimization and it becomes stable in the end by discovering nature.

These BAnts traverse backward and they update respective routing tables at each intermediate node and on reaching the destination (generator of the FAnts) they update its routing table and released. For each adaptive algorithm, probability calculation is a most vital issue but in given algorithm it plays an important responsibility. Successful accomplishment of given algorithm is affected by proper calculation of path probability. As dispatching the packet from one router to one more router, the path is selected with this probability. Probability of a path in this algorithm is computed on the pheromone intensity. Trail probability is described as,

$$P_i = \frac{I_{ij}}{I_t}$$

where,P_i = probability of the path connected with router. I_{ij}=intensity of path connected with router and node j.
I_t=summation of intensity of entire path connected with router.

- **Route Maintenance:** Proposed method maintains its path in an adaptive way. If a router gets down, it would be impossible to deliver any packets through the path containing that router. So due to pheromone evaporation the selection probability of thatpath would be decreased and other paths would begin to get priority over the path containing the broken link. If the down router gets back to network then that link would be established again.
- **Route selection:** Selection of routes is done on the basis of the probability calculated. Selection of each path is directly proportional to probability of that path. Pheromone deposition and evaporation process is defined earlier in this section.

3. SIMULATION AND RESULT ANALYSIS

All the simulation work has been done over NS 2. It is an event-driven tool for simulation. We present the simulation results as well as compare the proposed protocol with exiting protocol. This simulation and result demonstrate the correctness of propose scheme. Performance evaluation of the proposed protocol is done through following matrices:

- **Throughput:** This parameter defines the overall performance of the algorithm in terms of packet delivery ratio.
  In Figure 6 Xgraph shows a comparison between AODV and proposed algorithm. Where x-axis presents Simulation time (Sec) and y-axis represents packet delivery ratio.
- **End to end delay:** In Figure 7 on x-axis simulation time and on y-axis End to End Delay for each node (ms) is presented.
- **Jitter:** Jitter represents the variation in delay of received packets, i.e., when data packet flow from one node to another node some packet take longer time in transmission that cause variation in latency on data flow.
  Figure 8 represents graphical representation of jitter in AODV versus that in proposed algorithm. In graph x axis shows packets and y axis shows jitter value.

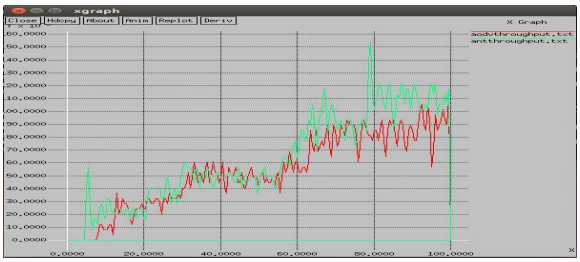


Figure 6. Xgraph for Throughput

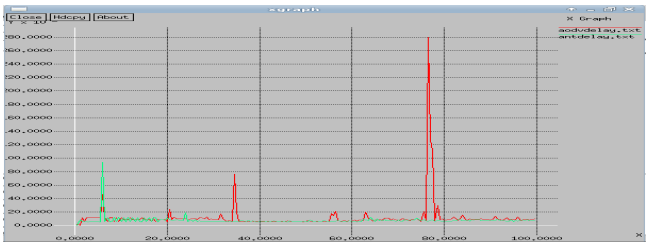


Figure 7.XGraph plotted for End to End Delay

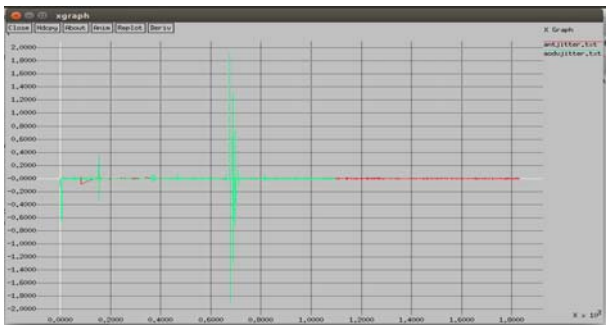


Figure 8.XGraph plotted for jitter

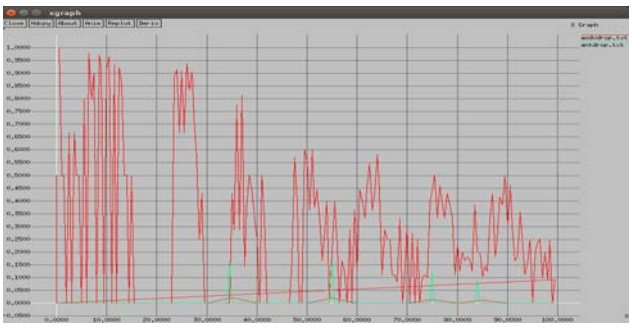


Figure 9. XGraph plotted for packet drop

Packet drop: In Figure 9 XGraph for packet drop (AODV v/s proposed algorithm) is drawn in which simulation time is taken on X-axis and drop rate is taken on Y-axis

CONCLUSION

We have proposed a multipath ANT-based routing protocol and a new ant-based algorithm. The hybrid algorithm is what it's called. The hybrid algorithm combines proactive route probing and exploration with reactive route setup. The routing method is based on swarm intelligence, specifically a meta-heuristic based on ant colonies. With little routing overhead, the protocol can identify numerous routing paths. Our suggested approach uses practical trust theory-based pheromone updating and also introduces a new method of initializing pheromone values, which improves the proto-overall col's performance. Simulation studies are also used to validate the protocol's usefulness. By considering delivery ratio and average delay, we show that the suggested protocol can outperform AODV in tough cases during the simulation process. The proposed approach also outperforms the competition in terms of delay jitter. We plan to expand and emphasize proactive ants' inquisitive behavior in future research. More information regarding probable path enhancements will be offered in the nodes as a result of enhancing the idea of pheromone diffusion, and this knowledge can guide proactive ants. As a result, you'll get better outcomes with less effort. Additionally, this technique can be improved by including safeguards against assaults such as black holes. In addition, an attempt will be made to develop a virtual circuit-based method. This could lead to greater path control, resulting in more dependable data transmission. Because many disjoint routing channels give robustness to mobility, our investigation reveals that the suggested ant-based approach outperforms existing pertinent techniques.

Reference

1. PreetiGulia, SumitaSihag.**Enhance Security in MANET using Bacterial Foraging Optimization Algorithm. International Journal of Computer Applications.**2013 December, 84(1), pp. 32-35 (International Journal of Computer Applications (0975 – 8887)) .
2. Seon Yeong Han, Dongman Lee.An Adaptive Hello Messaging Scheme for Neighbor Discovery in On-Demand MANET Routing Protocols IEEE Communications letters, Volume 17, Issue 5, 2013.
3. Jatinder Pal Singh, Anuj Kr. Gupta.**A Review on Dynamic MANET On- Demand Routing Protocol in MANETs. International Journal of Advanced Trends in Computer Science and Engineering.**2013 March – April, Volume 2, No.2.
4. Swati Dhawan, Vinod Saroha.**Review on Performance Issues of Routing Protocols of Mobile Ad-hoc Networks. International Journal of Advanced Research in Computer Science and Software Engineering**.2013 June ,Volume 3, Issue 6..
5. Ehsan Mostajeran, Rafidah Md Noor et al. **A Novel Improved Neighbor Discovery Method for an Intelligent-AODV in Mobile Ad-hoc Networks.**2013 May, IEEE, 978-1-4673-4992.
6. C. E. Perkins and P. Bhagwat.**Highly Dynamic Destination-Sequenced Distance-Vector Routing (DSDV) for Mobile Computers. ACM SIGCOMM'94,** 1994.
7. Haas ZJ, Pearlman MR, Samar P.The Zone Routing Protocol (ZRP) for Ad Hoc Networks. IETF draft, July 2002.
8. C. E. Perkins and E. M. Royer.**Ad-Hoc On-Demand Distance Vector Routing. 2nd IEEE Workshop. Mobile Computer Systems and Applications,**1999, pp. 90–100.
9. Mesut Günes, Udo Sorges, and Imed Bouazizi.**ARA-the ant-colony based routing algorithm for MANETs.Parallel Processing Workshops**, 2002., pp. 79-85.
10. John S. Baras, and Harsh Mehta.**A probabilistic emergent routing algorithm for mobile ad hoc networks. WiOpt'03: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks**, 2003, pp. 10.
11. O. Hussein, and T. Saadawi.Ant routing algorithm for mobile ad-hoc networks (ARAMA).Performance, Computing, and Communications Conference, 2003, pp. 281-290. (IEEE)

12. Marc Heissenbüttel, and Torsten Braun.**Ants-Based Routing in Large Scale Mobile Ad-Hoc Networks**.**KiVSKurzbeiträge,**2003, pp. 91-99.
13. Gianni Di Caro, Frederick Ducatelle, and Luca Maria Gambardella.**AntHocNet: an ant-based hybrid routing algorithm for mobile ad hoc networks.** Parallel Problem Solving from Nature-PPSN VIII, 2004, pp. 461-470.
14. Lianggui Liu, and Guangzeng Feng.**A novel ant colony based QoS-aware routing algorithm for MANETs. Advances in Natural Computation**, 2005, pp. 457-466.
15. Gianni Di Caro, Frederick Ducatelle, and Luca Maria Gambardella.**AntHocNet: an adaptive nature-inspired algorithm for routing in mobile ad hoc networks. European Transactions on Telecommunications 16,**2005, pp. 443-455.
16. Osama H. Hussein, Tarek N. Saadawi, and Myung Jong Lee.**Probability routing algorithm for mobile ad hoc networks' resources management.Selected Areas in Communications, IEEE Journal**, 2005, pp. 2248-2259.
17. ShahabKamali, and Jaroslav Opatrny.**Posant: A position based ant colony routing algorithm for mobile ad-hoc networks.Wireless and Mobile Communications,** 2007, pp. 21-21.
18. R.Asokan, A. M. Natarajan, and C. Venkatesh.**Ant based dynamic source routing protocol to support multiple quality of service (QoS) metrics in mobile ad hoc networks.International Journal of Computer Science and Security 2, no. 3,**2008, pp. 48-56.
19. P. Deepalakshmi, and S. Radhakrishnan.**Ant colony based QoS routing algorithm for mobile ad hoc networks.International Journal of Recent Trends in Engineering 1, no. 1,**2009, pp. 459-462.
20. Shivanajay Marwaha, Jadwiga Indulska, and Marius Portmann.Biologically inspired ant-based routing in mobile ad hoc networks (MANET): a survey.Symposia and workshops on ubiquitous, autonomic and trusted computing, IEEE, 2009, pp. 12-15.
21. Ehsan Khosrowshahi Asl, MortezaDamanafshan, MaghsoudAbbaspour, Majid Noorhosseini, and Kamran Shekoufandeh.EMP-DSR: An enhanced multi-path dynamic source routing algorithm for MANETs based on ant colony optimization. Third Asia International Conference on Modelling & Simulation, 2009, pp. 692-697.
22. Jianping Wang, Eseosa Osagie, ParimalaThulasiraman, and Ruppa K. Thulasiram. HOPNET: A hybrid ant colony optimization routing algorithm for mobile ad hoc network. In: Ad Hoc Networks 7, no. 4, 2009, pp. 690-705.
23. Ahmed M., Abdel-Moniem, Marghny H. Mohamed, and Abdel-Rahman Hedar. An ant colony optimization algorithm for the mobile ad hoc network routing problem based on AODV protocol.10th International Conference on Intelligent Systems Design and Applications (ISDA), 2010, pp. 1332-1337.
24. Zheng-Yu Wu, and Han-Tao Song. **Ant-based energy-aware disjoint multipath routing algorithm for MANETs.The Computer Journal 53, no. 2,**2010, pp. 166-176.

Biographies

Rashmi Mishra completed her B. Tech. (CS&E) from Uttar Pradesh Technical University in year 2013 and the master's degree in computer engineering from Madan Mohan Malviya University of Technology in 2016. She is currently working as an Assistant Professor at the Department of Computer Science and Engineering in Tula's Institute Dehradun. Her research areas include Mobile Ad hoc Networks and IoT.

A Comparative Study on Association Rule Mining and Its Preliminaries

Aditya Shukla[1], Pankaj Kumar Gond[2], Dr. Harvendra Kumar[3]

[1]B.Tech 4th Year, Dept. of Information Technology, ITM Gorakhpur, U.P, India

[2]B.Tech 4th Year, Dept. of Information Technology, ITM Gorakhpur, U.P, India

[3]Associate Professor, Dept. of Computer Science and Engineering, ITM Gorakhpur, U.P, India

harvendra.patel81@gmail.com

Abstract.

The process of extracting and identifying nuggets of findings from outsized amounts of records is known as "data mining (DM)". It consists of several approaches, such as clustering, data summarization, association mining, and classification. Association Rule Mining (ARM), in particular, aims to extract common patterns, interesting connections, associations, or structures that can be adjusted between sets of objects or other statistics. This technique plays a significant part in the route of refining strong rules to demonstrate the stable association between several itemset present in the database. With combined rule mines, different types of techniques and measures have been designed, but it is important to know which way is the best to extract appropriate association rules. Therefore, in this document, we assess the procedures used in an ARM to test the ability to extract high-dimensional data. The paper contains the following sections: section first is the introduction, section second is literature survey, section third is preliminary concepts, section fourth is DM and DM tasks, section fifth is association rule mining, and the last section is the conclusion.

Keywords Data Mining, Frequent Itemset, Data mining Model, Association rule mining

1. INTRODUCTION

DM is the methodology of finding sensible, new patterns related to trends by filtering large amounts of archived data using Pattern Recognition (PR) techniques and mathematical techniques. According to researchers, the two important DM models are predictive and descriptive. The predictive uses a variety of available databases to predict unknown results, while the descriptive focus on finding patterns that define data. Each model for further classification has four functions, as represented in figure 1. Out of these eight functions, the ARM is the most commonly used DM function for studying trends or patterns in a dataset. It also provides rules which help in understanding customers' behavior. Nevertheless, these document sprits around a relative study of different measures of ARM. The measures are support, lift, confidence, and conviction.

This document presents a relative study of the principle and methods used in each measure. The representations provided in this document provide a further understanding of the effectiveness of the measures. Tests are performed on these scales and the results are seen in the provision of time to act and use memory in addition to the patterns produced by it.

2. LITERATURE SURVEY

K. Solanki Surbhi and T. Patel Jalpa [1], write about the trouble faced by individuals while using the frequent pattern mining strategies. This Mining strategy work by examining the database several times, which eventually results in higher process costs. Not only this, this strategy leads to the production of candidate itemsets, which ultimately requires more memory and becomes more sophisticated in handling when the database is outsized. So, to reduce the downside of this problem of candidate $s$et generation, a tree-based strategy came for mining periodic patterns. However, the tree-based strategy produces numerous conditional fp trees. So, in progression to vanquish this problem of producing numerous based fp-trees, fp is DAG for normal pattern mining is enhanced, whereas fp-tree is constructed as a DAG. In addition, to find effectual mining, fuzzy deviation could be applied to the assayable database to provide the optimal patterns.

P. Amaranatha Reddy and MHM Krishna Prasad [2], on how to obtain ARs for the differing types of data-objects and requisitions involved in them. Some of the ways of data sets reasoned in this document are Boolean, Quantitative, weighted, time-series, stream data, infrequent, Diversified, and fuzzy item-sets. So the idea of differentiating ARs amongst correlated items like milk and banana will not be considered enterprise comprehension but differentiating unknown ARs between distinct items such as liquor and diapers will be beneficent in enterprises expansion and finding such an unknown category of ARs requires in-depth information about the data.

Mrs. Geeta S. Navale and Drs. Suresh N. Mali [3] proposed a variety of methods to hide association rules on the database and to develop support and confidence measures. The intention of this paper is to conceal the critical association rules of the DM on the following conditions: no production of false rules, no loss of information, Modification Degree and robustness against intentional or unintentional attacks. Except for the three conditions above, the proposed method will be compatible with a measurable database. Therefore, the proposed method of data encryption of the operating system will be evaluated and verified with regard to the various parameters as set out in the conditions and it is necessary to assess the satisfaction of the above-mentioned conditions.

L Greeshma and Dr. G Pradeepini [4] focused on developing the latest Apriori-*b*ased algorithm, which satisfies positive aspects of constrained *i*temset based mining such as anti-*m*onotonicity. The problem of ARM is to retrieve relevant itemsets for which it presents a new constraint, called relation-based constraints, applicable to relevant data. In the CIM algorithm, it helps us to recognize the main components of a candidates' key itemsets and generate frequent itemsets, which satisfy the anti-*m*onotonicity properties, which means small coverage and cardinal size limited to a particular dataset.

Table 1: A comparative study of algorithms used in DM:

S. NO.	Algorithm Name	Application	Advantages	Disadvantages	Year
1.	AIS	Not frequently used, but when used is used for small problems.	1. Better than SETM. 2. Easy to use	1. Candidate sets generated on the fly. 2. Size of candidate set large.	1994
2.	SETM	Not frequently used.	1. Separates generation from counting.	1. Very large execution time and the size of candidate set is large.	1994
3.	Apriori	Best for closed item sets.	1. Less candidate sets. 2. Generates candidate sets from only those items that were found large.	1. Takes a lot of memory.	2003
4.	AprioriTID	Used for smaller problems.	1. Better than SETM, Apriori for small databases. 2. Fast & Time saving	--	2013
5.	Apriori Hybrid	Used where Apriori and AprioriTID used.	1. Better than both Apriori and AprioriTID.	--	2013
6.	Eclat	Best used for free item sets.	1. Less memory usage. 2. Lower minimum support.	1. Apriori wins in cases where candidate sets are more.	2004
7.	Recursive Elimination	Effectively select the most relevant itemset.	1. Better than Apriori in all cases.	1. Less than éclat in all cases.	2005
8.	FP Growth	Used in cases of large problems as it doesn't require generation of candidate sets.	1. Only 2 passes requires. 2. No candidate set generation required.	1. Using tree structure creates complexity.	2003

P. Naresh and Dr.R. Suguna [5] explain a relative study of four ARM algorithms, namely $Apriori$, $FPGrowth$, LCM and FIN. These algorithms are discovered in terms of their purpose, the way they instigate recurring itemsets, and their appearance on the organization's data. The performances of all these algorithms are evaluated by means of time. The algorithms presented in the DM configuration, namely "Sequential Pattern Mining Framework" are used to make an investigational assessment of the algorithms. These investigational results on these algorithms disclosed that the amount of data has a great significance on the implementation time and at last these led to the conclusion that: the FIN algorithm displays the least implementation time while $FPGrowth$ shows minimum memory utilization.

3. PRELIMINARIES

1) Frequent Set: T is a transactional database and σ is the "minimum support threshold" specified by the user or domain experts. An itemset is frequent if they satisfy the min support threshold,

$$s(M)_{T\geq\sigma}$$

2) Maximal Frequent Set: In order to be a "$maximal\ frequent\ set$", a frequent set must be recurring and no superordinate of it must be recurring.

3) Support(s): In a database D, s is the proportions of agreements that comprise both M and N itemsets. In s of an ARs $M \longrightarrow N$, there is,

$$supp(M \longrightarrow N) = supp(M \cup N) = P(M \cup N)$$

4) Confidence(c): In database D, this is the percentage of agreements that contain itemsets M and N. 'c' is calculated by considering the conditional probability as well as the itemset support. Confidence can be calculated using the equation,

$$conf\ (M \longrightarrow N) = P(N/M) = supp\ (M\ U\ N)\ /\ supp\ (M)$$

Here, $supp(M \cup N)$ indicates the number of agreements considering both sets of items M and N, and $supp(M)$ indicates the number of agreements considering just set M.

5) Lift: It is used to analyze the frequency M and N together, if both are precisely in different. The lift of rule $M \longrightarrow N$ is defined as,

$$lift\ (M \longrightarrow N) = \frac{confidence}{expected\ confidence} = \frac{conf(M \longrightarrow N)}{supp(N)}$$

6) Conviction: Conviction analyses the implication stability of the rule from statistical independent. Conviction is defined as,

$$conv\ (M \longrightarrow N) = \frac{1 - supp(N)}{1 - conf(M \longrightarrow N)} = \frac{P(M) * P(\overline{N})}{P(M \cup \overline{N})}$$

Where $P(\overline{N})$ is the probability that N doesn't appear in an agreement.
It compares the probability that M appears without N if they were dependent on the actual frequency of the appearance of M without N [6, 7].

4. DM AND DM TASKS

DM can often be categorized into two categories based on what a particular project is trying to accomplish. Those two categories are descriptive model and predictive model. There are a number of DM functions, such as ARM, Time Series Analysis, prediction, Neural Network, etc. Either of these functions falls under the predictive model or descriptive model. The DM system can perform one or more of the above functions as part of the DM.

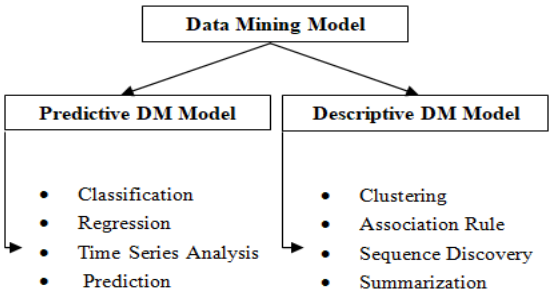


Fig 1: DM Model and its Tasks

4.1 Descriptive Model: The descriptive model defines a domain that represents a way that can be described or find the relationship among data. It can be used for many purposes. It may include ethical, structural, and other definitions that establish reasonable relationships about the system, such as its component relationships, the interactions between its components, and the distribution of its ethical properties to structural elements. Descriptive models are not usually constructed in a way that directly supports imitation, animation or performance, but can be considered compliance with grammatical rules, and logical relationships can be considered for them. This model basically relies upon an unsupervised learning approach. Some of the major tasks of descriptive models are as follows: Association Rule, Clustering, Sequence Discovery, and Summarization.

4.2 Predictive Model: The Predictive-model is a mathematical method that is generally used by DM technologies to define $m$odeling prediction as to the process of predicting the subsequent time behavior of an analysis of factual and recent data. A predictive model makes theories based on what has previously occurred and what is currently occurring. If new data shows that the current situation has changed, the likelihood of the outcome must be recalculated. This model relies upon the supervised learning approach. Some major tasks of predictive models are regression, classification, prediction, and Time Series Analysis.

5. ASSOCIATION RULE MINING

ARM [3] is an event for determining organizations, patterns, and relationships between sets of objects on dataset. The law of association is form $M \longrightarrow N$ [support, confidence]. The support and confidence are two measures that are used for assurance of the rule. AR is said to be strong if it satisfies both the minsupp(minimal support) and the minconf(minimal confidence) that is defined by the user. These ARs are facile to establish due to the $s$mall database but become more complex as the database transforms. Some general values and concepts are required in order to better understand DM in large data sets. A set contains n items and is called an n-itemset. So set {A, B} is a set of 2- itemset. Based on the frequency of itemsets, the number of active functions is calculated. So now, let us assume the value of minsupp=0.25 and minconf=0.10.
The dilemma of mining ARs can be fragmented into two sub-dilemmas:

1) Find all sets of itemsets whose support is greater thanσ. These itemsets are known as $frequent\ itemsets$.
2) Use these $frequent\ itemsets$ to develop the desired rules. The conventional ideology is that if, say p, q, r, s are $frequent\ itemsets$, then we can decide on the rule $pq \rightarrow rs$ that holds by checking the following inequality

$$\frac{\mathrm{s}(\{\mathrm{p},\mathrm{q},\mathrm{r},\mathrm{s}\})}{\mathrm{n}} \geq \sigma$$

Where n is the total number of transactions and σ is a minsupp.

Measures for Association Rule Mining

Table2 presents Transactional Super Market Data

TID	Items
T_1	A_1, A_2, A_3
T_2	A_1, A_2, A_4
T_3	A_1, A_3, A_5
T_4	A_2, A_4, A_6
T_5	A_3, A_6, A_7
T_6	A_1, A_2, A_3, A_4
T_7	A_2, A_3, A_4
T_8	A_3, A_7
T_9	A_2, A_3, A_4, A_8
T_{10}	A_6, A_7

Here: Ten transactions and eight items in a transactional dataset are shown in table 1.

5.1 $\boldsymbol{Support(s)}$: $Support$ is measured as the amount of logs that contain $M \cup N$ from the entire logs in the database. The proportion for each item is altercation by one, when so ever the item is crossover in dissimilar transaction in a database across the period of scanning.
$Support$ is used to measure the quantity or frequency of an itemset in a database. This measure gives an idea of how common itemset is in all activities. Support can be measured as:

$$Supp(M \longrightarrow N) = P(MN) = Freq\frac{(M,N)}{n}$$

Where: Freq (M, N) = Transaction containing M and N, and n = Total number of transitions.
It helps us to identify the rules that need to be considered for further analysis or not. If support of the rule is greater than minsupp, then find the confidence of the rule.

$$Support\ (A_2, A_4 \longrightarrow A_3) = 3/10 = 0.3$$

Thus the support value of items A_2, A_4 & A_3 is 0.3.

5.2 **Confidence(c):** Confidence is defined as the [9] amount of the numbers of transactions that contains $M \cup N$ to the entire logs that contain M, where, if the fraction exceeds the kick-off of $confidence$, an ARs $M \longrightarrow N$ can be obtained. Confidence is an indication of how often the rule is true.
Confidence $(M \longrightarrow N)$ in relation to the set of functions n, the part of the functions containing M and N.

$$conf\ (M \longrightarrow N) = \frac{supp(M\ U\ N)}{supp\ (M)}$$

Confidence explains how N is frequently occur when already buying M. This describes the link between two things. For example, if a person buys jam there is a good chance to buy bread. It is calculated as part of the number of operations where both M and N occur to support the M object.

$$Supp(A_2, A_4 \longrightarrow A_3) = 3/10 = 0.3$$
$$Supp(A_2, A_4) = 5/10 = 0.5$$
$$conf\ (A_2, A_4 \longrightarrow A_3) = 0.3/0.5 = 0.6$$

An association rule $M \longrightarrow N$ will be strong if, $conf(M \longrightarrow N) \geq minconf)$ and here $conf(A_2,\ A_4 \longrightarrow A_3)$ is 0.6 which is greater than the minconf, therefore the rule can be reasoned as a strong rule because it met with minsupp and minconf conditions. But we need to check it further than support and confidence alone cannot be sufficient to find a strong rule. In the above association rule($A_2,\ A_4 \longrightarrow A_3$), support of the consequent $(s(A_3) = \frac{7}{10} = 0.7)$ is greater than the confidence of the rule (0.6). This is not feasible. Therefore, this may be a misleading rule. Misleading rules can be generated from irrelevant datasets. Therefore, additional steps are needed to avoid misleading rules.

So to solve this problem of misleading rules another two measures can be used i.e. lift and conviction.

5.3 **Lift:** The Lift [10] is defined as one of the measures of ARM which define how far the inter-dependence in between M and N. The measure lift is not sensitive to rule i.e. $(lift(M \longrightarrow N) = lift(N \longrightarrow M))$. A Lift could be formulated as:

$$lift\ (\mathrm{M} \longrightarrow \mathrm{N}) = \frac{conf(M \longrightarrow N)}{supp(N)} = \frac{supp(MUN)}{\mathrm{supp}(M) * supp(N)}$$

So,

$$lift(A_2, A_4 \longrightarrow A_3) = \frac{0.3}{0.5 * 0.7} = 0.35 < 1$$

The fraction of the observance $support$ and the predicted $support$ if M and N are free for each other is known as lift. It has three possible values:

- If Lift = 1, the probability of occurrence and outcome are independent of each other.
- If Lift >1, then the itemsets are dependent on each other.
- If Lift < l, tells us that one thing replaces other things, which means one thing has a negative effect on another thing.

An ARs $M \longrightarrow N$ is engaging if it is $strong$ and $lift(M \longrightarrow N) > 1$.

Here as $lift(M \longrightarrow N) < 1$, thus the rule is not valid to consider it as a strong rule.

5.4 **Conviction:** The conviction is defined as one of the measures of ARM which undertake to analyze the magnitude of execution of the rule, to evaluate the conviction. [8, 11].

Unlike Lift, Conviction is tactful to rule direction i.e. $(conv\ (M \longrightarrow N) \neq conv\ (M \longrightarrow N))$.

A Large conviction value shows that the obtained result is largely relying on the predecessor. The conviction can be formulated as:

$$conv(M \longrightarrow Y) = \frac{1-supp(N)}{1-conf(M \longrightarrow N)} = (1 - supp\ (N))/(1 - conf(M \longrightarrow N))$$

It correlates the likeliness that M exists without N when they are relying on the factual frequentness of the existence of M without N. In that scenario, it is similar to $lift$. However, $conviction$ have monotonousness in $confidence$ and $lift$.

So,

$$conv(A_2, A_4 \longrightarrow A_3) = \frac{1 - 0.7}{1 - 0.6} = 0.75$$

$$conv(A_3 \longrightarrow A_2, A_4) = \frac{1 - 0.5}{1 - 0.42} = 0.86$$

Here, as the value of $conv(A_2, A_4 \longrightarrow A_3)$ has a value less than the value of $conv(A_3 \longrightarrow A_2,\ A_4)$, therefore the $conv(A_3 \longrightarrow A_2, A_4)$ can be considered as a strong rule.

Table 3: Quality measures and range of feasible values:

Name	Equation	Feasible values
Support	P_{MN}	[0,1]
Confidence	$\frac{P_{MN}}{P_M}$	[0,1]
Lift	$\frac{P_{MN}}{P_M * P_N}$	[0,1]
Conviction	$\frac{P_M * P_{\bar{N}}}{P_{M\bar{N}}}$	$\left[\frac{1}{n}, \frac{n}{4}\right]$

6. CONCLUSION

In this paper, there is a preliminary of the DM and a detailed discussion on ARM. ARM is facing the problem of finding the most efficient and strong rules which are suitable for any dataset due to the presence of numerous rules. In most literature, the fascinating steps of governance in ARM algorithms are based on $support$ and $confidence$. Depending on the types of application, different measures could be used to compute the interestingrules. As described in the above sections of the paper, the main issues are to trust the method for support is low forecasting capability and similar support issues. Whereas, the previous work retains presuming solutions to these issues, as an add-on the rate of promotion or sentencing and the use of an unusual support barrier, where it remains without guidance in defining support. Without such an order, users may set the wrong support limit and suffer from combinatorial explosion or loss of new cognitive arrangement. Our plan for constructing the solution for this issue is to exclude clients from determining a support limit. Therefore, we need to refine the rules in sort to obtain strong rules that must satisfy the following parameters.
Association rules will be strong if they satisfy the following conditions: (i) $supp(M \longrightarrow N) \geq minsupp,$
(ii) $conf(M \longrightarrow N) \geq minconf$ (iii) $lift(M \longrightarrow N) > 1$ and higher $conv(M \longrightarrow N)$ value.

7. REFERENCES

1. Surbhi K. Solanki and Jalpa T. Patel, “A Survey on Association Rule Mining”, Fifth International Conference on Advanced Computing & Communication Technologies, Volume 5, pp. 212-216, 2015.
2. P. Amaranatha Reddy and MHM Krishna Prasad, “Challenges to find Association Rules over various types of data items: a Survey”, International Conference on Computing, Communication and Automation, pp. 180-184, 2017.
3. Mrs. Geeta S. Navale and Drs. Suresh N. Mali, “A Survey on Sensitive Association Rules Hiding Methods”, Third International Conference on Computing, Communication, Control And Automation (ICCUBEA), IEEE, 2017.
4. L Greeshma and Dr. G Pradeepini, “Unique Constraint Frequent Item Set Mining”, 6th International Conference on Advanced Computing, IEEE, pp. 68-72, 2016.
5. P.Naresh and Dr.R.Suguna, "Association Rule Mining Algorithms on Large and Small Datasets: A Comparative Study", Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 587-592, 2019.
6. Hemant Kumar Soni, “Multi-objective Association Rule Mining using Evolutionary Algorithm”, IJARCSSE, Volume 7, Issue 5, May 2017.
7. H. K. Soni et al., “Frequent Pattern Generation Algorithms for Association Rule Mining: Strength and Challenges”, IEEE International Conference on Electrical, Electronics and Optimization Techniques (ICEEOT), pp. 3744-3747, 2016.
8. Dinesh J. Prajapati et al., “Interesting association rule mining with consistent and inconsistent rule detection from big sales data in distributed environment”, Future Computing and Informatics Journal 2, pp. 19-30, 2017.
9. J. M. Luna et al., “Optimization of quality measures in association rule mining: an empirical study”, International Journal of Computational Intelligence Systems, Volume 12, pp. 59-78, 2018.
10. H. K. Soni et al., “Association Rule Mining: A data profiling and prospective approach”, International Journal of Current Engineering and Scientific Research, Volume 3, pp. 57-60, 2016.
11. Memoona Khanum and Tahira Mahboob, “A Survey on Unsupervised Machine Learning Algorithms for automation, Classification and Maintenance”, International Journal of Computer Applications, volume 119-No.13, June 2015.
12. Sikha Bagui and Probal Chandra Dhar, "Positive and negative association rule mining in Hadoop’s MapReduce environment", Journal of Big Data, pp.1-16, 2019.
13. Lichun Li et al., “Privacy-Preserving-Outsourced Association Rule Mining on Vertically Partitioned Databases”, IEEE Transactions on Information Forensics AND Security, pp. 1-15, 2016.
14. Mandeep Mittala et al., “Loss profit estimation using association rule mining with clustering”, Management Science Letters, pp. 167–174, 2015.

Machine Learning in Cloud Securing data and Cloud Cryptography

Ramnarayan[1], Akansha singh[2], Sandeep Sharma[3], Sakshi koli[4], Anuj Kumar[5], Kapil Joshi[6]

Assistant Professor CSE Department UIT Uttaranchal University Dehradun[1, 5,6], Assistant Professor CSE Department MIET Meerut[2] , Assistant Professor CSE Department Tula's Institute Dehradun[3,4]

ram000010@rediffmail.com[1], Akansha.Singh@miet.ac.in[2], sanintel123@gmail.com[3], Kolisakshi84@gmail.com[4], kannojia.anuj@gmail.com[5], kapilengg0509@gmail.com[6]

Abstract

Cloud computing is the next upcoming era for data storage and abstraction area for next generation. Cloud is a very complex system so its security is major issue so we need to develop the technologies to secure. The intricacy of cloud computing make many issues connected with security just as all parts of Cloud processing. So the most important is information security at correct part of the cloud. Understanding the associations and interdependencies between the different circulated registering sending models and organization models is essential to understanding the security dangers suggested in cloud computing. Machine learning play very important role with different existing algorithms because these technologies are playing lead role to automate and secure to everything in this field. Cloud security can be used at all the level of cloud today.

Keywords: Cloud computing, Architecture, Modern cloud security, ML in cloud.

1.0 Introduction

Cloud computing is an extremely intricate region. It is giving a high reach office to save and deal with the information at every one of the levels. So this is vital to get the information utilizing some ML, AI and IOT based innovations.The upsides of cloud computing conveyed registering in-cloud reasonableness, adaptability, and moderateness. Besides, conveyed capacity has establishment characteristics on demand, economy, universality, ease, leasing pluralism, reliability, and adaptability.Cloud computing's widespread acceptance will be hampered by security concerns. In fact, the difficulty of keeping disseminated registered organizations secure and protected from unwanted access or usage is addressed by sharing them. [1] IAAS deal with the hardware to facilitate at low or high level of the user so that PAAS provide the platform for better service, similarly SAAS provide the different type of software facility to access the data or to use the data in better ways.

1.1 Private Cloud

Private Cloud processing is runs and administered inside the server ranch of an affiliation, which is suggested as a private cloud. [2]Because the system is owned and operated by a similar organization, customer and supplier relationships are easy to discern in a private cloud.

1.2 Public cloud

Public cloud endeavors, the academic world or government affiliations have a public cloud environment, which can make many issues since customers don't have even the remotest clue about the areas then again owners of resources, which assembles the difficulty of safeguarding resources from attacks. [3]

2.0 Cloud Architecture

Cloud computing technology is an integration of different technology and components. Cloud architecture include two phases, one is front end and second is back end. The main components are SAAS, PAAS, and IAAS. [3]

SAAS	Google app, zoho and salesforce.com etc.
PAAS	Google app engine, Aptana cloud and windows azure etc.
IAAS	Web services, mozy, Akamai, amazon and drop box etc.

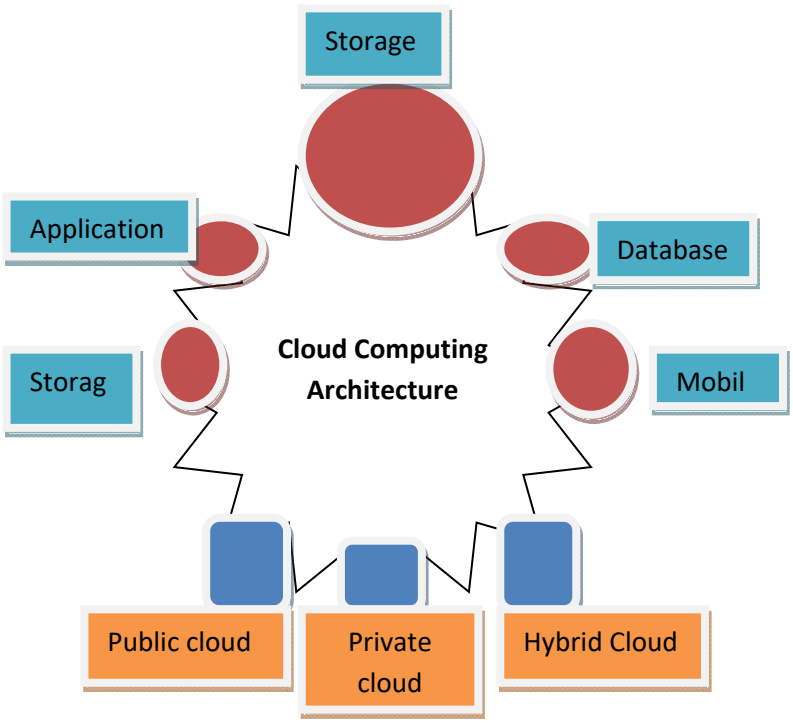


Fig 1: Architecture of cloud computing

2.1 Key Terms:

- Cloud Infrastructure
- Internet
- Application
- Service
- Storage
- Management
- Security

There are different sides of the cloud computing. The front end is what's apparent to the end client; at the end of the day, it's the UI. The back-end foundation runs the cloud. This back end is comprised of server farm equipment, virtualization, applications, and administrations. The front end speaks with the back end through middleware. While there are varieties of cloud engineering in view of how you're attempting to treat, mists require equipment, middleware, the executives, and robot programming. [4] Most mists additionally use virtualization to digest the equipment assets into halfway oversaw information lakes, while certain mists known as uncovered metal mists associate customers straightforwardly to equipment. [15]

3.0 Security on clouds

Security of the data on cloud is a critical issue for all the organization. There are many methods to protect the data for different servers. [5] Most of the major point regarding security of data on cloud are data integration, confidentiality, reliability, scalability, authentication and protection of data on cloud.

3.1 Cloud security and challenges

3.1.1 Data Protection and prevention

Data security is an important component that should be examined. Attempts are hesitant to purchase a merchant's assurance of corporate data security. They are concerned about losing data in the event of a challenge, as well as the data order of buyers. The true amassing region isn't provided in many models, which adds to the security concerns of undertakings. [6]Firewalls across server ranches (asserted by tries) ensure this sensitive information in current models. Service providers are at danger in the cloud model for remaining cognizant of data security, and tries would have to rely on them. [13]

3.1.2 Data Recovery and availability

All business applications have Service level game plans that are unbendingly followed. Utilitarian gatherings expect a basic part in organization of organization level plans and runtime organization of employments. In progress conditions, utilitarian gatherings support [14]

- Replication of data
- Recovery of data
- Management of data
- Monitoring of data

3.1.3 Administrative and Compliance Restrictions

Government restrictions in some European countries prohibit customers' personal information and other sensitive data from being shared outside of the state or country. To comply with such requirements, cloud companies must plan a server ranch or a limited-access site just within the country. [7] Having such a structure may not perpetually be conceivable and is truly hard for cloud providers.

3.1.4 The board Capabilities

Despite the fact that there are numerous cloud providers, stage and foundation administration is still in its early stages. Features like as "Auto-scaling," for example, are a must-have for some projects. The flexibility and burden adjustment characteristics mentioned today have a lot of room to grow. [8]

3.2 Symmetric view to secure the cloud- Machine learning

Machine Learning intelligence, a system and set of developments that usage AI thoughts, is clearly associated with plan affirmation and computational learning. It's an old thought, first described in 1959 as empowering PCs to learn without reproducing. Man-made intelligence was once out of the range of most endeavor spending plans, yet today, public cloud providers' ability to offer AI organizations makes this development sensible. [9] I should bring you outstanding on AI and its importance to the current IT improvement and sending needs, especially for those working inside a cloud environment. More applications and methods can be relevant in terms of industry 4.0 [16] [17] [18] [19].

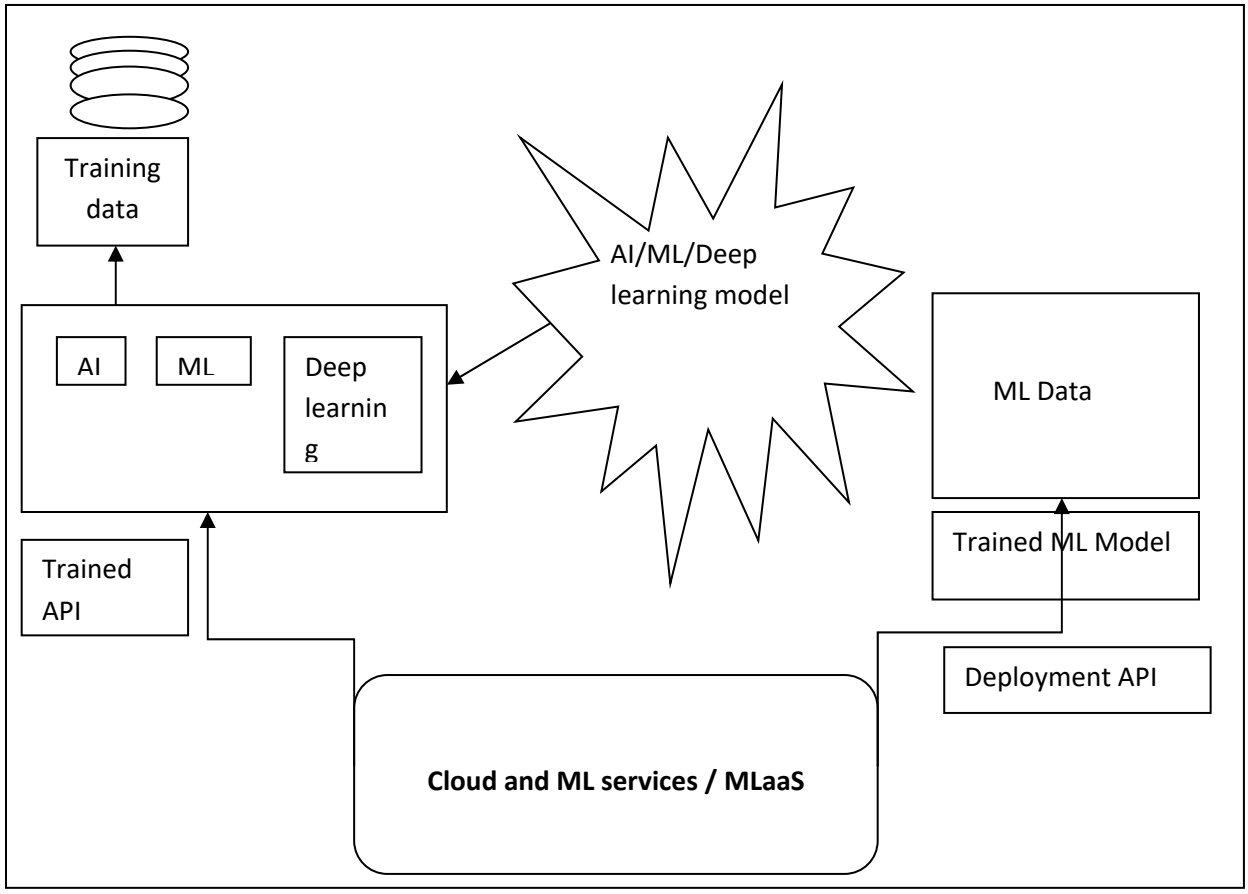


Fig 2: Role of ML in cloud computing security

4.0 Some prediction of different organizations for cloud computing:

4.1 Binary Prediction

This kind of ML forecast manages "yes" or "no" reactions. It is essentially utilized for extortion discovery, proposal motors, and request handling, to give some examples. [12]

4.2 Category prediction

It this kind of expectation, a dataset is noticed and in view of the accumulated data from it, the dataset it set under a particular classification. [10] For example, insurance agencies use classification forecast to order various sorts of cases.

4.3 Value prediction

This sort of expectation tracks down designs inside the aggregated information by utilizing learning models to show the quantitative proportion of the relative multitude of likely results. Organizations use it to anticipate an unpleasant number of the number of units of an item will sell sooner rather than later (e.g., the following month). [11] It permits them to shape their assembling plans likewise.

5.0 Conclusion and future Aspects

Cloud computing is the deep and future serving technology that can use for all the data for data servers and services. This is a great challenge to secure the Meta data using securing technologies. The information security assumes an essential part in customer side and furthermore in cloud supplier side. The significant security issue that we have examined in this paper was information security. The greater part of individuals use cloud to save their information because of its adaptability.

Those associations with practically zero distant innovation framework set up experienced the most effect because of the pandemic. However, one of the positive results of COVID-19 is it has assisted associations with seeing how they can carry on with work all the more successfully and effectively with innovation. Generally speaking, this has made most organizations more vigorous and tough to future interruptions. What has empowered this change? Cloud.

Now the technology have to implemented on "however, what security challenges has this colossal shift to cloud and remote work caused? Furthermore how does this influence the eventual fate of work?".

References:

[1] Dinh, P.T.; Park, M. Dynamic Economic-Denial-of-Sustainability (EDoS) Detection in SDN-based Cloud. In Proceedings of the 2020 Fifth International Conference on Fog and Mobile Edge Computing (FMEC), Paris, France, 20–23 April 2020.

[2] Han, J.; Zang, W.; Chen, S.; Yu, M. Reducing Security Risks of Clouds through VirtualMachine Placement. In Proceedings of theIFIP Annual Conference on Data and Applications Security and Privacy, Philadelphia, PA, USA, 19–21 July 2017.

[3] Saravanan, N.; Umamakeswari, A. Lattice based access control for protecting user data in cloud environments with hybridsecurity. Comput. Secur. 2021, 100, 102074.

[4] Rao, R.V.; Selvamani, K. Data Security Challenges and Its Solutions in Cloud Computing. Procedia Comput. Sci. 2015, 48, 204–209. [CrossRef]

[5] Wani, A.R.; Rana, Q.P.; Saxena, U.; Pandey, N. Analysis and Detection of DDoS Attacks on Cloud Computing Environment usingMachine Learning Techniques. In Proceedings of the 2019 Amity International Conference on Artificial Intelligence (AICAI),Dubai, United Arab Emirates, 4–6 February 2019.

[6] Mohiuddin, I.; Almogren, A.; Alrubaian, M.; Al-Qurishi, M. Analysis of network issues and their impact on Cloud Storage. InProceedings of the 2019 2nd International Conference on Computer Applications & Information Security (ICCAIS), Riyadh, SaudiArabia, 1–3 May 2019.

[7] L. Tawalbeh, N.S. Darwazeh, R.S. Al-Qassas, and F. AlDosari, "A secure cloud com- putting model based on data classification, "2015, doi: 10.1016/j.procs.2015.05.150.

[8]S.S. Khan, P.R. Tuteja, Security in cloud computing using cryptographic algo- rithms, Int. J. Innov. Res. Comput. Commun. Eng. (2015), doi: 10.15680/ijir- cce.2015.0301035.

[9]D.P. Timothy, A.K. Santra, A hybrid cryptography algorithm for cloud computing se- curity, 2017 International conference on Microelectronic Devices, Circuits and Sys- tems(ICMDCS), Vellore (2017) 1–5, doi: 10.1109/ICMDCS.2017.8211728 .

[10] Kaur R, Kaur J. Cloud computing security issues and itssolution: A review. 2015 2ndInternational Conference onComputing for Sustainable Global Development (INDIACom);2015. p. 1198–200.

[11] Moyo T, Bhogal J. Investigating security issues in cloudcomputing. 2014 8th International Conference on Complex,Intelligent and Software Intensive Systems; 2014. p.141–6

[12] Narula S, Jain A, Prachi MS. Cloud computing security:Amazon web service. 2015 5thInternational Conference onAdvanced Computing and Communication Technologies;2015.p. 501–5.

[13] Vasanth C, Bhagawat B, Arul D, Kumar LS. Survey ondata security issues in cloudenvironment. IJIRAE. 2015;2(1):31–5.

[14] Cyril BR, Kumar SBR. Cloud computing data security issues,challenges, architecture andmethods - A survey. IRJET.2015; 2(4):848–57.
[15] Puthal D, Sahoo BPS, Mishra S, Swain S. Cloud computing features, issues and challenges:A big picture. International Conference on Computational Intelligence and Networks; 2015. p. 116–23.
[16] Singh, R., Gehlot, A., Rashid, M., Saxena, R., Akram, S. V, Alshamrani, S. S., & Alghamdi, A. S. (2021). Cloud server and internet of things assisted system for stress monitoring. Electronics (Switzerland), 10(24). https://doi.org/10.3390/electronics10243133.
[17] Goel, G., Tiwari, R., Rishiwal, V., & Upadhyay, S. (2018). Data preservation by hash algorithm for matrix multiplication over venomous cloud. PDGC 2018 - 2018 5th International Conference on Parallel, Distributed and Grid Computing, 210–214. https://doi.org/10.1109/PDGC.2018.8745851.
[18] Pandey, N. K., Chaudhary, S., & Joshi, N. K. (2017). Resource allocation strategies used in cloud computing: A critical analysis. 2nd International Conference on Communication, Control and Intelligent Systems, CCIS 2016, 213–216. https://doi.org/10.1109/CCIntelS.2016.7878233.
[19] Pathak, N., Anwar, S., Singhal, V., Sharma, N., & Shukla, A. K. (2019). Trends and augmentations of cloud computing. Proceedings of the 2019 6th International Conference on Computing for Sustainable Global Development, INDIACom 2019, 373–377. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85080966578&partnerID=40&md5=511517f503385b51b8b6d00d3b4cbe54.

Biographies

Ramnarayan received the bachelor's degree in Information Technology from Uttar Pradesh Technical University in 2010, the master's degree in Information Technology from Graphic Era University Dehradun in 2013, and perusing Ph.D. from Uttaranchal University Dehradun in 2022, respectively. He is currently working as an Assistant professor at the department of computer science and engineering, faculty of Uttaranchal University. His research area include cloud computing, IOT, and Machine Learning.

Software Fault Prediction Models Using Soft Computing - A Systematic Review

Gurmeet Kaur[1], Dr. Jyoti Pruthi[2], Dr. Parul Gandhi[3]

[1]*Research Scholar CST, MRU, India, grmtkaur02@gmail.com*
[2]*Professor CST, MRU, India, jyoti@mru.edu.in*
[3]*Professor FCA, MRIIRS, India, parul.fca@mriu.edu*

Abstract.

The most dynamic exploration zone in software engineering is software fault prediction because it offers the benefits regarding time complexity, low cost budget, testing effort, and increases the reliability along with the quality of the software if it is applied at the starting phase of traditional and agile based software development life cycle. This study of literature review is conducted in a systematic manner to understand the trends and techniques used for software fault prediction (SFP) problem and synthesis the qualitative results to present technical and methodological information, success and usefulness of SFP model. From this study, it analyzed that Neuro-Fuzzy as a soft computing methodology provides more accurate result for prediction of faults at the initial and later stage of software development, using a variety of metrics for feature selection process, identified datasets and reporting the statistical significance of different accuracy measure parameters, and comparing the performance of existing models with designed model.

Keywords. Software Fault Prediction, Agile Software Development, Soft computing, Software Metrics, Software defects.

1. INTRODUCTION

The Software Quality Assurance and Software Reliability is the key to guarantee the superior quality of software. Both these concepts are attracted all throughout the development of the software and measure[1]. A software bug is a defect, error, failure, or flaw during the execution of a code of software that permits it from acting as unexpected function (e.g., delivering an erroneous outcome). The software defect may be a defect that gives rise to failure of software functionally. This failure may occur due to the existence of one quiet fault in the software. Defective software modules cause software failures, reduce customer satisfaction, increased development, and maintenance costs. A software fault prediction (SFP) is frequently suggested because it incorporates the activities inside the development interaction which assists with anticipating the defective modules or data at initial stages of the development of software. The forecasting of defective modules at the initial stage makes the process of testing simple and quick. With the expansion to the present, it likewise improves the norm of programming.

With the fast progression of the software business, the use of the agile development technique proposed in recent years stress on timely reacting to the changes in prerequisites

as describing the insufficiency of the conventional programming development process. Agile development involves different structures or systems, specifically Scrum, XP, Crystal strategies, Lean Software Development, Feature Driven Development and so on. Agile-based process effectively deals with the truth of variation.

Estimating performance and project progress is a motivating field in programming practice. Inside the traditional programming development process, programming measurements are gathered into product metrics and process metrics. To evaluate the complexity of product, TSD measure normally utilizes Cyclomatic Complexity Metric (CCM), Halstead Complexity Metric (HCM), and Lines of Codes (LOC. Defect Density estimates the defect per function point or defect per KLOC and might be a product reliability metric. Defect Removal Efficiency is one among the significant estimations of programming standard. There are various kinds of soft computing modelling based approaches like a Fuzzy Logic System and Neural Network employed for Software Fault Prediction.

In recent years, various researchers try to automate the fault prediction process by designing computer-based models that can perform learning from existing prediction data. The objective of this research is to take correct decisions on the development of the prediction models on basis of finest knowledge and execution of many previous related studies. The outcomes of this research paper explore the existing fault prediction techniques for agile software. This research paper describes the review of literature in different ways by recognizing the basic studies on SFP and the key features of prediction models as follows:

- Prediction techniques: Based on soft computing techniques like Artifical Neural Network, Fuzzy Inference System, Machine learning techniques, and Neuro-fuzzy Hybrid system;
- Software metrics: Either Product and Process Metrics and Agile based metrics;
- Datasets: Public and Private data sets;
- Performance evaluation methods: Either Continuous and Categorical studies;

2. LITERATURE REVIEW

The goal of the systematic review is to understand technical works that describe modelling and identification of metrics to predict the fault for agile software development (ASD) methodologies in the area of prediction methods in soft computing. The research performed in the following four steps:

Step 1: Conduct manual and automatic search to get an initial list of studies. Always try to discard the duplicate studies.

Step 2: To identify the inherent relevant studies on the basis of title and abstract, and reject studies which are not related to the concern topic.

Step 3: Then, shortlist the research paper reviewed on the basis of the introduction, methodology section and conclusion. If reading of the selected research paper is not complete to confirm results, the study of the research paper read in detail.

Step 4: Finally, perform essential investigation using the esteemed standard of achieved and administered list of research paper.

This section describes a systematic method for the reviews of existing research paper on the fault prediction in source code by identifying research issues.

2.1 Research Issues

To explore the models applied to forecast faults in code, the following research questions are described as follows:

RI1. Which soft computing approaches/methods have used to predict fault for traditional and agile development software?

RI2: Which software metrics have been applied in the fault prediction models for traditional and agile development software?

RI3: What is the level of accuracy of fault prediction models?

2.2 Results of Concern Research

RI1. Which soft computing approaches/methods have used to predict fault for traditional as well as agile development software?

As the growing number of faults affects development time, cost and quality of a software package, so software fault prediction is the way of detecting faulty components in units in the prior of the implementation of the software. This study is predicated on systematic review that provides a comprehensive picture in the field of software fault prediction administered by many researchers for agile based software development.

Catal and Diri (2009) [2] described research on types of methodology or techniques used, metrics, and datasets for software fault prediction. The outcomes shows that the use of machine learning computation increased successively after 2005 then PROMISE data store was designed.

Hall (2012) [3] reported the application of feature selection algorithm and the combination of independent variables gives a better outcome in performance.

Begel, (2007) [4] studied based on an experiment performed to find out regarding agile based development, and its implementation in development, executives and testing by individuals, it had been found that 33% of the research respondents make the use of agile strategies. Around 65% of respondents used scrum in their software teams. The test-driven development and pair programming were the least used practices.

Table 1. Technical description of RI1

Ref. no.	**Objective**	**Methodology**	**Data Set/Features**	**Findings**	**Performance measurement tool**
[5]	The fuzzy inference system proposed to calculate the efficiency of metrics in defect predicting for agile software projects.	Neuro-Fuzzy hybrid approach used	From PROMISE repository	The proposed framework gives superior accuracy (specially for large size projects) as evaluated from performance measurement tools	MMRE, BMMRE, RMSE, NRMSE,
[6]	Proposed new variations of WOA as wrapper algorithms to deal with the feature (metric) selection issues in SFP applications.	Roulette wheel, Tournament, Linear position, and random based, stochastic universal sampling.	From the PROMISE archive.	The deep investigation presented that the suggested TBWOA (tournament approach) exceeded the principal WOA.	Average AUC and running time.

[7]	Proposed a system using inter-version and inter-project assessment, to recognize the product defect.	Based on fuzzy logic, the past assignments or adventure variations are taken for preparing sets, and, the present version or exercises are taken as testing sets.	From PROMISE store, even as on PDE and JDT adventures and other more nine open source adventures.	The assessment results concluded that the suggested system showed outstandingly good outcomes for Eclipse-PDE and Eclipse-JDT based projects.	AUC and GM, Root mean square error RMSE.
[8]	Proposed an exact framework to anticipate programming deficiency and assisted the feature selection algorithm for classification issues.	Proposed upgraded binary moth flame Optimization and adaptive synthetic sampling	It makes use of the wrapper feature selection and improves the data set.	ADASYN attempts to beat the imbalanced data issue while BMFO perform as a feature selection.	Average AUC
[9]	Depicted adequately anticipate the product defect proneness in the soft modules with a tendency to lessen programming maintenance cost.	A DDN (deep neural network) with BAPS (Bound particle swarm optimization) dimensional reduction suggested.	From real-world programming projects (NASA and Eclipse) was made use.	The analysis result showed that the BPSO-based dimensionality reduction innovation can enhance the organization structure and acquire better execution.	F-measure, area under the curve (AUC), and probability of detection (pd).

The study describes many techniques and approaches which are soft computing in nature and used for fault prediction in traditional and agile based software, but one of the study from [5] relates the field of research as Software fault prediction in the best manner. The author presented framework using 21 process metrics incorporating different phases in the life cycle of software development using fuzzy inference system to compute the number of faults and applied the back propagation algorithm to train the fuzzy set of rules to improve the accuracy of prediction and also validate the presented model using 29 projects from the PROMISE repository. The author computed diverse performance criteria such as MMRE (0.0539), BMMRE (0.0585), RMSE (18.69), and NRMSE (0.010). The observed and calculated values concluded that the presented model provides better fault prediction capability as comparing to models discussed in their literature.

There are three categories of datasets as Private, Public, and Unknown. Public datasets are available freely from PROMISE and NASA repository. Private Datasets normally belong to private organizations. The unknown datasets are those which are neither public nor private.

RI2: Which software metrics have been applied in the fault prediction models for traditional and agile development software? Metric can be classified into class-level, method/product-level, process-level, and component-level. Halstead (1977) and McCabe (1976) in 1970 proposed Method/Product-level metrics, and they are still in use.

"Product Metrics"

"Defect per Function Point": It can be calculated using Function Point, which helps in Software fault prediction.

"Defect per SKLOC": It can be helpful in Software fault prediction but this metric will be dependent on language used for implementation [10].

"Class-level metrics" It can be applied to object oriented programming only and proposed by Chidamber–Kemerer in 1994 and known as Chidamber & Kemerer (CK) metrics suite is still used by various researchers and software vendors. CK metric suite is perfect for prediction of fault in source code. These metrics are WMC, DIT, CBO, RFC, NOC, and LCOM [11].

"Process Metrics"

Process measurements incorporate the arrangement of measurements, which relies upon the qualities gathered over the life cycle of software development. The classes of process measurements are Code Churn, Requirement Metrics, Change Metrics, and Code Delta.

"Agile Software Development (ASD) based metrics" Changing needs is one among the principle issues that emerge inside the process of software development. To select software metric for Agile-based software, there is need for deep understanding to get the difference between the Agile-based process and Traditional Software Development (TSD), and their measurements methods.

"Task effort": The team produces and evaluates effort based on software professional hours for every task [12].

"Number of stories": The number of stories within the sprint can be basic, medium, and sophisticated on the basis of the story. The measurement is determined as a value in term of count or weight of the complexity of the story [13].

"Story point": Fibonacci format is used to describe the problem size of a story. There might be difficulty associated with uncertainty, efforts, and complexities involved in the process of measurement to forecast the software effort and size needed [13].

"Velocity": It is computed as the sum of all accepted works [13].

For example if for a project completion there are 5 sprints and story points completed from every sprints are 12,34,30,22,and17, then velocity is computed as follows:

$$\text{Velocity} = (12+34+30+22+17)/5 = 115/5 = 23$$

"Work capacity": The work capacity is the number of chores revealed during the run of the sprint, with respect to the finished feature. [13].

"Focus factor": The proportion among velocity and Work capacity, the great worth reaches approximate 80% for a group [13].

$$\text{Focus factor} = \text{Velocity} / \text{Work capacity} \quad (1)$$

"Open defect severity index": The defect evaluated at the completion of the sprint. This metric measures quality of feature for each and every iteration of the sprint [13].

"Defect per iteration": Calculate the entire defect that was presented during the sprint. The defect that was found, and will not be compile within the sprint [13].

"Error density": Number of mistakes detailed by the client after delivery (determined per sprint) [12].

C. Jones [10] (2008) the investigation inspected that there were a couple of difficulties in estimating the performance of the agile based development. On account of using the knowledge of the TSD Process, numerous project manager and developers are favoured to make use of a comparable measurement for the ASD process.

Padmini (2015) [12] the investigation was intended to gather data about the metric used in the agile at present. The investigation distinguished measurements like unit test coverage, ideal delivery, and defect seriousness index, bug correction time, in ASD measures.

Arisholm, E., L.C., Johannessen, Briand, E.B. 2010 [14] made sensible models with high inadequacy probability to recognize parts of a Java based structure. Because of the

extended demonstration of object-oriented development in companies, an extended use of object-oriented metrics has been proposed.
Briand L., Ikonomovski S., Wüst J., and Lounis H. (1998) [15] assembled an adequate data and information regarding the utilization of design measures. The author examined the coupling and cohesion estimation information gathered for the 83 framework classes. The author likewise concludes that some of the aspects to be estimated on the selected data sets: method innovation versus aggregate coupling, coupling to application classes versus library classes, and export versus import coupling.

RI3: What is the level of accuracy of fault prediction methods?

The prediction results of SFP model depend on and vary according to dependent variables. To measure the level of accuracy of the SFP model, there are various methods available that can be applied to analyze the performance of the designed SFP model. These measurements focused on predicting whether a given part of software is fault free or not.
Magnitude of Relative Error (MRE): The most commonly used accuracy metrics are the magnitude of relative error. It can be calculated on the basis of either the mean or the median. MRE is an always less than 1[5].

$$\text{MRE} = \frac{(\text{Actual defect} - \text{Predicted defect})}{\text{Actual defect}} \quad (2)$$

Mean MRE (MMRE): It is the average of magnitude of relative error values N projects. The problem of the Mean MRE is its sensitivity to anomaly.
Balanced MMRE (BMMRE): As MMRE is unbalanced, so for this reason BMMRE is used as [5][16][17]:

$$\text{BMMRE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|\text{Actual defect} - \text{Predicted defect}|}{\min(\text{Actual defect}, \text{Predicted defect})} \quad (3)$$

Table 2. Accuracy values of the prediction methods used.

Ref. No.	Prediction Techniques	Accuracy Model	Accuracy Value
[6]	Machine learning	Accuracy	0.982
[5]	Fuzzy Inference System	MMRE	0.0539
		BMMRE	0.0585
		RMSE	18.69
[7]	Neuro- Fuzzy System	RMSE	0.1266
		GM	0.8128
		Accuracy	84.4941
[16]	Artifical Neural Network	MMRE	0.0247
		BMMRE	0.0244
[17]	Fuzzy Inference System	MMRE	0.36343
	Artifical Neural Network	BMMRE	0.36534
		MMRE	0.015397
		BMMRE	0.007258

RMSE (Root MSE): The RMSE is computed as square root of the mean of squared difference between actual and observed/predicted defect values [5][7]. It is evaluated as:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\text{Actual defect} - \text{Predicted defect})^2} \quad (4)$$

Where, n denotes the extent of the actual dataset. It gives the quadratic based value to compute the average measure of the error.

Accuracy: It is applied to compute and equate the precision of prediction models, and calculated as the ratio of number of fault free predictions executed by the total number of prediction executed [7].

Accuracy = (True Positive +True Negative) / Total number of instances executed (5)

Where, True Positive –the number of fault free cases executed by the certain class of software and True Negative - the number of cases rejected by the certain class of project.

G-Mean: For imbalanced datasets, G-Mean is applied to compute the efficacy of the prediction method [7]. It is calculated as:

$$\text{G-Mean} = \sqrt{Precision * Recall} \quad (6)$$

$$\text{Where, Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad \text{and}$$

$$\text{Recall} = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

3. RESULT AND DISCUSSION

Since the predictability of fault is the essential step for development of any software, and this is also required for projects based on agile methodologies. The literature review study explore that there is very small quantity of efforts has been done in software fault prediction for agile-based software. Applying soft computing techniques for fault prediction in agile software development may result better performance in concern of precision. The SFP model affects in various areas of software development which in turn offer benefits in concern of reliability, quality, time and cost of completion.

4. CONCLUSION

The present research paper conclude that Neuro-Fuzzy (Hybrid) as a soft computing methodology provides more accurate result for prediction of faults, but it requires a real and huge data set to implement a model for agile software project from industry. There is a variety of metrics (Product, Process, and Agile based) that are used for feature selection process. The metric selection would be depending on the size or structure, and number of defects, effort on inspection activities, and experience of team members. This research paper also identified and reported the statistical significance of different accuracy measure parameters. This study also presents the vision of fault prediction field of software engineering by exploring the latest view of literature and a distinctive feature of fault prediction.

5. REFERENCES

[1] Parvinder S. Sandhu, Sunil Khullar, Satpreet Singh, Simranjit K. Bains, ManpreetKaur, Gurvinder Singh, 'A Study on Early Prediction of Fault Proneness in Software Modules using Genetic Algorithm', World Academy of Science, Engineering and Technology Vol. 72, pp. 648-653, 2010.

[2] Catal C, Diri B, 'A systematic review of software fault prediction studies', Expert System Application, Vol. 36, pp. 7346–7354, 2009. doi: 10.1016/j.eswa.2008.10.027.

[3] Hall T, Beecham S, Bowes D, et al, 'A Systematic Literature Review on Fault Prediction Performance in Software Engineering' IEEE Transaction of Software Engineering Vol. 38, pp. 1276–1304, 2012. doi: 10.1109/TSE.2011.103
[4] Begel, A., &Nagappan, N., 'Usage and Perceptions of Agile Software Development in an Industrial Context: An Exploratory Study' in Proceedings of the 1st International symposium on empirical software engineering and measurement IEEE, pp. 255-264, 2007,doi: 10.1109/ESEM.2007.12.
[5] Sharma, P., Sangal, A.L., 'Building and Testing a Fuzzy Linguistic Assessment Framework for Defect Prediction in ASD Environment Using Process-Based Software Metrics', Arabian Journal of Science and Engineering, Vol. 45, no.12 pp.10327–10351, 2020 https://doi.org/10.1007/s13369-020-04701-5.
[6] Y. Hassouneh, H. Turabieh, T. Thaher, I. Tumar, H. Chantar and J. Too, 'Boosted Whale Optimization Algorithm With Natural Selection Operators for Software Fault Prediction', IEEE Access, Vol. 9, pp. 14239-14258, 2021 doi: 10.1109/ACCESS.2021.3052149.
[7] KapilJuneja, 'A fuzzy-filtered neuro-fuzzy framework for software fault prediction for inter-version and inter-project evaluation', Applied Soft Computing Elsevier, Vol. 77, pp. 696-713, 2019.https://doi.org/10.1016/j.asoc.2019.02.008
[8] I. Tumar, Y. Hassouneh, H. Turabieh and T. Thaher, 'Enhanced Binary Moth Flame Optimization as a Feature Selection Algorithm to Predict Software Fault Prediction', IEEE Access, Vol. 8, pp. 8041-8055, 2020,doi: 10.1109/ACCESS.2020.2964321.
[9] Wang Geng, 'Cognitive Deep Neural Network prediction methods for software fault tendency module based on Bound Particles Swarm Optimization' Cognitive System Research Elsevier, Vol. 52, pp. 12-20, 2018.
[10] C. Jones, 'Measuring defect potentials and defect removal efficiency', Journal of Defense Software Engineering, Vol.21, no. 6, pp. 11-13, 2010.
[11] Chidamber S, Kemerer C, 'A metrics suite for object-oriented design'. IEEE Transaction of Software Engineering, Vol. 20 no. 6 pp.476– 493, 1994.
[12] K. V. J. Padmini, H. M. N. DilumBandara, and I. Perera, 'Use of software metrics in agile software development process', Moratuwa Engineering Research Conference (MERCon), pp. 312–317, 2015.
[13] M. Agarwal and P. R. Majumdar, 'Tracking Scrum projects Tools , Metrics and Myths About Agile', International Journal of Emerging Technology and Advanced Engineering, Vol. 2, no. 3, pp. 97–104, 2012.
[14] Arisholm, E., Briand, L.C., Johannessen, E.B., 'A systematic and comprehensive investigation of methods to build and evaluate fault prediction models', Journal of System and Software, Vol. 83, no. 1, pp. 2–17, 2010.
[15] Briand L., Wüst J., Ikonomovski S., and Lounis H., 'A Comprehensive Investigation of Quality Factors in Object-Oriented Designs: An Industrial Case Study,' Technical Report, International Software of Engineering Research Network-98-29, 1998.
[16] T. Sethi, 'Improved approach for software defect prediction using artificial neural networks' in the Proceedings of the 5th International Conference on in Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), pp. 480-485, 2016.
[17] Ravi Kumar, T., SrinivasaRao, T., 'Software Defects Prediction based on ANN and Fuzzy logic using Software Metrics', International Journal of Applied Engineering Research, Vol. 12, no.19, pp. 8509-8517, 2017.

A Systematic Review on Emotion Recognition System Using Physiological Measures

Ms. R. Selvi[1], Dr. C. Vijayakumaran[2]

[1]*Research Scholar, Department of Computer Science and Engineering, SRM Institute of Science and Technology, Chennai, sr9323@srmist.edu.in*

[2]*Associate Professor, Department of Computer Science and Engineering, SRM Institute of Science and Technology, Chennai, vijayakc@srmist.edu.in*

Abstract.

Emotion detection is the process of identifying and classifying different types of emotions. The systems can understand, recognize, identify and display emotions with the help of affective computation. The type of emotion expressed by a person is recognized via affect recognition. The Artificial Intelligence (AI) and Human-Computer Interaction (HCI) fields are realized effectively in emotional computing. The on-going research is in place to reveal an effective emotion recognition model through identifying the correlation of physiological signals and their highest contribution with other modalities like audio, video, and eyeball movement. The survey on this domain is conducted through following Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA), which acts as a standard guideline for conducting a qualitative review and meta-analyses. This study discloses 35 relevant documents over 513 articles on the web of sciences. Now a day, machine learning techniques are widely used to improve the classification result as well prediction. A brief assessment of the building emotion models with modern technologies like deep learning, its advantages, limitations, and future scope is discussed with a case study addressing the effective emotion model of interest..

Keywords. Affective computing, human computer interaction, emotion recognition, physiological signals, deep learning.

1. INTRODUCTION

Although computers do not have emotions, in human- computer relationship, they frequently communicate their emotions, to be as natural as possible when people connect with computers; the computers do not recognize them. Human emotions are organized using two methods based on conscious response and unconscious response in conscious response identify the emotions using PANAS (Positive and Negative Affect Schedule) which is the self-questionnaire that can estimate both positive and negative affects. In unconscious response, their emotions are identified from physic measures like electrocardiogram (ECG), electroencephalogram (EEG), electromyogram (EMG) activity, skin temperature, blood pressure, respiration rate, and facial expressions. The non-physiological signals are not able to reflect the innermost mental states of humans [1].Emotion recognition has its application towards various sectors from education, marketing, healthcare, and ensuring human safety while engaged in driving and so on [2].

The remaining part of the paper follows the structure as outlined as below In section 2, an exhaustive survey is carried out to ascertain the necessary facts about emotion recognition, its input modalities, and technologies that help in evaluating accurate emotions. The survey is conducted in strict adherence to the guidelines of PRISMA [3-5]. Section 3 discusses the knowledge inferred from the related literal work of resultant documents out of the qualitative analysis driven through PRISMA. Section 4 concludes the work and the scope for enhancement.

2. MATERIALS AND METHODS

The entire process flow of the systematic survey is captured in this section. This section addresses the research questions that initiate the study, the sources of documents, search keywords, and the final list of relevant documents using inclusion criteria in detail.

2.1. Research Questions

The key research goal of the study is to reveal the facts on emotion recognition, machine learning practices in predicting exact emotions and to analyse the real-time challenges associated with emotion models. The research questions are formulated in addressing the issues in the emotion model which in turn presents the gap to be addressed by the researcher in the future and listed as below.

Q1: What are the types of emotion models?

Q2: What are the input modalities of the emotion model? Q3: Are there any unimodal promises of better efficiency?

Q4: How machine learning can be integrated with emotion recognition?

Q5: What is the level of exploitation of machine learning algorithms and their efficiency?

2.2 Search Strategy

In order to select the documents of study, a search is devoted to three major electronic data sources (EDS) like IEEE Xplore, WoS, and Scopus. The articles which are published from 2010 through 2021 are accounted for selection. The keywords are refined according to the research questions. The major keywords of choices are Emotion Recognition, Physiological Signals, Facial Expressions, Multimodal, Machine Learning, and Deep Learning. These keywords are exercised with Boolean operators like “AND” and “OR” to formulate a query that in turn retrieves the document from the sources of interest. Furthermore, the scope of new sources is investigated, and to portray the market trend in emotion recognition, trusted website content is added with.

2.3 Criteria for Selection

The document selection is executed using the exclusion criteria listed below to extract the more relevant articles for the inclusion of the study.

EC1: Short summary research papers

EC2: Articles beyond the reputed publishers

EC3: No wider contribution on emotion recognition

The entire process of article selection process for the study aligned with updated PRISMA guidelines is depicted in Figure 2.1. Further, the quality assessment of extracted documents is scrutinized by manual investigation through analysing their content

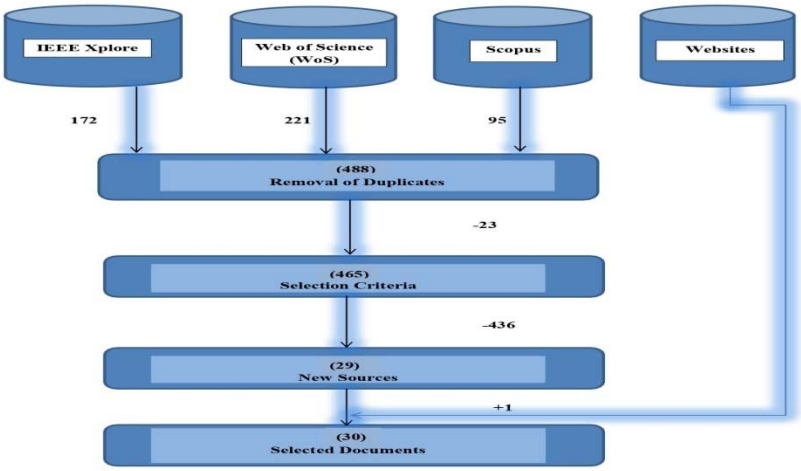


Figure 2.1. Selection Process of Systematic Study on Emotion Recognition

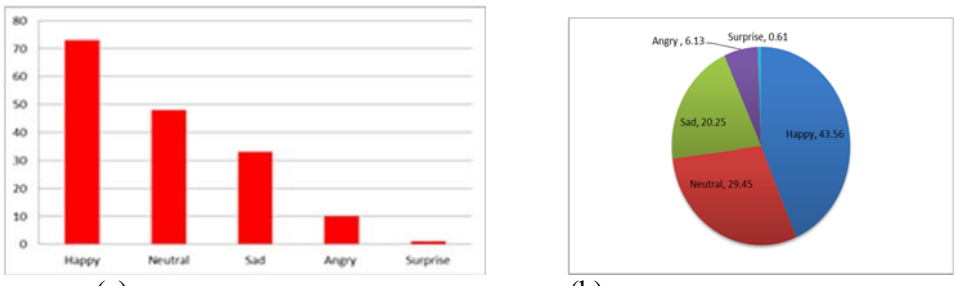
3. RESULTS AND DISCUSSIONS

The summary of the study based on the research questions is elaborated here. Further, a novel design for emotion recognition incorporating a deep learning algorithm is discussed.

3.1 Emotion Models

Human emotions are categorized into two types in emotional space namely discrete emotion model and dimension emotional model.

Mathura Prakash et al. [6] contributed an emotion recognition system that detects different emotions of the face and obtains the average emotional state during a particular event and assessed the same via the audience feedback system. The resultant of the assessment process is illustrated in Figure 3.2. The bar graph shows the number of times the occurrence of a particular emotion, and the pie chart describes the percentage of having a gathered emotion.



(a) (b)

Figure 3.1. Evolution Result of Emotion Recognition System Presenting

a) Number of Times an Emotion Occurred and b) Percentage of Gathered Emotion [6]

3.2 Deep Learning-based Emotion Recognition Models

Deep learning (DL) and machine learning are becoming the next evolution in artificial intelligence. The trained model learns new things like the human brain through a neural network in deep learning. DL is employed in most province emotion recognition studies due to its efficacy in deep feature extraction.

Chao and Don [7] developed a convolutional neural network (CNN) to recognize human emotions using the ResNet algorithm. They analysed emotions from multichannel EEG signals in the time domain. Based on the position of electrode sensors the features are extracted and treated as a 3- D feature matrix. Whereas a pre-processing of EEG signals is carried out by principal component analysis before applying the classification with CNN is captured in [8].Similarly, an emotion recognition system based on EEG signals is experimented with tuned CNN, which adjusted with the tuning parameter in convolution layer and it produced 73.4% of accuracy in successful classification of emotions [9]. In most of the previous studies,. They classify the emotions obtained from multimodal physiological cues with the help of DL and ML algorithms to achieve better results. Table 3.1 summarizes the key contributions of the existing endeavour.

Table 3.1. Summary of significant emotion recognition models in practice

Reference	Modalities	Emotions	Techniques Used	Dataset/Accuracy
X Zhang et al. [11]	EEG, EMG,GSR, RES, MEG, EOG	Arousal, Valence	SVM-FLF (Support VectorMachine-Future Level Fusion), Fully Convolutional Network (FCN),SVM-DLF(Decision Level fusion)	DEAP-63.8% DECAF-64.2%
K. Zhang et al. [12]	Text, audio,visual	Happy, sad, neutral, angry, excited, frustrated, average	Gated Recurrence Unit (GRU)	IEMOCAP –61.8 % CMU-MOS-81.2%
M. D. Hssayeni et al. [10]	Respiration rate, ECG, Skin temperature , EMG and acceleration	Stress, amusement	Gradient tree boosting	WESAD-79%
H. Chao et al. [7]	EEG	Arousal,Valence	CNN-ResNet, Sliding window	DEAP, Binary -85.5% Four-Class-75.7%
B. Nakisa et al. [13]	EEG, BVP signals	Arousal,Valence	ConvNet, Long Short-term Memory	MAHNOB-71.6%

Reference	Modalities	Emotions	Techniques Used	Dataset/Accuracy
G. Du et al.[14]	ECG, EEG	excitement, anger,sadness, calmness	BLSTM, CNN	Kinect2.0-87.3%

In most of the previous studies, the convolutional neural network is widely used for classifying emotions in Arousal- Valence space Hence the researchers employed CNN in producing high-end emotion detection mode with promising accuracy.. Also, the performance analysis is captured in Figure 3.1. Further, they can achieve an increased rate of accuracy by combining more than one deep learning algorithms.

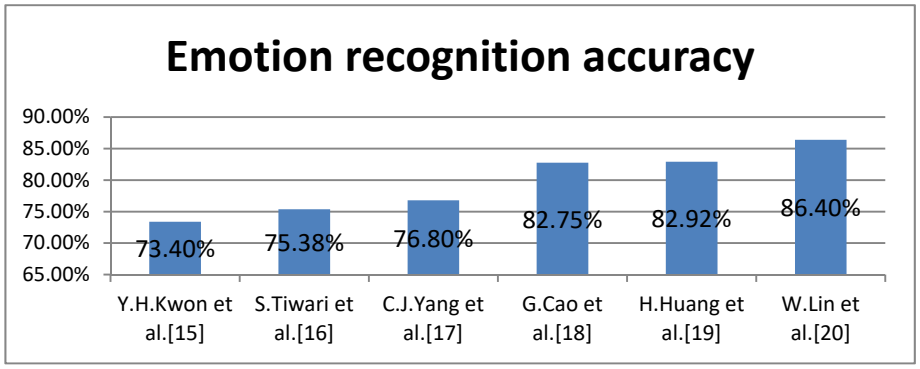


Fig 3.1. The Efficiency of Neural Network Assisted Emotion Detection Models

4. CONCLUSION

Building deep architecture-based real-time multimodal emotion identification system is an emerging topic of research. Affect recognition is becoming increasingly popular among researchers because of its growing applicability in the education and healthcare industry. Physiological signals exhibit more advantages in emotion recognition, as they are uncontrollable, unaffected by culture or education, and strongly linked to people's emotional states. A review of the study broadly addresses the emotion models, interesting modalities to be considered as input in the process of emotion evaluation, and the scope for the multimodal emotion detection system. The emerging scope of deep learning practices in this domain and its exploitation level and attained efficiency is well investigated.

5. REFERENCES

[1] M. Ali, A.H. Mosa, F. Al Machot, K. Kyamakya, 'Emotion Recognition Involving Physiological and Speech Signals : A Comprehensive Review', In: K. Kyamakya, W. Mathis, R. Stoop, J. Chedjou, Z. Li (eds) Recent Advances in Nonlinear Dynamics and Synchronization. Studies in Systems, Decision and Control, Springer, Cham, 109, 287-302, 2018.

[2] M. Ali, F. Al Machot, A. H. Mosa, and K. Kyamakya, 'CNN Based Subject-Independent Driver Emotion Recognition System Involving Physiological Signals for ADAS', In: Advanced Microsystems for Automotive Applications, Springer InternationalPublishing.pp.125–138, 2016.

[3] A.A. Selcuk, 'A Guide for Systematic Reviews: PRISMA', Turkish Archives of Otorhinolaryngology.57(1),pp.8,2019, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6461330/

[4] D. Moher, A. Liberati, J. Tetzlaff, D.G. Altman, 'The PRISMA Group (2009), Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement', PLoS Med, 6(6): e1000097, 2009

[5] M.J. Page, J.E. McKenzie, P.M. Bossuyt, I. Boutron, T.C. Hoffmann, C.D. Mulrow, L. Shamseer, J.M. Tetzlaff, E.A. Akl, S.E. Brennan, R. Chou, J. Glanville, J.M. Grimshaw, A. Hrobjartsson, M.M. Lalu, T. Li, E.W. Loder, E. Mayo-Wilson, S. McDonald, L.A. McGuinness, L.A. Stewart, J. Thomas, A.C. Tricco, V.A. Welch, P. Whiting, D. Moher, 'The PRISMA 2020 statement: an updated guideline for reporting systematic reviews', BMJ, 372, pp.1-9, 2021.

[6] Madhura Prakash, Aman Kumar, Anmol Saxena, and Somil Agrawal, Twinkle Singh, 'Audience Feedback Analysis using Emotion Recognition', Int. J. Eng. Res., 9(5),pp.1090–1094, 2020.

[7] H. Chao, L. Don, Emotion Recognition Using Three-Dimensional Feature and Convolutional Neural Network from Multichannel EEG Signals, IEEE Sensors Journal. 21(2) (2021) 2024-2034.

[8] G. Cao, Y. Ma, X. Meng, Y. Gao, M. Meng, Emotion recognition based on CNN, *Chinese Control Conf. CCC*, 2019(61372023) (2019) . 8627–8630.

[9] Y.H. Kwon, S.B. Shin, S.D. Kim, Electroencephalography based fusion two dimensional (2D)-convolution neural networks (CNN) model for emotion recognition system, Sensors. 18(5):1383 (2018) 1-13.

[10] M.D. Hssayeni, B. Ghoraani, Multi-Modal Physiological Data Fusion for Affect Estimation Using Deep Learning, IEEE Access. 9 (2021) 21642-21652.

[11] X. Zhang, J. Liu, J. Shen, S. Li, K. Hou, B.H.J. Gao, T. Zhang, B. Hu, Emotion Recognition From Multimodal Physiological Signals Using a Regularized Deep Fusion of Kernel Machine, IEEE Transactions on Cybernetics. 51(9)(2021)4386 – 4399.

[12] K. Zhang, Y. Li, J. Wang, Z. Wang, X. Li, Feature Fusion for Multimodal Emotion Recognition Based on DeepCanonical Correlation Analysis, IEEE Signal Processing Letters. 28 (2021) 1898 -1902.

[13] B. Nakisa, M.N. Rastgoo, A. Rakotonirainy, F. Maire, V. Chandran, Automatic Emotion Recognition Using Temporal Multimodal Deep Learning, IEEE Access. 8 (2020) 225463 -225474.

[14] G. Du, S. Long, H. Yuan, Non-contact Emotion Recognition Combining Heart Rate and Facial Expression for Interactive Gaming Environment, IEEE Access.8(2020),11896 – 11906.

[15] Y.H. Kwon, S.B. Shin, S.D. Kim, Electroencephalography based fusion two dimensional (2D)-convolution neural networks (CNN) model for emotion recognition system, Sensors. 18(5):1383 (2018) 1-13.

[16] S. Tiwari, S. Agarwal, K. Adiyarta, M. Syafrullah, Classification of physiological signals for emotion recognition using IoT, Int. Conf. Electr. Eng. Comput. Sci. Informatics. (2019) 106–111, 2019.

[17] C.J. Yang, N. Fahier, C.Y. He, W.C. Li, W.C. Fang, An AI-edge platform with multimodal wearable physiological signals monitoring sensors for affective computing applications, IEEE International Symposium on Circuits and Systems (ISCAS). (2020) 3–7.

[18] G. Cao, Y. Ma, X. Meng, Y. Gao, M. Meng, Emotion recognition based on CNN, *Chinese Control Conf. CCC*, 2019(61372023) (2019) . 8627–8630.

[19] H. Huang, Z. Hu, W. Wang, and M. Wu, "Multimodal Emotion Recognition Based on Ensemble Convolutional Neural Network," *IEEE Access*,vol.8, no. 2, pp. 3265–3271, 2020.

[20] W. Lin, C. Li, S. Sun, Deep convolutional neural network for emotion recognition using EEG and peripheral physiological signal, In: Zhao Y., Kong X., Taubman D. (eds) Image and Graphics. ICIG 2017. Lecture Notes in Computer Science, Springer, Cham. 10667 (2017) 385-394.

Service Oriented Implementation of Information and Communication Technologies in Education

Dr. Sunil Ghildiyal[1], Dr. Ashish Gupta[2,] Dr.Kapil Joshi[1], Dr. Minakshi Memoria[1], Thanh-lan Thi Nguyen3, Mr. Sagar Bhatt[4]

Uttaranchal University, Dehradun, Uttarakhand INDIA[1]
Devbhoomi Uttarakhand University, Dehradun, Uttarakhand INDIA[2]
Assistant Professor, Department of Liberal Arts, Wonkwang University, Korea[3]
Assistant Professor, GNA University Phagwara, Punjab,India[4]

Abstract:

Internet has enabled the people to involve the use of information and communication technologies (ICT) in every routine. Easy graphical interface of such technologies has made it so popular that anyone can utilize technologies and associated applications by just visualization. It has attracted educationists to utilize the ICT in the routine education delivery and evaluation system. It has been observed that the teaching and learning process has been very effective while ICT and its associated tools are included. Many government policies throughout the globe have been focusing on providing the ICT based education but many analysis reported that educational system were lacking in the use of ICT for its routines. Due to COVID19, regular educational routines are not possible to avoid contacts. ICT has been continuing the teaching learning process in the education system. But ICT technologies have their own issues and face many challenges while being implemented as a service. in developed and developing countries. This paper aims to find out the role of ICT to enhance the education delivery specially in the pandemic situations.

Keywords:
Information Technology Education Government Service Society

I. INTRODUCTION

Education system demand a very rapid sharing of knowledge, research outcomes and opportunities through web resources. Involvement of technologies requited for it are computing, storage and communication, leading to ICT[1] [17]. In the era of computing we cannot even think for education delivery without using IT aided technologies. A very simple form of IT utilization in the education system has been presentations for last about three decades and so on. Traditionally, over head projector with transparencies as storage media and multimedia projectors with electronic storage media like CD,DVD and pen drive have been used for this purpose. But when

we talk about the global scenario of education system, web resources and web communication is involved with use of email, social networking. ICT driven education is based on sharing and transmission of knowledge, research facts transferred through electronic mode without physical contact. Inclusion of ICT methodologies and services are making major differences in the pedagogical approaches. ICT enhanced learning environment facilitates collaborative, creative and evaluating learning as an advantage over the old methodologies[22]. ICT based education is successor of computer-aided teaching and learning which is based on the computer-aided instruction over communication channels. Use of such instructions are flexible, broad viable and more opportunistic. Involvement of Information and Communications Technology (ICT) has motivated the teaching, learning and research tremendously for last about one decade. Its ease of use, access has made it a major component of educational system with the use of handouts, documents or presentations.ICT supportive tool have been very popular as a pedagogical means in education. It drives the cognitive, pedagogical and affective benefits of using ICT for teaching and learning process[2]. Higher education system in all countries undergoing a big transformation phase for access, equity and quality of educational are research facts. It has been possible the rapid evolution of information and communication technologies (ICT) around the world. ICT has influenced almost all the components of educational process like access, equity, management, efficiency, pedagogy and quality[3]. ICT has transformed the learning process of education system to the next level where teaching learning tools are based on technology and helping students, teachers and also updating to the parents or guardians. ICT has transformed the method of dissemination of knowledge. ICT has also helped to modify the teaching learning process model. ICT has replaced the traditional methods of research, education and provided the more feasible and reliable education delivery system concept instead of the traditional methods. Even nowadays people also prefer the ICT tools and methodologies to access the knowledge as it can also be recorded latest advances in different areas of education and research and allied domain. Education is backbone of a country's social and economic growth to increase the productive capacities and upliftment of individual and society individually. Education helps to develop the sense of a good social living and many other intangible benefits. ICT has a variety of products available related to the education like online platforms, teleconferences, online sessions, web based sessions and many storage devices.

II. ICT IN EDUCATION

Technological enhancements have affected the industrial sector and different productive/service centric sectors including education[23]. ICT is viewed as a "major tool for building knowledge societies" (UNESCO 2003) and particularly, as a mechanism for education that could provide new method for rethinking and redesigning of the educational contents and processes, thus leading to quality education for all[4] . Appropriate use of ICT has already been proven in Europe school education and has been considered a key factor in improving quality of educational level. It motivated them to introduce the promotion ICT in teaching-learning processes through their e-Learning plan. UNESCO and UNICEF with the association of government of India, have started a pilot project in India for internet access in every school, college. But due to constraints like no electricity connections,

poor geographical conditions, non availability of public infrastructure, it got limited in urban areas only. ICT has many applications in area of education. ICT has been proved as booster to enhance the quality as well as delivery scope of education to any of the nation. ICT enables the remarkable change in the nature and quality of education by facilitating students or audience with many methodologies associated to ICT. ICT technologies attracted its audience such as students, teachers, educationists and the community even who were never been habitual of such technologies. ICT improves the teaching-learning process through the more interactive education methodologies and tools to attract the learner's motivation. It is a way of level placing of each and everyone to get access of same education platform, bringing education to the doorstep of everyone, living remote or rural geographic areas[5].

ICT strengthens the student learning and better teaching skills. It has been proved that more involvement of ICT in education, curriculum designing and integration have resulted much better goals achievement by student and teacher both. It has been remarkably recorded in subjects like science, physics, and mathematics [6]. However use of technology between teacher and students requires both to be digitally literate. Use of ICT in the education system can lead better content addressing, mutual information exchange between both and to create better teaching environment [7]. However budget constraints are always there to be considered for it. Digital literacy enables higher order thinking skills, positive creations, smooth exchange of the dialogues between teacher and student. Use of ICT in education also leads to meet the with ongoing technological change in society and surroundings[8][18]. ICT in education leads inclusion of interesting, colourful and much graphical images in the delivery content for better understanding, pictorial and animated presentation of complex instruction by the use of appropriate technology. It helps to create interactive classes with much enjoyable contents to attract more attention of the students to improve attendance. ICT helps a lot to the distance learners to use online educational materials and ease access of the online resources, leading to the resource assist learning. But implementation of ICT meets with many issues like rapid changes in the society, economy, and parallel technologies. Introduction of ICT in the system cannot become fruitful until the users are also literate to access and avail its benefits. General crowd of the developing countries are still not much aware and capable of using technology or its advanced tools. It has been recorded as serious issue while implementing ICT [9].

III. CHALLENGES FOR ICT IMPLEMENTATION

a. Reengineering of Education System:

Use of ICT in education should follow use in society, not lead or impose it. Programmes that use cutting-edge technologies have been observed failed in long term. The teachers, students and parents or guardians have been habitual of traditional education systems for many decades. Formulation of new technique or approach for education contents delivery and its evaluation need to create a zeal of adapting new trends. In general, societies

lack on initiations for observing, understanding and adapting new concepts. They limit the use of the internet based technologies for entertainment or social networking only. The main reason for it is the lack of trust on technology. Hence there is a demand of reengineering of education system with the involvement of government, administration and society itself to motivating the trust and inclusion the technology in daily routines and education system and avoid the contacts in case of disasters like COVID19.

b. Infrastructure Building:

Inclusion of ICT technology need infrastructure at hardware, software, courseware levels. In most of the countries existing old infrastructure cannot sustain the load of electrical wiring, heating/cooling and security systems required for implementation of ICT technology. In many developing countries, obstacles as no constant and steady supply of electricity, communication channel are far away. These countries have very poor wireless telecom infrastructure. Many schools, universities in the developing countries lack computers, electricity and telecom connectivity and their sustainability[10][19]. Geographical inequity and imbalance make it tough to provide and maintain ICT infrastructure. However developed countries take advantage in this area.

c. Integration of Technology and pedagogical innovation:

New innovations in the current pedagogy are required to get integrated with ICT technology driven education. Policymakers at government, institution, and administration level are required to make appropriate regional, national and global policies to motivate the integration of ICT with education pedagogy. The effectiveness of ICT in education depends on how ICT is included, used and for what purpose. ICT does not for everyone, everywhere and in the same manner. Enhancing gradual but constant access , enhancing the quality of ICT based educational contents and placing those contents at same level at diversified geographical, economical areas of the continents. Using proper audio, video and animation technologies, interesting contents can be developed to attract the attendance. Contents and evaluating assessments based on ICT are very much helpful for distance learners. A research has already claimed that there was “no significant difference” between the grade scores of learners taking education through ICT-based distance courses and those who were getting it in a class room or auditorium in face-to-face instruction method.

d. Capacity Building:

Administrators of education organisation, proper fiscal management, motivating policies for teachers also play important role for implementation of ICT in education system. Teachers of the educational institutes do not take interest to learn new methodologies to deliver the contents and assessment subsequently. It is due to the lack of interest to teach ICT, lack of ICT literacy among teachers and students, lack of knowledge of development of pedagogy and its modification for ICT, limited access to ICT facilities. Technical trainings are also needed for teachers, staff and management team of the organisation.

e. Cost, Equity and Sustainability:

The importance of information and communication technology (ICT) in education in the current era cannot be overemphasized. Many literature reviews has shown that ICT tools and methodologies have become indispensable in today's information age, drastically impacting on the global living. ICT has efficiently supported entire globe during latest faced pandemic also when there is no any provision of physical contact anyhow [11]. ICT infrastructure needs lot of investments in areas like purchase of devices, equipments , trainings to the technicians, teachers and maintenance of the equipments. In developing countries money investments matter a lot as those face its scarcity. Cost effective solution in the area of ICT are to be deployed, especially in the countries where economy is too low or government policies don't prefer such investments on high priority. Equity is related to the implementation of ICT enabled environment globally among the developing, less developed and developed countries. All these three have their own meaning and ways of money investments for education system and then implementation of ICT in it. ICT solutions have to be sustainable for meeting the economical growth, environmental protection, society, eco-system and democracy of any country or region. The use of technology in any sector should not affect any type of sustainability.

IV. CONSTRAINTS FOR SOLUTION

While designing or thinking for an effective solution, many factors are to considered at teacher, student and infrastructure level. For example, perceptions about ICT and its associated curriculum by student, teachers parents and also by society. Awareness and literacy about ICT in all. Pedagogical history and approach of students and teachers. Availability and sustainability of ICT infrastructure for students and teachers[12][19][20].

There is need of complete guidelines on successful inclusion of the ICT based curriculum in the education system or curriculum. Educating and creating positivity for use of ICT among the teachers and students is also secondary task for successful implementation of ICT. It has been observed that many education system and their counterparts avoid ICT or resist it due to lack of full information about its productivity. It has been observed for last about two decades that multimedia based animated teaching contents, animated case studies of the subjects, bring to more learning. However lack of ICT infrastructure and poor quality of available infrastructure hinder it. When we talk about the home or family support for ICT, it totally differs from the college, universities. For example children have more interest in animated contents, girls have another approach to use ICT contents. Many parents or guardian find animated contents useless as teaching aids. Main concern about avoiding the ICT is its out of reach either technically or economically specially in developing countries[13]. Governments also have a major role to promote the ICT usage by providing better infrastructure, education or awareness

programmes for teachers, students and even entire society for its usage and providing the subsidies to the related authorities to promote ICT infrastructure.

Many myth and worries with the implementation of ICT are also there. These are to be addressed while implementing ICT solution in the education system. For example will exposure and use of ICTs affect the future employment in education sector? Or What will be the impact of computer aided instruction in education Or How gender will impact on ICTs in education Or Is ICT suffice for use to present, comment on and discuss student work or are all educational setup and their curriculum are suitable for ICT integration than others? Many findings suggest that ICT companies may result enriched in education sector also not only for its deployment in boosting circular economy in the service-oriented technology sectors [24]. Further ICT era may be involved with the help of IOT in village 4.0 and machine learning will also help for better communication [25-30].

CONCLUSION

ICT has been proved as necessity and an opportunity in education system specially in pandemics when there is no option except dependency on digital mode of connectivity. While using any technology, issues and challenges are there to limit the scope of it. In developing countries implementation of ICT technologies and support for it is a big obstacle14]. Introduction of ICT methodologies motivated the teachers as well as learner to explore more components of it for its optimum utilization. Though the ICT technologies have faced many social, ethical, technological, economical challenges, cost and power solutions but it has attracted people to explore and enjoy its factors which were hidden for a long time.

Many developing countries face lack of resources for implementation of ICT technologies specially in educational sector educational is one of the hindrance in its implementation[15]. Governments, Management bodies avoid to such investments. Secondly, less exposure to the computer, ICT and technologies of the students and teachers is also responsible for it. The education system needs to be alert for ICT enabled technologies in its system for regular routines of education as enhancement and as a preventive action in the case of pandemic. Since, implementation of ICT major dependent on the mentors and academic management, government, allied corporate sectors, can help by providing infrastructure, in-service service centric training to the society. Private sectors can be collaborated for this novel cause.

References:

1. “ICT and higher educational system in Nigeria”, Adeyemi I. Idowu and Mary Esere, Vol. 8(21), pp. 2021-2025, 10 November, 2013
2. “ICT in education: Benefits, Challenges and New directions”, Sotiria FoutsitziGeorge Caridakis IEEE *Xplore* November 2019
3. “The Role of ICT (Information & Communication Technology) in Higher Education” Naila Zia Multidisciplinary Higher Education, Research, Dynamics & Concepts February 2018
4. “The role of information and communication technologies in improving teaching and learning processes in primary and secondary schools” , Albert Sangrà &Mercedes González-Sanmamed, December, 2016 https://doi.org/10.1080/09687769.2010.529108
5. “Issues And Challenges in Bringing ICT Enabled Education To Rural India”, Dr. Shradha H. Budhedeo, IJSAE, Vol.4, Issue 01January-2016
6. “Information Technologies and Management”, C Koryuhina1, T Shamshina, April 26-27, 2018, ISMA University, Riga, Latvia
7.”ELMO 2017 What'sICT in education? Using ICTs and Blended Learning in Transforming TVET in ed”, C. Latchem, UNESCO and COMMONWEALTH of Learning, 2017
8. “Use of Tablet Technology in the Classroom Strathfield”, Godwin K, NSW Curriculum and Learning Innovation Centre 2012

9. “Futures for ICT and Higher Education: Changes Due to the Use of Open Content”, UNESCO 2016
10. “Challenges Faced by Schools when Introducing ICT in Developing Countries”, NomsaMndzebele, IJHSSI Vol 2, Issue 9, September. 2013
11. “Challenges facing information and communication technology implementation at the primary schools”,
Mark Valentine and Emmanuel Arthur, Nyarko
Education and Research reviews Vol. 14(13), pp. 484-492, July, 2019
12. “Benefits and Challenges of Information and Communication Technologies (ICT) Integration in Québec English Schools”, Jihan Rabah, TOJET, April 2015
13. “Challenged Facing in Implementation of ICT Curriculum in Primary School”, A Kassimu, Inaugural International Conference on Open and Flexible Education, Hong Kong-China, 16th-17th January, 2014
14.” Challenges and solutions when using technologies in the classroom”, Amy M. Johnson, Matthew E. Jacovina Arizona State University.
15. “The use of internet by secondary school teacher: A Case of twelve secondary schools in Dar es Salaam”, Mapunda, M. (2004).
16. “The Role and Impact of ICT in Improving the Quality of Education: An Overview”, Koushik Das (IJISSH Volume: 4 Issue: 6 | June 2019
17. “New Challenges for ICT in Education Policies in Developing Countries: The Need to Account for the Widespread Use of ICT for Teaching and Learning Outside the School”, J. Enrique Hinostroza et.al. , 2018 – Springer Educational Communications and Technology: Issues and Innovations book series (ECTII)
18. “Impact of ICT on Education: Challenges and Perspectives”, Hernandez, Ronald M.,
Journal of Educational Psychology - Propositos y Representaciones, v5 n1 p337-347 Jan-Jun 2017
19. “Teachers’ Views on the Use of Information and Communication Technologies (ICT) in Education Environments” Zhiyasheva Shoraevnai et.al., IJET Volume 16, Number 3, Feb 12, 2021 ISSN 1863-0383 Publisher: International Journal of Emerging Technology in Learning, Kassel, Germany
20. “Information and Communication Technology-Based Education Planning and Attitude of College Students”, Yamunah Vaicondam , IJIM, Vol. 15 No. 04 (2021) https://doi.org/10.3991/ijim.v15i04.20365
21. Technological Pedagogical Content Knowledge (TPACK) with Information and Communication Technology (ICT) Integration: A Literature Review”, Syaeful Malik, Dedi Rohendi, Isma Widiaty, February 2019, Proceedings of the 5th UPI International Conference on Technical and Vocational Education and Training (ICTVET 2018)
22. Information and Communications Technology for Teacher Training in India , Sharma, Anchal ICT India Working Paper No. 64, 2022, Columbia University, Earth Institute, Center for Sustainable Development (CSD), New York, NY
23. The core components of education 4.0 in higher education: Three case studies in engineering education
Jhonattan Miranda et.el. Elsevier Volume 93, July 2021, 107278
24. Developing and implementing circular economy business models in service-oriented technology companies, Graeme Heyes et.el. , Elsevier, Journal of Cleaner Production, Volume 177, 10 March 2018, Pages 621-632.
25. Malik, Praveen, Singh, R., Gehlot, A., Akram, S. V., & Das, P. K. (2022). Village 4.0: Digitalization of Village with Smart Internet of Things Technologies. Computers & Industrial Engineering, 107938.
26. Bagwari, S., Gehlot, A., Singh, R., Priyadarshi, N., & Khan, B. (2022). Low-Cost Sensor-Based and LoRaWAN Opportunities for Landslide Monitoring Systems on IoT Platform: A Review. IEEE Access, 10, 7107–7127.
https://doi.org/10.1109/ACCESS.2021.3137841
27. Malik, P, Gehlot, A., Singh, R., Gupta, L. R., & Thakur, A. K. (2022). A Review on ANN Based Model for Solar Radiation and Wind Speed Prediction with Real-Time Data. Archives of Computational Methods in Engineering. https://doi.org/10.1007/s11831-021-09687-3
28. Singh, R., Gehlot, A., Rashid, M., Saxena, R., Akram, S. V, Alshamrani, S. S., & Alghamdi, A. S. (2021). Cloud server and internet of things assisted system for stress monitoring. Electronics (Switzerland), 10(24). https://doi.org/10.3390/electronics10243133
29. Karunanidy, D., Ramalingam, S., Dumka, A., Singh, R., Rashid, M., Gehlot, A., ... & AlGhamdi, A. S. (2022). JMA: Nature-Inspired Java Macaque Algorithm for Optimization Problem. Mathematics, 10(5), 688.
30. Sharma, V. K., Singh, R., Gehlot, A., Buddhi, D., Braccio, S., Priyadarshi, N., & Khan, B. (2022). Imperative Role of Photovoltaic and Concentrating Solar Power Technologies towards Renewable Energy Generation. International Journal of Photoenergy, 2022.

Dr. Sunil Ghildiyal is working as Associate Professor in Department of Computer Science and Engineering, UIT, Uttaranchal University, Dehradun. He has obtained Ph.D. in area of Wireless Sensor Networks. He has more than 15 years experience in area of Teaching and Research. His area of interest are Computer Networks, Java, ICT and Service Oriented Computing.

Retrofitting and Restrengthening of RC Beam Using FRP Technology

Kandula Kiran[1]*, Nimmagadda Venkata Rao[2]

[1]*Civil engineering department, Institute of Aeronautical Engineering, Hyderabad, India*

[2]*Civil engineering department, Institute of Aeronautical Engineering, Hyderabad, India*

Email: kirankandula1998@gmail.com; venkatnimmagadda@iare.ac.in

**Corresponding Author*

Abstract.

This External interaction of creative FRP (fiber reinforced polymer) polymer composite material has developed in popularity in recent years because it offers a supplementary profitable and theoretically excellent compared to conventional procedures in several circumstances due to its durability, lightweight, resistance to corrosion, fatigue resistance, easy and quick construction, and least modification in structural configuration. Because several in-situ RC (Reinforced concrete) beams were consecutive in structure construction, there was not enough scientific interest in beam strengthening using FRP mechanism. The foremost objective of this investigation involved is to propose procedures for restoration of RC beam utilizing FRP material. The characteristics of FRP and their impact on beam strengthening are illustrated. The research concentrate on self-compacted RC beams that have been strengthened by wrapping CFRP (Carbon fiber reinforced polymer) & GFRP (Glass fiber reinforced polymer) layers in epoxy adhesive.

Keywords. Self-compacting concrete, Steel bars, FRP sheets, Epoxy resin and Hardener.

1. INTRODUCTION

From the past two decades construction of structures has increased significantly. Concrete has a significant part in building engineering due to its advantages in several areas like as economy, strength, and durability. SCC (Self compacting concrete) is used to improve the structural performance and sufficient filling of narrow sections and massively reinforced structural components reported by Okamura et al.,[1]. An elementary proposal for the SCC was given by Nan Su et al., [2]. In this mix design Portland cement was utilized. The aggregate percentage required for the concrete is determined first, and then binder numbers required is determined for which concrete can conquer decent flow capability, self-compressing capability and remaining SCC assets. More knowledge of a good performance SCC is required for such a performance concrete to reach wider acceptability for casting tough and crowded structural components with significant reinforcements, particularly in seismic zones.

1.1. FRP Technology for Strengthening and Rehabilitation of Structures

Since major renovation or rebuilding of the structure is not cost efficient, reinforcing or remodelling is an effective technique to improve it. The use of FRP sheets for strengthening of RC beams will reduce the deflections and increase the load carrying capacity described by Grace [3]. Renovating concrete constructions with FRP sheets provides a supplementary cost-effective and properly excellent advantages over traditional methods in various circumstances because it provides heartiness, light mass, rusting defiance, higher exhaustion resilience, to wrap up quick construction, and slight alteration in physical geometry. Strengthening of RC beam using CFRP sheets will decrease the ductility when compared to conventional RC beam member designated by E1-Refaie [4]. N. Pannirselvam [5] stated that the load bearing of FRP reinforced beams contained, GFRP sheets were enhanced compared to controlled beam. The increasing in load sustaining ability of a structural associate dependent upon the number of FRP layers wrapped on RC beam. FRP systems can also be employed in regions where traditional methods would be impossible due to limited access. However, due to a absence of suitable understanding on the structural performance of reinforced concrete constructions, the application of FRP materials for renovation existing concrete structures falls short of the expectations.

2. MATERIALS

- Portland cement: Across the experiment, OPC (ordinary Portland cement) of 53 Grade approving IS:12269-2013[6] was chosen from a single collection. It was fresh and free of lumps. Cement is carefully stored to avoid deterioration of its characteristics due to humidity interaction.
- Fly Ash: By-product which is formed by combustion of coal that mostly contains of silicon dioxide (SiO_2) and calcium oxide (CaO). When it was mixed with concrete, fly ash undergoes a primary pozzolanic reaction with the moisturized cement paste, resulting in a denser microstructure over time. NTPC (National Thermal Power Corporation) fly ash was obtained from Ramagundam during the experimentation.
- Coarse Aggregate: Samples approving IS:383-2016[7] retained at 10mm and passing through 12.5mm were used. The specific gravity is 2.78. The aggregates were cleaned to eliminate dirt and dust before being dried to the surface.
- Fine aggregate: Sand authorising IS: 383-2016[7] with 2.70 fineness modulus and the specific gravity is 2.65.
- Longitudinal steel (Top &Bottom bars) and Lateral steel(stirrups):Steel bars with Fe415 yield strength were used. The beam's reinforcing design is determined in accordance with IS456:2000[12]. Figure 2.1 represents the beam reinforcement details.

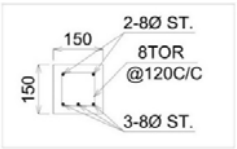


Figure 2.1.Beam Reinforcement Details

- CFRP sheets: CFRP refers for composite materials. The composite in this experiment is made up of two segments: a matrix and reinforcement. Generally, the matrix is a polymer resin, for example epoxy, which provides bond between two components. The CFRP sheet is bidirectional with tensile strength of 3500N/mm².Figure 2.2(a) shows the image of CFRP sheets used in the present work. Two layers of CFRP was wrapped to beams.
- GFRP sheets:GFRP composites were the frequently used materials in complex product manufacturing. Polyester, thermostable, vinyl ester, phenolic, and epoxy resins were applied in the matrix. Afiber reinforced composite's mechanical performance is essentially determined by strength of fiber and magnitude, chemical strength, matrix stability, and interactive bond between the fiber and matrix to permit stress transmission. GFRP composites offered the equivalent physical and efficient assets as steel, were stiffer than aluminium with a specific gravity one-quarter to steel. The GFRP sheet is bidirectional with tensile strength of 3400N/mm². Figure 2.2(b)shows the image of GFRP sheets used in the present work. Two layers of GFRP was wrapped to the beams.

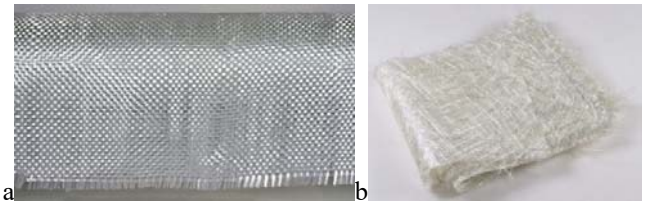

a b

Figure 2.2. (a)CFRP Sheets;(b) GFRP sheets

- Epoxy Resin and Hardener: Epoxy resin is a liquid chemical compound that is utilized for a number of purposes across a wide range of industries. The epoxy resin and hardener pair are the ideal working combination, giving excellent durability, tensile strength, and working compatibility.
- Super Plasticizer: The chemicals which reduce water content in manufacturing of high strength concrete referred as super plasticizers (SP). By using these chemicals in concrete mixing, we can reduce the water content up to 30%. To ensure high workability, Bento polymix PCE 3000 was utilized.

3. METHODOLOGY

3.1. Mix Design of Concrete

The design mix for M35 grade SCC is done as per IS 10262:2019 guidelines [8]. Mix ratio of the SCC is 1:2.71:2.15. The W/C ratio (water cement) is locked at 0.40 and a slump of 70 to 75cm and properties were represented in Table 1.

Table 1 SCC mix proportion(kg/m³)

Cement	*Fine aggregate*	*Coarse aggregate*	*W/C*
341.25	925.00	734.32	195
1	2.71	2.15	0.40

3.2. Mixing

I combined the materials according to the mixed design after measuring the materials quantity. To begin, mix coarse aggregates and fine aggregates for a few min to obtain a uniform mix, then add cement and fly ash to the mixture and mix for another few minutes to obtain the same mix throughout the material. To produce newly prepared concrete for M35 grade concrete, added water according to the calculations from the mix design.

3.3. Workability tests

Workability tests were conducted on freshly prepared SCC to establish the capacity under filling and passing according to IS:1199 part-6[9].

3.3.1Slump Flow Test

The filling performance of SCC is measured by using slump flow test. It evaluates two mechanisms: flow distribution and T50 flow time. The flow distribution denotes unrestricted, clear deformability, whereas T50 denotes deformation rate within certain flow distance.

3.3.2 V-Funnel Flow

Duration for the definite capacity of SCC passes through a thin introductory is called as V-funnel flow, and it specifies SCC's filling ability in the absence of blockage and/or segregation. To some extent the flow duration of the V-Funnel trial is related towards plastic velocity.

3.3.3 L-Box Test

Passing capability for SCC is examined via L-Box trial. It measures the height ratio achieved by newly prepared SCC subsequently passing over the detailed holes of steel bars and passing inside a definite passage distance.

3.4. Casting of Spacemen

Three beam moulds were connected, and the inside of the mould was carefully greased. To maintain the desired effective depth of the beam, cover blocks of sufficient depth were placed at the bottom of the mould. The required volume of the materials for casting one batch of beams were mixed thoroughly on a platform to get a uniform mix. First, the reinforcement cage was kept on cover blocks in the mould. Then, the concrete is situated in the beam. Nine concrete cubes and nine cylinders were cast. The beam moulds, cylindrical moulds and cubes moulds were stripped after 24 hours of concreting.

3.5. Curing

After casting, all specimens were stored for curing to preserve the environmental conditions, namely a temperature of 27±2ºC later a day. Casted members were demoulded and placed in water for 28 days.

3.6. Compressive Strength Test

Compressive strength testing machine by means of a capacity of 2000kN referenced by IS:516-1959[10] was used to test the cube specimens.

The area of the sample that can sustain the greatest force applied to it during the test is used to calculate compressive strength. The compressive testing sample represented in Figure 3.1(a).

3.7. Split Tensile Strength Test

Trail was directed referenced by IS5816-1999[11]. Split tensile strength is an alternative test for direct tension. The standard 150mmx300mm chambers were utilized. The Split tensile strength testing illustration represented in Figure 3.1(b).

a  b  c 

Figure 3.1. (a)Compressive strength test;(b) Split tensile strength test;

(c) Flexural strength test

3.8. Flexural Strength Test

According to IS:516-1959[10] the flexural strength test was executed. As per Indian standards, I prepared the concrete specimen with a dimension of 150mm width, 150mm depth, and a span of 700mm.The test was carried out immediately after removing from the water and even when they are still wet. The load must be supplied without creating any disturbance. The load must be raised until the sample breaks, and the peak load attained during the experiment must be noted. Pictorially beam flexural strength testing shown in Figure 3.1(c).

3.9. Strengthening of Beams with FRP sheets

After flexural strength testing, an adhesive with a thickness of about 2mm was applied on the concrete surface, followed by two layers of CFRP sheets with a thickness of approximately 1mm. To eliminate any air bubbles and establish a strong adhesion, the sheets were applied with moderate pressure using a roller. Similarly, two layers of GFRP sheets were applied on beams. After wrapping FRP sheets, beams were maintained at room temperature for 24 hours before subjected to a flexural strength test. Figure 3.2(a) &Figure 3.2(b) signifies the flexural strength of RC beam strengthened by wrapping GFRP and CFRP sheets. Flexural Strength (fb)= pl/bd^2; Here, p = Extreme load induced in N; l = beam span; b = beam breadth; d = beam depth.

a  b 

Figure 3.2.(a) Flexural strength test after wrapping GFRP sheets;(b) Flexural strength test after wrapping CFRP sheet

4. RESULTS AND DISCUSSIONS

4.1. Workability

Filling and passing capability of SCC properties of SCC were measured according to code provisions IS:1199 part6, results were represented in Table 2.

Table 2 SCC workability test results

S.No	*Description*	*Experimental Results*
1.	Slump Flow(mm)	715
2.	T_{50} Slump flow in (sec)	3
3.	V-Funnel in (sec)	9.60
4.	T_{50} V-Funnel in (sec)	11.80
5.	L-Box ratio ($H2/H1$)	0.90
6.	$H1$ (mm)	104
7.	$H2$ (mm)	93

4.2. Compressive Strength Test

After 7, 21, and 28 days of curing the cubical members were assessed under compression loading and the mean value was reported as the 41.25MPa optimum after 28days of testing, experiment results were represented in Table 3.

Table 3Compressive strength in MPa

7 days	*21 days*	*28 days*
20.56	38.49	41.255

4.3. Split Tensile Strength Test

Examined at 7, 21, and 28 days after required curing finished. After 28days of testing, the mean value of these cylinders was reported as 4.65MPa optimum, Table 4 represents the experimental results.

Table 4Split Tensile strength in MPa

7 days	*21 days*	*28 days*
1.76	3.45	4.65

4.4. Flexural Strength Test

The RC beams were designed as under reinforced section and beams tested under three-point bending after 28 days of curing. As the steel provided was less than balancing steel then the concrete beam failure was observed at flexural zone and a crack initiated at support section was extended from bottom to the load point with 45 degrees angle.

The widening of cracks was observed after peak load was attained. It was observed that the maximum load value noted for conventional reinforced beam was 75 KN, whereas beams wrapped up with CFRP sheets were sustained a maximum load of 77.5 KN and load carrying capacity greatly increased in beam member GFRP wrapped resulted 78.9 KN.

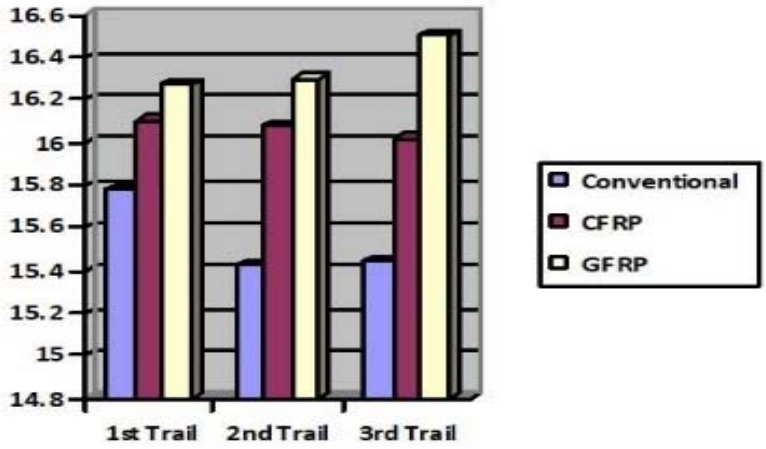


Figure 4.1Flexural strength test results

By Figure 4.1 it was clearly noted that strength value was enhanced by application of CFRP and GFRP, strength value noted for conventional reinforced concrete beam was 15.55 MPa, 16.07 MPa noted for beam member wrapped with CFRP sheets strength gain was 3.34 %. Beam members with GFRP sheets resulted 16.36 MPa strength was 5.20% more as contrasted to conventional RC beam.

5. CONCLUSIONS

The present research concluded on flexural performance of SCC-based beams of RC that were reinforced by wrapping FRP sheeting. The following conclusions are obtained from the experimental observations.

- SCC mix attained high flowability under filling and passing capability.
- Optimum compressive strength 41.25MPa was noted for cubes after 28days of testing.
- Maximum split tensile strength 4.65MPa observed after 28days trail investigation.
- Load carrying capability of RC beam remained enhanced after wrapping FRP sheets. It was observed that GFRP give more strength when compared CFRP.
- The flexural strength was 3.34% enhanced by wrapping CFRP sheets whereas there is 5.20% over conventional beam by wrapping GFRP sheets. Strength value improved for a flexure member as results of adding FRP sheets.
- Flexural strength is higher in RC beams with GFRP sheets than in RC beams with CFRP sheets.

6. REFERENCES

[1] Okamura H & Ozawa K, “Mix-design for self-compacting concrete”, Concrete Library of JSCE, No. 25, June 1995, pp.107-120. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[2] Nan Su, Kung-Chung Hsu, His-Wen Chai, “A simple mix design method for self-compacting concrete”, Cement and Concrete Research, 6 June 2001, pp1799-1807.

[3] Grace NF, Sayed GA, and Saleh KR, “Strengthening of continuous beams using Fiber reinforced polymer laminates”, American Concrete Institute, Farmington Hills, Mich, pp. 647-657, 1999.

[4] El-Refaie SA, Ashour AF, and Garrity SW, “CFRP strengthened continuous concrete beams”, Proceedings of the ICE - Structures and Buildings, pp. 395 - 404, 2003.

[5] N. Pannirselvam., V. Nagaradjane, K. Chandramouli.,” Strength behavior of fiber reinforced polymer strengthened beam” ARPN journal of Engineering and Applied science, vol. 4, no. 9, pp. 34-39, 2009.

[6] IS: 12269- 2013 “Ordinary Portland Cement 53 Grade- Specifications” Bureau of Indian Standards.

[7] IS: 383-2016 “Coarse and fine aggregate for concrete specification” Bureau of Indian Standards. New Delhi.

[8] IS: 10262-2019 “Concrete mix proportioning- guide lines” Bureau of Indian Standards. New Delhi.

[9] IS: 1199-2018 part 6 “Method of sampling and analysis of concrete” Bureau of Indian Standards. New Delhi.

[10] IS: 516-1959 “Methods of tests for strength of concrete” Bureau of Indian Standards. New Delhi.

[11] IS: 5816-1999 “Method of test splitting tensile strength of concrete” Bureau of Indian Standards.

[12] IS: 456-2000, Code of practice for plain and reinforced concrete. Bureau of Indian Standards, 2000.

Biographies

Kandula Kiran received the bachelor's degree in civil engineering from JawaharlalNehruTechnologicalUniversity in 2019, pursuing the master's degree in structural engineering fromInstitute of Aeronautical engineering. His research area includesretrofitting and restrengthening of RC beams using FRP technology.

Nimmagadda Venkata Rao received the bachelor's degree in civil engineering from AndraUniversity, the master's degree in structural engineering from JawaharlalNehruTechnologicalUniversity and pursing Ph.D.from VIT University.He is currently working as an Associate Professor in the Department of Civil Engineering, Institute of Aeronautical Engineering, Hyderabad. His research areasincludeconcrete and special concretes. He published 40 research papers in various national and international journals.

Efficacy of GWO Algorithm by Varying One Algorithm-Specific Parameter

[1] P. D. Dewangan, [2] N. Patnana, [3] Lalbihari Barik, [4] P. J. Krishna, [5] V. P. Meena, and [6] V. P. Singh

[1]*Dept. of Electrical Engineering NIT Raipur, Raipur, India, pddewangan.ele@nitrr.ac.in*

[2]*Dept. of Electrical, Electronics and Comm. Engg., GITAM University, Vizag, India, npatnana@gitam.edu*

[3]*Dept. of Information Systems, King Abdulaziz University, Kingdom of Saudi Arabia, lalbihari@gmail.com*

[4]*Department of Electrical Engineering, MNIT Jaipur, Jaipur, India, puthi.jayakrishna@yahoo.com*

[5]*Department of Electrical Engineering, MNIT Jaipur, Jaipur, India, vmeena1@ee.iitr.ac.in*

[6]*Department of Electrical Engineering, MNIT Jaipur, Jaipur, India, vinay.ee@mnit.ac.in*

Abstract.

This paper compares the performance of the grey wolf optimization (GWO) algorithm for four distinct exploration approaches. These exploration strategies are used throughout the iterative procedure, with the co-efficient vector being modified for different dimensions of each solution. The co-efficient vector, $\overline{X}$ is linearly changed in the first exploration strategy from 1 to 0. While, in second, third, and fourth strategies, the co-efficient vector, $\overline{X}$ is varied linearly from 2 to 0, 10 to 0, and 20 to 0 respectively. Thus, modified GWO algorithm is applied on five different unimodel benchmarks functions for performance comparison. The performance comparison is done on the basis of results obtained for the value of objective function (i.e. figure of merit), standard deviation, mean, minimum and maximum values of the figure of merit.

Keywords. Grey Wolf Optimization, Exploration; Meta-heuristics, Swarm Intelligence, Unimodal Benchmark Functions.

1. INTRODUCTION

Over the last few decades, meta-heuristic optimization approaches have been increasingly popular. These optimization techniques are used in a wide range of research fields as well as in a large number of applications. This is because of advantages such as simplicity, flexibility, avoiding local optima, no requirement of derivation. Meta-heuristics optimization techniques are easy to learn and apply to the existing optimization problems. These optimization techniques are implemented by numerous scholars because of ease of use in replicating many natural behaviours. Also, new meta-heuristics by merging two or more meta-heuristics or strengthening the current meta-heuristics with an improvement are suggested in literature. Furthermore, the simplicity of the meta-heuristics helps researchers from inter-disciplinary branches in quickly understanding and applying them to their own optimization problems.

The next most important advantage of using these meta-heuristics to optimize the problems is their flexibility. Meta-heuristics are easily adaptable to wide range of issues as they presume the problems to be black boxes. In simple context, meta-heuristics only examine the system's input(s) and output(s). A designer just has to know how to articulate his or her problem by understanding these meta-heuristics.

The vast majority of meta-heuristics include procedures that do not need derivation. Meta-heuristics, in contrast to gradient-based optimization methods, approach issues in a stochastic way. The optimization process starts with random solution(s) to find the best, and there is no need to compute the derivative of such search spaces. As a result, meta-heuristics are well-suited for real-world scenarios where derivative information is not known.

Meta-heuristics are better than conventional optimization techniques in avoiding local optima. This is because meta-heuristics are stochastic in nature. This stochastic nature allows them to avoid trappings in local solution and thoroughly search the whole search space.

Swarm intelligence (SI) is one of the most important branch of population-based meta-heuristics. Beni and Wang [1] proposed SI for the first time in 1993. Natural colonies are a primary source of SI based methods [2]. These SI optimization algorithms generally emulate the social behaviour of swarms, herds, flocks, or schools of organisms in nature. SI optimization techniques provide a number of benefits, including:

- These optimization techniques are simple to implement.
- These optimization methods contain fewer operators than evolutionary strategies.
- There are usually fewer parameters to configure with these optimization techniques.
- These optimization methods maintain track of information about the search space during the course of iterations.
- These optimization methods usually require memory to preserve the best solution identified.

Along withstanding the benefits indicated above, SI optimization strategies have certain disadvantages too. The search agent activity seems to be noisy because the decision action is stochastic. Without understanding how the search agent works, it is impossible to know the functions of colony. Anticipating behaviour based on a set of specified rules is tough. Even little modifications to the basic rules have a significant influence on group behaviour.

Within the due course of time, many other SI optimization techniques are proposed in the literature providing good results. Some of them are particle swarm optimization [3], [4], grey wolf optimization [5], [6], cat and mouse based optimizer [7], elephant herding optimization [8], [9], jaya algorithm [10], [11], rock hyraxes swarm optimization [12], teacher-learner-based optimization [13], symbiotic organisms search optimization [14], [15], differential evolution algorithm [16], honey badger algorithm [17], sine cosine algorithm [18], [19], whale optimization algorithm [20], etc. These algorithms have also been modified and published in the literature. Furthermore, hybridising two or more algorithms improves the performance of these algorithms.

In this article, the efficacy of grey wolf optimization (GWO) algorithm is investigated by varying one algorithm-specific parameter. This algorithm-specific parameter is required in exploration. In parent GWO algorithm [5], this algorithm is varied linearly from 2 to zero as the number of iterations increase. However, in this article, three other ranges are taken for this parameter. Then, efficacy of GWO algorithm is tested for four variations (one parent and other three suggested in this article). The performance is tested for five benchmarks functions.

The organization of this contribution is as follows: In section II, GWO is described in detail and modification in GWO algorithm is proposed. Section III deals with the simulation results and discussion. Finally, the whole work is concluded in section IV.

2. GREY WOLF OPTIMIZATION ALGORITHM

Grey wolf optimization (GWO) is a revolutionary swarm intelligence system based on grey wolf leadership hierarchy and prey hunting. Mirjalili et al. [21] produced GWO in 2014. Canis lupus is the scientific name for grey wolves, and they belong to the canidae family. These grey wolves are top-tier predators in the food chain. Grey wolves like to live in packs, which have an average of 5 to 12 members.

2.1. Description of GWO algorithm

The pack's social order, surrounding prey, hunting, attacking prey, and looking for prey are all examples of grey wolves' behaviour. These grey wolves in the pack have an extremely strict social hierarchy. Grey wolves are divided into four groups depending on social hierarchical dominance α, β, γ, and ω. The leader is at the α level and is in charge of making decisions for the pack. The decisions taken by α are dictated to the other wolves in the pack, and the entire pack must obey them. The wolf in this α level is the greatest at sustaining pack discipline and structure, demonstrating that group organisation is more essential than collective strength. The next level in the grey wolf social hierarchy is β, and the wolves in this level are in charge of aiding the α level in making decisions and other collective interests. In the absence of a α level, the wolves in the β level function as the leader.

The next level in the grey wolf social hierarchy is gamma, which includes scouts, sentinels, hunters, and caretakers. Scouts are in responsible of keeping an eye on the territory's boundaries and informing the pack if any threat persists, while sentinels are in charge of defending and ensuring the pack's safety. Hunters help α and β level wolves by hunting animals and providing food for the pack. The caretakers must look after the pack's weak, injured, and wounded wolves. The pack's lowest level is ω. At all times, the wolves at this rank must surrender to all other dominant wolves. The next most important social behaviour among grey wolves is that they always engage in groups while hunting. This encircling behaviour of grey wolves is given by

$$\overline{C} = \left|\overline{Y} \cdot \bar{J}_p - \bar{J}\right| \tag{1}$$

$$\bar{J}i + 1 = \bar{J}p - \overline{X} \cdot \overline{C} \tag{2}$$

where $\bar{J}p$ represents the position vector of prey, $\bar{J}$ indicates position vector of wolves in current iteration, $\bar{J}i + 1$ illustrates the position vector of wolves in next iteration, $\overline{X}$ and $\overline{Y}$ are co-efficient vectors which are calculated as

$$\overline{X} = 2\overline{x} \cdot \overline{r}_1 - \overline{x} \tag{3}$$

$$\overline{Y} = 2 \cdot \overline{r}_2 \tag{4}$$

where, $\overline{X}$ decreases linearly from 2 to 0 as the number of iterations increase, $\overline{r}_1$ and $\overline{r}_2$ are random vectors in the range of [0,1].

Grey wolves have the ability to track down and encircle their prey. In most cases, the α level is in charge of hunting. Hunting is a skill that β and δ level wolves can exhibit occasionally. However, in search space, we are

unable to determine where the best prey position is located. We suppose that the α level is the best candidate, although the $beta$ and $delta$ levels have a better understanding of the prey's likely position. We preserve the top three best solutions identified so far and need additional search components to update their positions in line with the best search component. This behavioural aspect is considered in order to mathematically simulate the hunting behaviour of grey wolves. The following formulae are stated in this regard:

$$\overline{J}i+1=\frac{\overline{J}\alpha+\overline{J}\beta+\overline{J}\gamma}{3} \tag{5}$$

$$\overline{J}\alpha=\overline{J}\alpha,i-\overline{X}1\cdot\overline{C}\alpha \tag{6}$$

$$\overline{J}\beta=\overline{J}\beta,i-\overline{X}2\cdot\overline{C}\beta \tag{7}$$

$$\overline{J}\gamma=\overline{J}\gamma,i-\overline{X}3\cdot\overline{C}\gamma \tag{8}$$

$$\overline{C}\alpha=\left|\overline{Y}_1\cdot\overline{J}\alpha,i-\overline{J}\right| \tag{9}$$

$$\overline{C}\beta=\left|\overline{Y}_2\cdot\overline{J}\beta,i-\overline{J}\right| \tag{10}$$

$$\overline{C}\gamma=\left|\overline{Y}_3\cdot\overline{J}\gamma,i-\overline{J}\right| \tag{11}$$

where, $\overline{J}i+1$ interprets position of wolves in next iteration, $\overline{J}\alpha$, $\overline{J}\beta$ and $\overline{J}\gamma$ represent new position of α level, β level and γ level wolves respectively; $\overline{J}\alpha,i$, $\overline{J}\beta,i$ and $\overline{J}\gamma,i$ indicate the current position of α level, β level and γ level wolves respectively; $\overline{C}\alpha$, $\overline{C}\beta$ and $\overline{C}\gamma$ refer to the encircling behaviour of α level, β level and γ level wolves respectively.

2.2. Amendment to GWO algorithm

In parent GWO algorithm, the co-efficient vector, $\overline{X}$ is presumed to be decreasing linearly from 2 to 0 as the iterations increase in number. The modification in the value of this co-efficient vector, $\overline{X}$ tends to modify the searching capability of the GWO algorithm. In this contribution, the efficacy of GWO algorithm is tested by modifying the exploration of algorithm. The exploration is modified by varying the value of $\overline{X}$ considering the following four cases.

1. By varying the value of $\overline{X}$ from 1 to 0 linearly.
2. By varying the value of $\overline{X}$ from 2 to 0 linearly.
3. By varying the value of $\overline{X}$ from 10 to 0 linearly.
4. By varying the value of $\overline{X}$ from 20 to 0 linearly.

3. RESULT AND DISCUSSION

Four different exploration strategies are proposed in previous section and are used with GWO algorithm one by one. For presenting the results, the modified GWO is used for minimizing following five benchmarks functions.

A. Sphere Function

$$F_1(x)=\sum_{i=1}^{n}{x_i}^2 \tag{12}$$

B. Schwefel 2.22 Function

$$F_2(x)=\sum_{i=1}^{n}|x_i|+\prod_{i=1}^{n}|x_i| \tag{13}$$

C. Schwefel 1.2 Function

$$F_3(x)=\sum_{i=1}^{n}\left(\sum_{j-1}^{i}x_j\right)^2 \tag{14}$$

D. Schwefel 2.21 Function

$$F_4(x)=max\{|x_i|,1\le i\le n\} \tag{15}$$

E. Rosenbrock Function

$$F_5(x)=\sum_{i=1}^{n-1}\left[100(x_{i+1}-{x_i}^2)^2+(x_i-1)^2\right] \tag{16}$$

The following four cases are considered for testing efficacy of the GWO algorithm:

1. Case 1: When $\overline{X}$ is varied from 1 to 0 linearly.
2. Case 2: When $\overline{X}$ is varied from 2 to 0 linearly.
3. Case 3: When $\overline{X}$ is varied from 10 to 0 linearly.
4. Case 4: When $\overline{X}$ is varied from 20 to 0 linearly.

A total of 100 solutions are explored while minimizing the five benchmark functions considered. The number of iterations, on the other hand, is taken to be 100. The algorithm is repeated five times in a row. The statistical analysis is based on the outcomes of five successive runs.

Table 1: Figure of merit (FOM) for dimension of $D = 10$

$D = 10$				
	Case-1	Case-2	Case-3	Case-4
F1	**3.6954e-113**	5.6107e-97	7.7927	806.4701
	2.3269e-114	5.1516e-20	9.3672	581.2886
	1.4317e-115	4.5816e-95	13.9832	230.8661
	3.8480e-115	2.1863e-98	24.0044	194.6897
	8.2018e-115	5.6902e-97	6.1811	360.3082
F2	**2.4177e-59**	6.0033e-50	1131.5887	6.3369e+6
	3.6649e-59	3.3865e-49	49.2205	3405.4965
	1.4559e-58	9.3874e-51	34.8568	1.1747e+7
	8.9506e-59	1.5391e-50	23.8319	1.9239e+5
	7.1150e-59	9.4559e-50	36.6730	1.9504e+5
F3	0	0	0	0
	0	0	0	0
	0	0	0	0
	0	0	0	0
	0	0	0	0
F4	**1.4726e-55**	4.0766e-47	5.2664	19.7313
	4.2422e-55	3.5396e-46	5.4806	25.6629
	3.3699e-54	4.6897e-48	5.6047	0.8569
	5.9530e-56	1.0689e-46	9.0666	0.6890
	5.7683e-55	6.4228e-46	6.1606	25.2770
F5	8.7541	**8.1003**	1.1269e+5	6.7109e+7
	8.5925	**8.1014**	4.7527e+4	2.2580e+7
	8.7685	**8.7292**	8566.5078	1.6725e+7
	8.0738	8.1002	8.1725e+4	4.2854e+6
	8.7003	**8.0723**	5.4835e+5	5.0827e+6

Table 2: Figure of merit (FOM) for dimension of $D = 20$

$D = 20$				
	Case-1	Case-2	Case-3	Case-4
F1	**1.1333e-103**	1.1491e-86	2019.8435	11011.4692
	1.9705e-104	2.3098e-87	2024.3171	7629.1472
	5.9303e-105	2.6469e-88	1523.5633	8539.1056
	3.1142e-104	2.6261e-89	1977.4169	8483.1347
	1.1380e-103	6.9748e-87	2403.7282	38627.4326
F2	**8.2755e-54**	3.8488e-45	1.0319e+14	6.9821e+16
	2.3608e-53	1.1079e-45	6.9370e+13	3.9841e+16
	2.5641e-53	1.6536e-45	2.2913e+12	5.5195e+15
	4.6818e-53	1.7741e-46	2.3862e+9	1.6333e+21
	2.6256e-53	1.8730e-45	3.8336e+12	3.4570e+17
F3	0	0	0	0
	0	0	0	0
	0	0	0	0
	0	0	0	0
	0	0	0	0
F4	**2.2649e-51**	3.9997e-42	0.3277	12.6963
	8.6926e-50	2.1907e-42	0.1535	7.3062
	4.4330e-51	5.0165e-41	25.1857	53.4712
	1.8159e-50	3.9018e-42	19.6380	2.9966
	4.4769e-51	3.1600e-41	36.0258	10.7821
F5	18.9093	**18.0925**	1.3002e+8	5.6713e+8
	18.8884	**18.6166**	5.1006e+7	9.5274e+8
	18.6499	**18.0848**	2.3117e+7	3.7050e+8
	18.8924	**18.6972**	8.4099e+8	8.0334e+8
	18.6815	**18.0722**	2.5234e+7	1.5402e+9

In this contribution, all benchmark functions are minimized with dimension of $D = 10$ and $D = 20$. The values of objective functions or figure of merit of five benchmark functions are represented in Table 1 for each of the four scenarios independently for dimension $D = 10$. Similarly, Table 2 presents the figure of merits of five benchmark functions for dimension $D = 20$ when all four cases are considered independently.

The statistical analysis of all four cases evaluated for dimension $D = 10$ is shown in Table 3. The values of standard deviation (SD), mean, minimum, and maximum of figure of merit (FOM) are tabulated in this table for each case. Table 4 shows similar findings for all four cases when the dimension is set to $D = 20$.

Table 3: Statistical Analysis for dimension of $D = 10$

$D = 10$					
		Case-1	Case-2	Case-3	Case-4
F_1	SD	**1.6137e-113**	2.3038e-20	7.1794	257.0936
	Mean	**8.1258e-114**	1.0303e-20	12.2657	434.7245
	Min	**1.4317e-115**	2.1863e-98	6.1811	194.6897
	Max	3.6954e-113	**5.1516e-20**	24.0044	806.4701
F_2	SD	**4.8086e-59**	1.3591e-49	489.9798	5.2434e+6
	Mean	**7.3414e-59**	1.0360e-49	255.2342	3.6951e+6
	Min	**2.4177e-59**	9.3874e-51	23.8319	3405.4965
	Max	1.4559e-58	**3.3865e-49**	1131.5887	1.1747e+7
F_3	SD	0	0	0	0
	Mean	0	0	0	0
	Min	0	0	0	0
	Max	0	0	0	0
F_4	SD	**1.3877e-54**	2.6789e-46	1.5728	12.6982
	Mean	**9.1555e-55**	2.2972e-46	6.3158	14.4434
	Min	**5.9530e-56**	4.6897e-48	5.2664	0.6890
	Max	3.3699e-54	**6.4228e-46**	9.0666	25.6629
F_5	SD	0.2901	**0.2845**	2.2066e+5	2.5769e+7
	Mean	8.5778	**8.2207**	1.5977e+5	2.3156e+7
	Min	8.0738	**8.07239**	8566.5078	4.2854e+6
	Max	**8.7685**	8.7292	5.4835e+5	6.7109e+7

Table 4: Statistical Analysis for dimension of $D = 20$

$D = 20$					
		Case-1	Case-2	Case-3	Case-4
F_1	SD	**5.2601e-104**	4.9332e-87	312.6342	1269.1422
	Mean	**5.6783e-104**	4.2134e-87	1989.7738	8858.0579
	Min	**5.9303e-105**	2.6261e-89	1523.5633	7629.1472
	Max	1.1380e-103	**1.1491e-86**	2403.7282	11011.4692
F_2	SD	**1.3724e-53**	1.3521e-45	4.7683e+13	7.3042e+20
	Mean	**2.6119e-53**	1.7321e-45	3.5737e+13	3.2677e+20
	Min	**8.2755e-54**	1.7741e-46	2.3862e+9	5.5195e+15
	Max	4.6818e-53	**3.8488e-45**	1.0319e+15	1.6333e+21
F_3	SD	0	0	0	0
	Mean	0	0	0	0
	Min	0	0	0	0
	Max	0	0	0	0
F_4	SD	**3.6150e-50**	2.1584e-41	15.7720	20.4718
	Mean	**2.3252e-50**	1.8371e-41	16.2661	17.4505
	Min	**2.2649e-51**	2.1907e-42	0.1535	2.9966
	Max	8.6926e-50	**5.0165e-41**	36.0258	53.4712
F_5	SD	**0.1272**	0.3156	4.3985e+7	4.4695e+8
	Mean	18.8043	**18.3127**	5.4074e+7	8.4680e+8
	Min	18.6499	**18.0722**	2.3117e+7	3.7050e+8
	Max	**18.9093**	18.6972	1.3002e+8	1.5402e+9

When comparing data in Tables 1-4, bold face data represents the better value than the equivalent normal face data. From Tables 1-2, it is clear that thirty-one times results are better in case 1. However, better results are obtained nine times in case 2. In cases 3 and 4, none result is better in comparision to case 1 and case 2. This proves that case 1, where $\overline{X}$ is varied from 1 to 0 linearly, is better than case 2, case 3 and case 4.

The same is also proved with the statistical analysis presented in Tables 3-4. In Tables 3-4, minimum value of FOM is obtained nineteen times in case 1. However, it is obtained only five times in case 2. Also, the maximum value of FOM is obtained only six times in case 2 in comparison to two times as obtained in case 1. So, with the results presented in 1-4, it is clear that by varying $\overline{X}$ from 1 to 0 linearly is better than varying $\overline{X}$ from 2 to 0 linearly, 10 to 0 linearly, or 20 to 0 linearly.

4. CONCLUSION

This article provides the comparison of performance of grey wolf optimization (GWO) algorithm with four different exploration strategies. This experimentation of initializing the co-efficient vector, $\overline{X}$ from different values simulate the divergence towards or away from the best solution. It is clear from results that the divergence of function is towards the best solution when the co-efficient vector, $\overline{X}$ declines linearly from 1 to 0. The performance of modified GWO algorithm with four different exploration strategies is utilized further for comparison of five unimodel benchmark functions. The results are tabulated in terms of value of objective function, standard deviation, mean, minimum and maximum values of objective function. Further, it can be concluded with results that we obtain the best solution when the co-efficient vector, $\overline{X}$ declines linearly from 1 to 0 instead of $\overline{X}$ declining linearly from 2 to 0, linearly from 10 to 0, or linearly from 20 to 0.

The study in this contribution should be extended to other optimization techniques as well. Furthermore, the same approach could be extended to different benchmark functions.

5. REFERENCES

[1] G. Beni and J. Wang, "Swarm intelligence in cellular robotic systems," in *Robots and biological systems: towards a new bionics?*, Springer, 1993, pp. 703–712.

[2] E. Bonabeau, G. Theraulaz, M. Dorigo, G. Theraulaz, and D. de R. D. F. Marco, *Swarm intelligence: from natural to artificial systems*. Oxford university press, 1999.

[3] T. M. Shami, A. A. El-Saleh, M. Alswaitti, Q. Al-Tashi, M. A. Summakieh, and S. Mirjalili, "Particle Swarm Optimization: A Comprehensive Survey," *IEEE Access*, 2022.

[4] D. Sedighizadeh, E. Masehian, M. Sedighizadeh, and H. Akbaripour, "GEPSO: A new generalized particle swarm optimization algorithm," *Math. Comput. Simul.*, vol. 179, pp. 194–212, 2021.

[5] N. S. Rathore, V. P. Singh, and B. Kumar, "Controller design for doha water treatment plant using grey wolf optimization," *J. Intell. Fuzzy Syst.*, vol. 35, no. 5, pp. 5329–5336, 2018.

[6] G. Vashishtha and R. Kumar, "An amended grey wolf optimization with mutation strategy to diagnose bucket defects in Pelton wheel," *Measurement*, vol. 187, p. 110272, 2022.

[7] M. Dehghani, Š. Hubálovskỳ, and P. Trojovskỳ, "Cat and mouse based optimizer: a new nature-inspired optimization algorithm," *Sensors*, vol. 21, no. 15, p. 5214, 2021.

[8] S. Gupta, V. P. Singh, S. P. Singh, T. Prakash, and N. S. Rathore, "Elephant herding optimization based PID controller tuning," *Int. J. Adv. Technol. Eng. Explor.*, vol. 3, no. 24, p. 194, 2016.

[9] G.-G. Wang, S. Deb, and L. dos S. Coelho, "Elephant herding optimization," in *2015 3rd international symposium on computational and business intelligence (ISCBI)*, 2015, pp. 1–5.

[10] S. P. Singh, T. Prakash, V. P. Singh, and M. G. Babu, "Analytic hierarchy process based automatic generation control of multi-area interconnected power system using Jaya algorithm," *Eng. Appl. Artif. Intell.*, vol. 60, pp. 35–44, 2017.

[11] W. Warid, H. Hizam, N. Mariun, and N. I. Abdul-Wahab, "Optimal power flow using the Jaya algorithm," *Energies*, vol. 9, no. 9, p. 678, 2016.

[12] B. Al-Khateeb, K. Ahmed, M. Mahmood, and D.-N. Le, "Rock hyraxes swarm optimization: A new nature-inspired metaheuristic optimization algorithm," *Comput. Mater. Contin.*, vol. 68, no. 1, pp. 643–654, 2021.

[13] V. P. Singh, T. Prakash, N. S. Rathore, D. P. Singh Chauhan, and S. P. Singh, "Multilevel thresholding with membrane computing inspired TLBO," *Int. J. Artif. Intell. Tools*, vol. 25, no. 06, p. 1650030, 2016.

[14] N. S. Rathore, V. P. Singh, and B. D. H. Phuc, "A modified controller design based on symbiotic organisms search optimization for desalination system," *J. Water Supply Res. Technol.-Aqua*, vol. 68, no. 5, pp. 337–345, 2019.

[15] A. E. Ezugwu and D. Prayogo, "Symbiotic organisms search algorithm: theory, recent advances and applications," *Expert Syst. Appl.*, vol. 119, pp. 184–209, 2019.

[16] V. P. Meena, N. Kumar, R. K. Lenka, R. K. Barik, and V. P. Singh, "DE Based Reduced-Order Modeling for Interval Modeling Doha Water Treatment Plant," in *2021 19th OITS International Conference on Information Technology (OCIT)*, 2021, pp. 428–432.

[17] F. A. Hashim, E. H. Houssein, K. Hussain, M. S. Mabrouk, and W. Al-Atabany, "Honey Badger Algorithm: New metaheuristic algorithm for solving optimization problems," *Math. Comput. Simul.*, vol. 192, pp. 84–110, 2022.

[18] V. P. Singh, "Sine cosine algorithm based reduction of higher order continuous systems," in *2017 International Conference on Intelligent Sustainable Systems (ICISS)*, 2017, pp. 649–653.

[19] V. P. Meena, S. Singh, M. Kandpal, R. K. Barik, and V. P. Singh, "SCA Assisted Reduced-Order Modelling of Interval Modelled Doha Water Treatment Plant," in *2021 IEEE 2nd International Conference on Applied Electromagnetics, Signal Processing, & Communication (AESPC)*, 2021, pp. 1–6.

[20] N. S. Rathore and V. P. Singh, "Whale optimisation algorithm-based controller design for reverse osmosis desalination plants," *Int. J. Intell. Eng. Inform.*, vol. 7, no. 1, pp. 77–88, 2019.

[21] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Adv. Eng. Softw.*, vol. 69, pp. 46–61, 2014.

Data Mining based Diseases Classification

[1]Archika Jain, [2]Devendra Somwanshi, [3]Barkha Narang

[1,2,3]Assistant Professor, [1,3]Department of CSE &[2]EC

[1,2,3]Poornima College of Engineering, Jaipur, India

[1]archikaagarwal@gmail.com, [2]imdev.som@gmail.com, [3]barkhanarang17@gmail.com

Abstract.

Data mining is fast gaining traction in a variety of fields, including organic chemical research, financial forecasting, healthcare, and weather forecasting. Data mining in healthcare is a rapidly growing discipline that may help with prognosis and a better understanding of medical data. Investigation of infirmary for finer well-being decision-making and determent of health centre mistake, early disclosure, and determent of ailment and avertible health centre demise, more utility for funds and cost effectives, and discernment of crooked security claims are all examples of data mining applications in healthcare. Data mining techniques are being used in the detection of a variety of ailments, including diabetes, stroke, cancer, and heart disease. We employ two types of datasets in this study: breast cancer and diabetes databases. We use the WEKA tool to put the techniques into practice. On breast cancer dataset, MLP is better error-free classifier in contrast of remaining with the highest accuracy i.e. 74.12%. On diabetes dataset, SMO is better accurate classifier in contrast of others with the highest accuracy i.e. 79.30%**.**

Keywords. WEKA tool, Data mining, Diseases**.**

1. INTRODUCTION

Data mining is an action of analysing enormous data bank to uncover previously unknown patterns, correlations, and information that would be difficult to identify using standard statistical approaches. Data mining is a computational approach that imply the use of intelligent retrieval, predictive analytics, and data bank arrangement to discover figure in enormous number of data file [5]. The extensive motive of the data extract action is to bring out facts from a data cluster and turn it into a formation. Data bank and data administration matter, data pre-processing imitation and inference deliberations, allure measures, difficulty deliberations, post-processing of establish forms, perceptions, and online refurbish are all part of it [3].

In terms of data mining applications, the healthcare business is essential since it generates an extensive range of repository that varies in magnitude, diversity, and rapidness. Condemnatory illnesses such as lymphoma, pulmonary disease, as well as diabetes are among the world's top causes of mortality [6]. Vital information may be obtained from a huge database using data mining tools and methodologies, providing an easy manner for

prophylactic educator to make dominant choices and enhance restorative aid [6]. WEKA is used in view of the fact that it allows ourselves quickly assess and collate knowledge discovery in data algorithms on actual facts [7]. It is now feasible to forecast many disorders more accurately because to developments in computing technology supplied by computer science technologies. As illustrated in Fig 1, data mining may be separated into sub-processes that include data selection, pre-processing, transformation, data mining, and ultimately data interpretation [8]. The classification approach is commonly employed in the health and medical fields. It gives a step-by-step method for creating a classifier model using training data, which is subsequently tested using test data and used to make predictions.

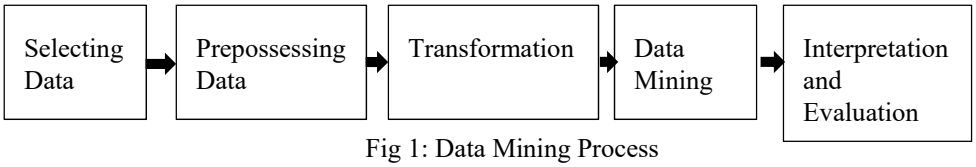


Fig 1: Data Mining Process

1.1 ***Data Excavate Job:*** Fact extraction job may be categorized:

- Prognostic imitation
- Depictive imitation

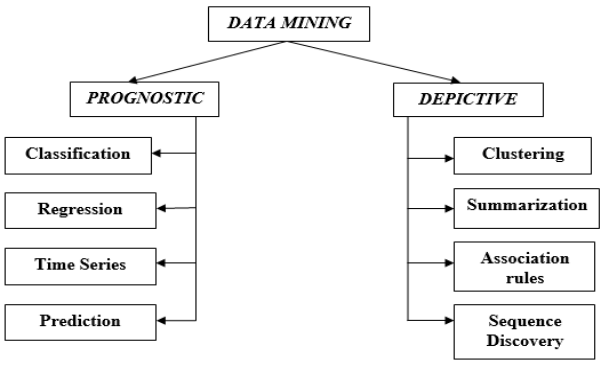


Fig 2: Data Excavate Job

2. CLASSIFICATION IS USED IN PROPOSED WORK

Table 1 List of Attributes

Dataset Name	No. of Attributes	No. of Instances	Attributes
Breast Cancer	9	296	Period, menopause, swell size, bosom canker, vertex
Diabetes	10	788	Pregnant, claret, pressure, skin steroid, paediatric

2.1 Flow Charts of Proposed System:

A flow chart and steps of proposed work as shown:

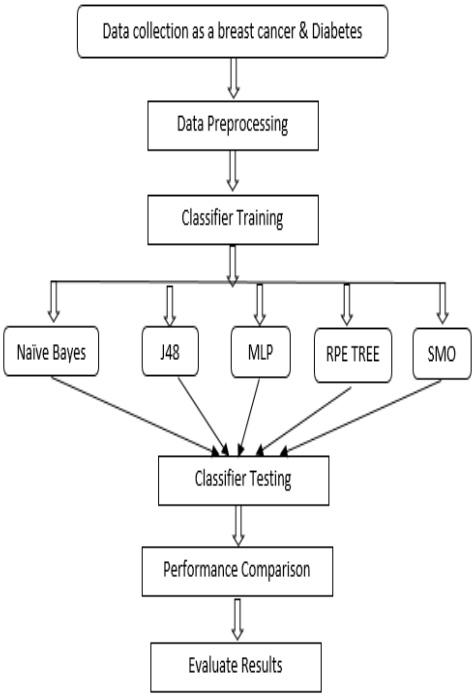


Fig 3: Flow Chart

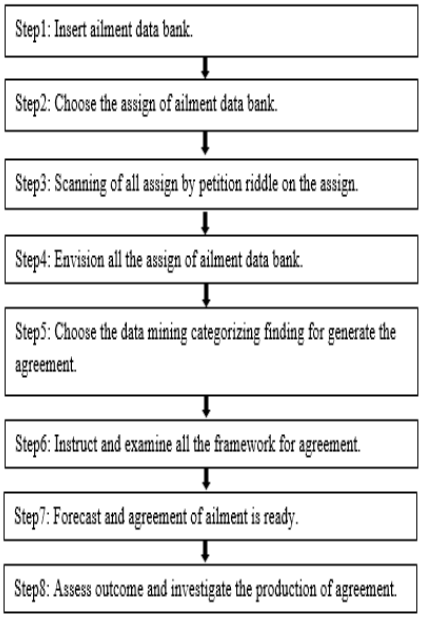


Fig 4: Steps of Proposed Work

3. EXPERIMENTAL WORK

3.1 Breast cancer dataset attributes description:

- Period [5], Menopausal, Quadruple-nodes [3], Fork top, Malignancy [7], Bosom canker [5], Bosom vertex, Betterment, Classification [1].

3.2. Diabetes dataset attributes description:

- Pregnant, Claret, Pressure, Skin, Steroid [3], Multitude, Pediatric [1], Lifetime, Class.

3.3. Performance of classifiers on Breast cancer dataset:

Table 2 Performance of classifiers on Breast cancer dataset

Evaluation Criteria	**Classifiers**				
	Naive Bayes	**J48**	**MLP**	**RPE Tree**	**SMO**
Correctly classified instances	69	66	72	64	68
Incorrectly classified instances	28	31	25	33	29
Accuracy (%)	71.10	68.0	74.12	65.95	70.00

From above table 2 we can conclude that on breast cancer dataset, MLP is better error-free classifier in contrast of others also it is clearly observed meaning it has a higher proportion

of correctly categorized occurrences and a lower proportion of mistakenly classified instances than Naive Bayes, SOM, Rep Tree and J48.

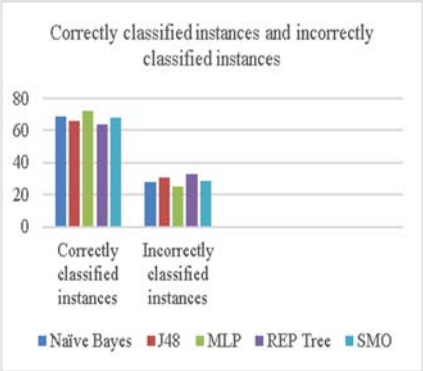


Fig 5: Representational of instances on Breast cancer

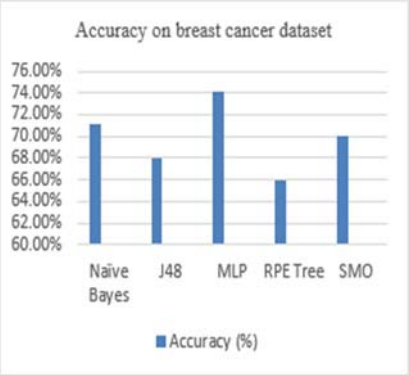


Fig 6: Diagrammatic representation of accuracy on Breast cancer

3.4. Performance of classifiers on Diabetes dataset:

Table 3 Performance of classifiers on Diabetes dataset

Evaluation Criteria	Classifiers				
	Naive Bayes	**J48**	**MLP**	**RPE Tree**	**SMO**
Correctly classified instances	201	199	194	197	207
Incorrectly classified instances	60	62	67	64	54
Accuracy (%)	77.00	76.2	74.30	75.45	79.30

From above table 3, SMO is better error-free classifier and has a higher proportion of correctly categorized occurrences and a lower proportion of mistakenly classified instances than Naive Bayes, MLP, REP Tree and J48.

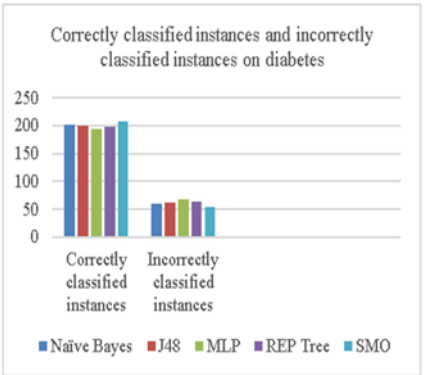


Fig 7: Representational of instances on Diabetes

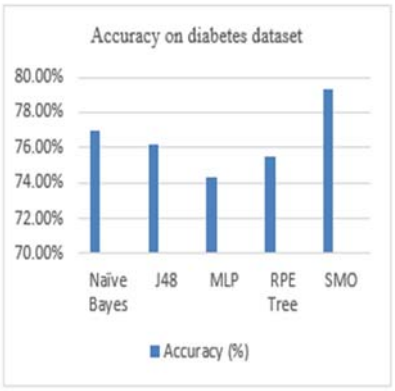


Fig 8: Diagrammatic representation of Accuracy on Diabetes

4. CONCLUSION & FUTURE WORK

We culminate that WEKA instrument is examine as one of the finest apparatus for data extraction categorizing. In this we have used two types of datasets breast Cancer datasets

and Diabetes datasets. And obtain the good results by using the all attributes. On breast cancer dataset, MLP is better error-free classifier in contrast of others also it is clearly observed meaning it has a higher proportion of correctly categorized occurrences and a lower proportion of mistakenly classified instances than Naive Bayes, SOM, Rep Tree and J48. On diabetes dataset, SMO is better error-free classifier in contrast of others also it is clearly observed meaning it has a higher proportion of correctly categorized occurrences and a lower proportion of mistakenly classified instances than Naive Bayes, MLP, REP Tree and J48.

We used percentage split test option for testing the parameters, so in future we will use other testing mode options to increase accuracy of classifiers.

5. REFERENCES

1. S. ALGHUNAIM AND H. H. AL-BAITY, "ON THE SCALABILITY OF MACHINE-LEARNING ALGORITHMS FOR BREAST CANCER PREDICTION IN BIG DATA CONTEXT," IN *IEEE ACCESS*, VOL. 7, PP. 91535-91546, 2019, DOI: 10.1109/ACCESS.2019.2927080.
2. V. MHETRE AND M. NAGAR, "CLASSIFICATION BASED DATA MINING ALGORITHMS TO PREDICT SLOW, AVERAGE AND FAST LEARNERS IN EDUCATIONAL SYSTEM USING WEKA," *2017 INTERNATIONAL CONFERENCE ON COMPUTING METHODOLOGIES AND COMMUNICATION (ICCMC)*, 2017, PP. 475-479, DOI: 10.1109/ICCMC.2017.8282735.
3. N. KUMAR AND S. KHATRI, "IMPLEMENTING WEKA FOR MEDICAL DATA CLASSIFICATION AND EARLY DISEASE PREDICTION," *2017 3RD INTERNATIONAL CONFERENCE ON COMPUTATIONAL INTELLIGENCE & COMMUNICATION TECHNOLOGY (CICT)*, 2017, PP. 1-6, DOI: 10.1109/CIACT.2017.7977277.
4. A. H.J., B. NIRMAL AND A. S. MAHESH, "PROGNOSTICATION OF DIABETES USING DATA MINING MODELS," *2021 6TH INTERNATIONAL CONFERENCE ON COMMUNICATION AND ELECTRONICS SYSTEMS (ICCES)*, 2021, PP. 1883-1887, DOI: 10.1109/ICCES51350.2021.9489061.
5. R. AL-DHAIBANI, M. A. M. BAMATRAF AND K. Q. SHA'AFAL, "DATA BENCHMARK COLLECTION OF PATIENTS WITH MALARIA FOR MACHINE LEARNING: A STUDY IN HADHRAMOUT- YEMEN," *2019 FIRST INTERNATIONAL CONFERENCE OF INTELLIGENT COMPUTING AND ENGINEERING (ICOICE)*, 2019, PP. 1-3, DOI: 10.1109/ICOICE48418.2019.9035166.
6. N. RAMKUMAR, S. PRAKASH, S. A. KUMAR AND K. SANGEETHA, "PREDICTION OF LIVER CANCER USING CONDITIONAL PROBABILITY BAYES THEOREM," *2017 INTERNATIONAL CONFERENCE ON COMPUTER COMMUNICATION AND INFORMATICS (ICCCI)*, 2017, PP. 1-5, DOI: 10.1109/ICCCI.2017.8117752.
7. S. S. RAYKAR AND V. N. SHET, "COGNITIVE ANALYSIS OF DATA MINING TOOLS APPLICATION IN HEALTH CARE SERVICES," *2020 INTERNATIONAL CONFERENCE ON EMERGING TRENDS IN INFORMATION TECHNOLOGY AND ENGINEERING (IC-ETITE)*, 2020, PP. 1-7, DOI: 10.1109/IC-ETITE47903.2020.442.
8. R. SYED, R. K. GUPTA AND N. PATHIK, "AN ADVANCE TREE ADAPTIVE DATA CLASSIFICATION FOR THE DIABETES DISEASE PREDICTION," *2018 INTERNATIONAL CONFERENCE ON RECENT INNOVATIONS IN ELECTRICAL, ELECTRONICS & COMMUNICATION ENGINEERING (ICRIEECE)*, 2018, PP. 1793-1798, DOI: 10.1109/ICRIEECE44171.2018.9009180.

Biographies

Archika Jain received the bachelor's degree in information technology from YIT, Jaipur in 2013, the master's degree in computer engineering from PU, Jaipur in 2016, and pursuing PhD in Computer Engineering. Her research areas include image processing, data mining, deep learning, and machine learning. She has published many research papers in different National and International journals and conferences.

Mr. Devendra Kumar Somwanshi completed his Bachelor’s degree in EC in 2007 from Govt. Eng. College, Bikaner. He did his Master’s Degree in EC from Thapar University, Patiala in 2009 and is currently pursuing his PhD. He authored 3 technical books. He has published more than 60 research papers in National and International conferences.

The Author, Barkha Narang received Bachelor’s Degree in Computer Science Engineering from Rajasthan University in 2004, Masters Degree from Banasthali University in 2009. She is pursuing Ph. D. She has been working on Block chain. She has an experience of 16 plus years into Academics. Also, she has written many research papers.

Cognitive Radio Engine Design using PSO and Firefly Algorithm

Himanshu Sharma[1], Anuj Kumar Goel[2], Ankur Singhal[3], Nemish Dikkala[4]
[1]*Associate Professor, ECE Department, J. B. Institute of Engineering and Technology, Hyderabad, India*
[2]*Professor, ECE Department, University Institute of Engineering, Chandigarh University, Punjab, India*
[3]*Professor, Department of ECE, Chandigarh Engineering College, Landran, India.*
[4]*Student, Department of ECE, J B Institute of Engineering and Technology, Hyderabad, India*
Email: himanshu.zte@gmail.com[1], anuj40b@gmail.com[2], asking2k@gmail.com[3], nemishscience@gmail.com[4]

Abstract – The cognitive radio is the basic cell that makes up the cognitive radio networks. The CRN and its multiple architectures and applications have received extensive attention in the literature. However, a look at the previous work on CR reveals that, in the field of cognitive radio, little is published about the complexities and specificities of the development of its core, the Cognitive Radio Engine (CRE). For proposed Cognitive Radio Engine, this paper discusses two metaheuristics that are Particle Swarm Optimization (PSO) and Firefly Algorithm (FFA) as approximate methods for the optimization of fitness function. For establishing evaluation of performance of CRE according to various criteria's that have been set, like Bit Error Rate (BER), Channel attenuation and Output Power.

Key words – CR, CRE, BER, DSA, FCC, FFA, PDA, PSO, SDR, CE.

1. INTRODUCTION

The increasing communications services and the heterogeneity of the networks have generated the formation of network architectures that respond to this diversity of technologies such as Personal Digital Assistant (PDA), laptops and smartphones that are being designed to make use of wireless standards and take advantage of the services offered by this technology, as in the case of 3G technology, which offers data and voice transmission, and additionally allows video calls to be made.

Cognitive Radio Engine Framework using two metaheuristic approaches; Particle Swarm optimization and Firefly Algorithm has been discussed. Second section presents the theoretical terms related to cognitive radio network. Section three describes the methodology regarding CR engine design implementation. Results are given in fourth section followed by the conclusion in fifth section.

2. COGNITIVE RADIO ENGINE

2.1. Cognitive Radio Engine Architecture

The Cognitive Engine CE of the Cognitive Radio CR consists basically of a programmer block, a core, detection interfaces, user and network (Figure 2.1). In turn, the core consists of a database and blocks of learning, reasoning and optimization.

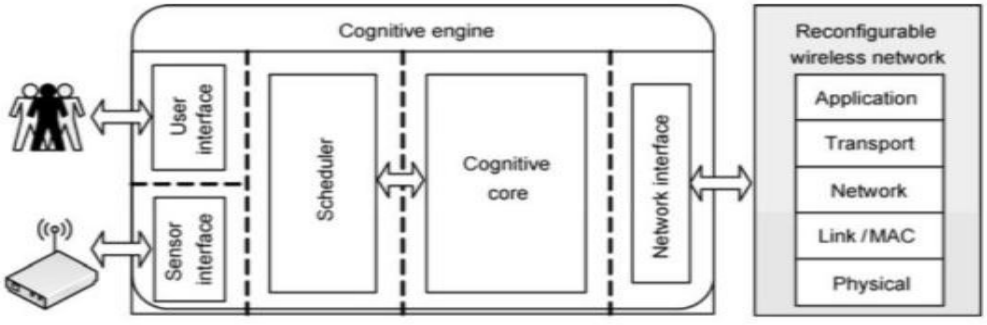


Figure 2.1: Basic block diagram of a CE [12]

3. METHODOLOGY

For its operation, a Cognitive Engine uses metaheuristics**,** so a comparative study between two metaheuristics that are Cuckoo search and Firefly has been done and an exact method called parallel dynamic programming. The main goal is to further develop spectrum management in the CRN framework. So, comparative study has been presented.

3.1. Objective Function (Fitness Function)

This function is to find the optimum result to a combinatorial optimization issue. In this case, to maximise cost is the utility of this function."

Following are the Parameters:

- n: SU numbers.
- m: free channels numbers possessed by PU.
- W: an array of size n. $W[i]$ is the channels numbers requested by SUi.
- C: a size chart n.
- Optimized function is: Max $\sum_{i=0}^{n-1} C[i]$
- The constraint to be respected is: $\sum_{i=0}^{n-1} W[i] \leq m$.

3.2. Particle Swarm Optimization

The PSO depends on a bunch of individuals initially arbitrarily and homogeneously organized, which we call particles from this point forward, which move in the research hyper-space and comprise, every one, a likely arrangement. Every molecule has a memory of its best arrangement visited and the capacity to speak with the particles that encompass it. From this data, the molecule will follow an inclination made, from one perspective, of its will to return towards its ideal arrangement, and then again, of its mimicry contrasted with the arrangements found in its area.

From nearby and observational optima, the arrangement of particles will ordinarily unite to the ideal in general arrangement of the issue being tended to.

Particle swarm is portrayed by [10]:

- The quantity of particles of the swarm is nb.
- The greatest velocity of a particle is $\vec{v}_{max}$.
- Particle inertia is given by Ψ.
- Confidence coefficients, given by ρ_1 and ρ_2,
- Particle is described at time t by:
- $\vec{x}_i(t)$: search space position.
- $\vec{v}_i(t)$: velocity
- $\vec{x}_{pbest_i}(t)$: the place of the best arrangement by which it has passed.
- $\vec{x}_{vbest_i}(t)$: the place of the most popular arrangement of its neighborhood.
- $pbest_i$: best solution fitness value.

3.3. Firefly Algorithm

The latest metaheuristic is Firefly Algorithm. It was created by Yang [13] [14].

Principle of Operation

Given below are the four important points in the firefly algorithm:

Intensity of Light: For simplest case for minimization problems, the luminosity of a firefly at specific location x to chosen as: $I(x)\ \alpha\ 1\ /\ f\ (x)$.

Attractiveness: The core form of this function is represented by any decreasing monotonic function like given:

$$\beta_{i,j} = \beta_0^* e^{-\gamma r_{i,j}^m} \tag{1}$$

Where the distance between two fireflies is r, β_0 at $r = 0$ is attractiveness and γ is constant coefficient for absorption of light.

Distance: The range between two fireflies i and j at x_i and x_j will be the Cartesian distance given:

$$r_{i,j} = \sqrt{\sum_{k=1}^{d} \left(x_{i,k} - x_{j,k}\right)^2} \tag{2}$$

Where $x_{i,k}$ will be the k^{th} component of i^{th} firefly.

Movement: Displacement of a firefly i attracted by more luminous firefly j, is given by:

$$x_i = \left(1 - \beta_{i,j}\right)x_i + \beta_{i,j}x_j + \alpha\left(rand - \frac{1}{2}\right) \tag{3}$$

Where the attraction refers as the first term and second term. Randomization is the third term. Random parameter is α and can be constant. "$rand$" is a random number generator distributed uniformly among [0, 1].

4. RESULTS OF SIMULATION

Graphs beneath address the results obtained:

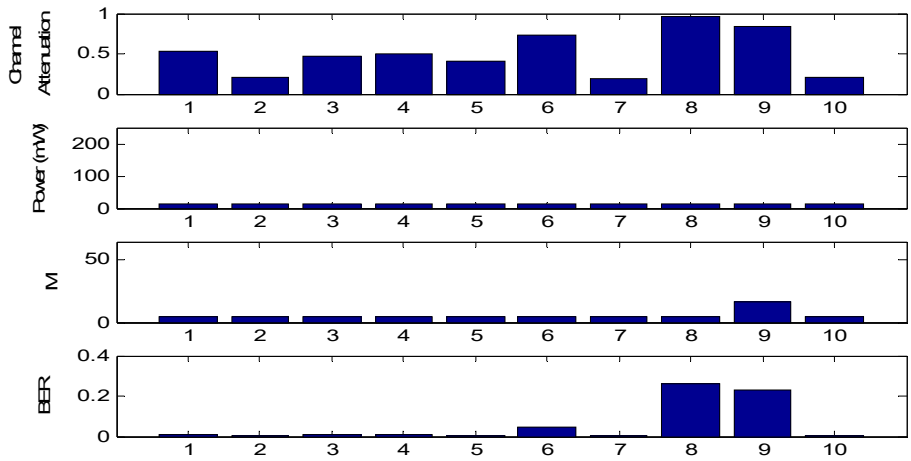


Figure 4.1. Result of low power mode using firefly algorithm

Figure 4.1 shows output of BER, modulation type, power, and channel attenuation for low power mode using firefly algorithm. It is found that the power consumption is minimized in low power mode while the BER is slightly high.

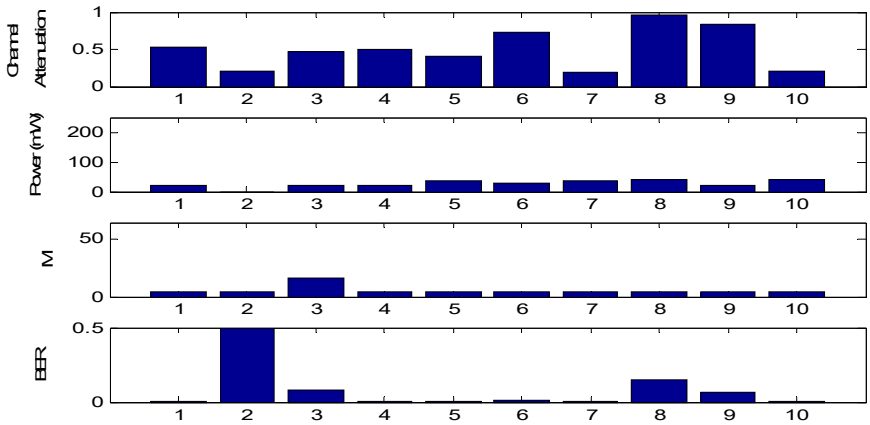


Figure 4.2. Result of emergency mode using firefly algorithm

Figure 4.2 shows output of BER, modulation type, power, and channel attenuation for emergency mode using firefly algorithm. This mode should minimize the BER for each channel. It was noticed that each channel BER was very low driven by the CE offered by the higher attenuation on each channel.

Figure 4.3 shows output of BER, modulation type, power, and channel attenuation for multimedia mode using firefly algorithm. It is under multimedia mode, therefore it is observed that the throughput is maximized at the cost of large power transmitted by proposed cognitive engine.

Figure 4.4 shows output of BER, modulation type, power, and channel attenuation for balanced mode using firefly algorithm. In balanced mode, weights of the three performance evaluation metrics are the same while BER and transmit power are relatively more.

Figure 4.5 shows a comparative analysis of BER, modulation type, power, and channel attenuation between PSO and firefly algorithm for low power mode. The power consumption is minimized while the BER is slightly high during low power mode. Here the firefly algorithm outperforms the PSO algorithm.

Figure 4.6 shows a comparative analysis of BER, modulation type, power, and channel attenuation between PSO and firefly algorithm for emergency mode. This mode should minimize the BER for each channel. It was noticed that for the firefly algorithm, the each channel BER was very low driven by the CE offered by the higher attenuation on each channel.

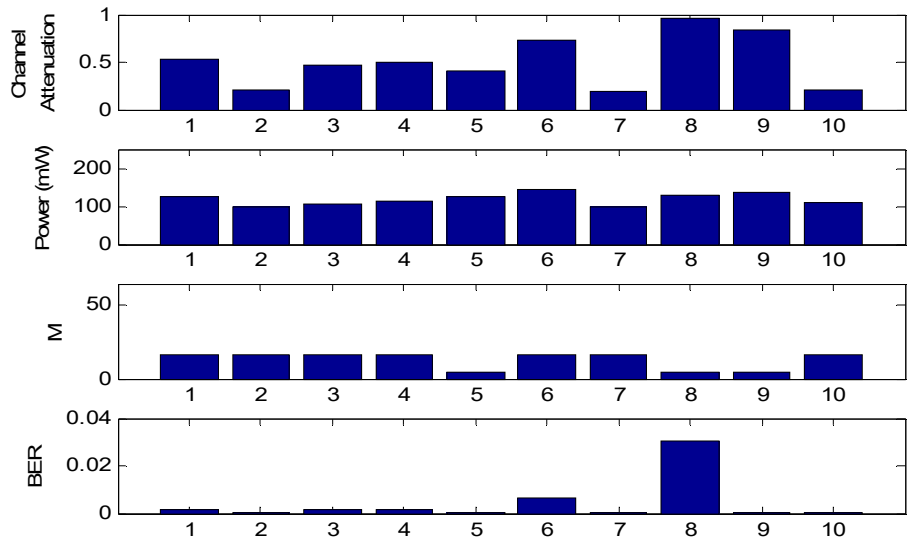


Figure 4.3. Result of multimedia mode using firefly algorithm

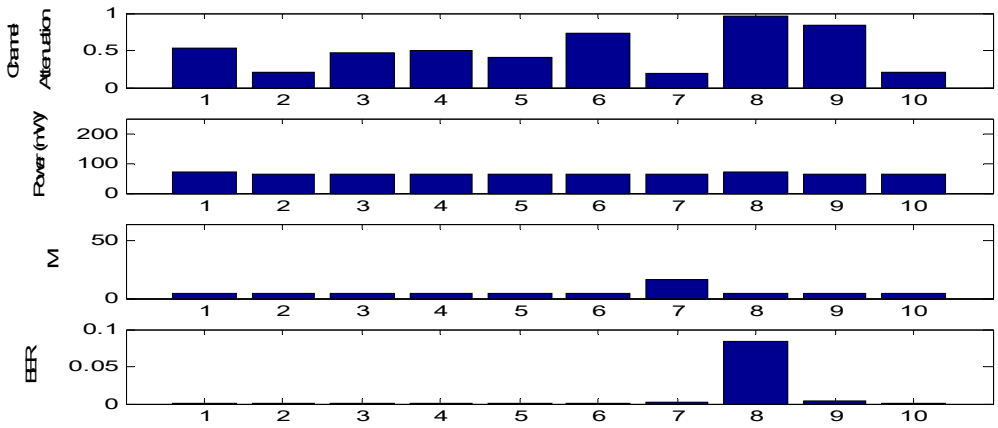


Figure 4.4. Result of balanced mode using firefly algorithm

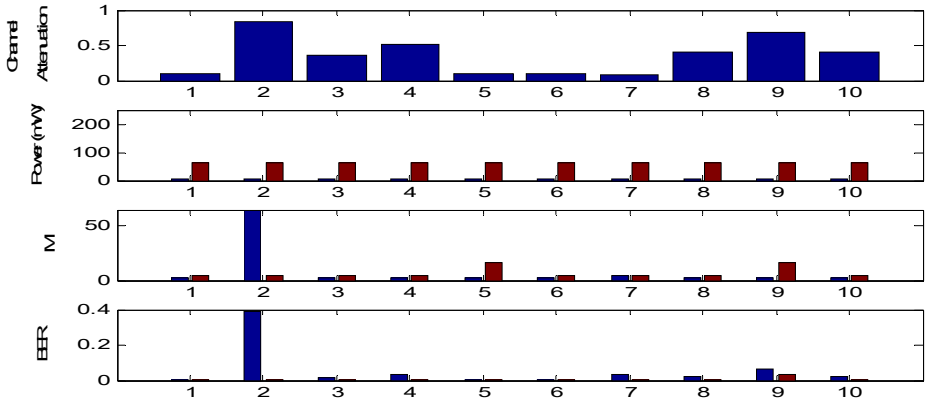


Figure 4.5. Comparative result of PSO and firefly algorithm in low power mode

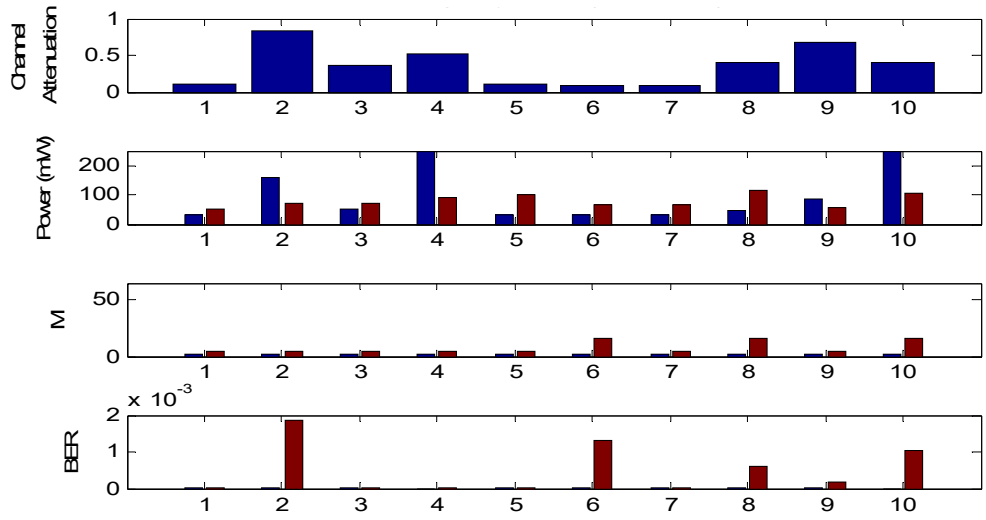


Figure 4.6. Comparative result of PSO and firefly algorithm in emergency mode

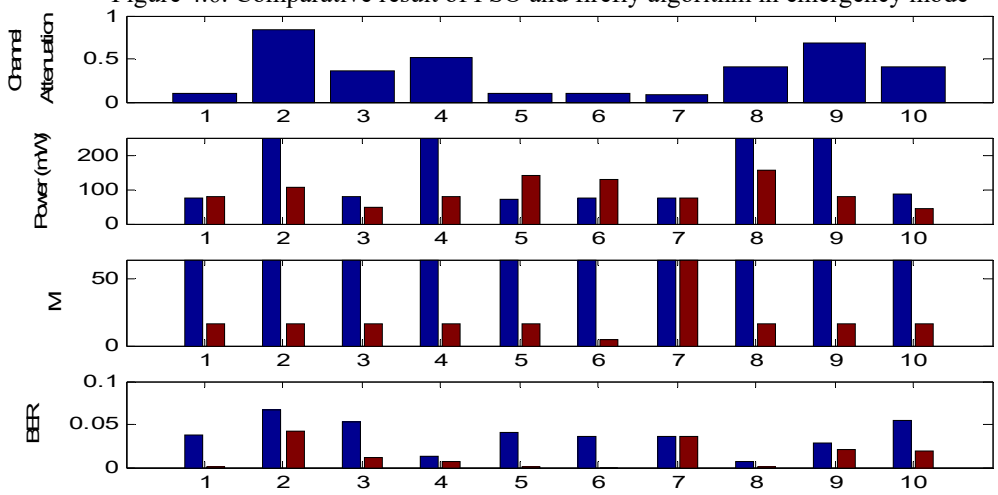


Figure 4.7. Comparative result of PSO and firefly algorithm in multimedia mode

Figure 4.7 shows a comparative analysis of BER, modulation type, power, and channel attenuation between PSO and firefly algorithm for multimedia mode. It is under multimedia mode, therefore it is observed that the throughput is maximized at the cost of large power transmitted by firefly algorithm based cognitive engine. Therefore, the firefly algorithm outperforms the PSO based approach.

5. CONCLUSION

Cognitive Radio is significant methodology which take care of access clashes and saturation of spectrum issues. The trades made in a range access arrangement among primary clients and secondary clients are seen as a sensitive activity and furthermore as conjugative optimization problem. Therefore in this review, we concentrated on a result that will nevertheless provides good pursuance to this negotiation, our work is based on metaheuristics that presents quick results. Two metaheuristic PSO and firefly algorithms that represent the optimization methods are presented. Simulation results clearly shows that the firefly algorithm provides better results as compared to PSO optimization.

REFERENCES

[1] M.A. McHenry, “NSF Spectrum Occupancy Measurements Project Summary”, *Shared spectrum co. report*, Vienna, USA Aug. 2005.

[2] Erpek, T., K. Steadman, and D. Jones. "Dublin Ireland spectrum occupancy measurements project summary." *Shared spectrum co. report, Vienna, USA* (2007).

[3] Federal Communications Commission. *FCC, ET Docket No 03-222 Notice of proposed rule making and order*. Tech. Rep., December, 2003.

[4] Akyildiz, Ian F., Won-Yeol Lee, Mehmet C. Vuran, and Shantidev Mohanty. "A survey on spectrum management in cognitive radio networks." IEEE Communications Magazine (2008).
[5] Setoodeh, Peyman, and Simon Haykin. "Cognitive Radio Networks." (2017).
[6] Alias, Dinu Mary. "Cognitive radio networks: A survey." In *2016 International conference on wireless communications, signal processing and networking (WiSPNET)*, pp. 1981-1986. IEEE, 2016.
[7] Venkataraman, Hrishikesh. *Cognitive radio and its application for next generation cellular and wireless networks*. Edited by Gabriel-Miro Muntean. Berlin: Springer, 2012.
[8] Mitola, Joseph, and Gerald Q. Maguire. "Cognitive radio: making software radios more personal." *IEEE personal communications* 6, no. 4 (1999): 13-18.
[9] Sahukar, Latha, and L. Madhavi. "Frequency domain based digital down conversion architecture for software defined radio and cognitive radio." *Int. J. Eng. Technol* 7, no. 216 (2018): 88.
[10] Khan, Hamza, and Sang-Jo Yoo. "Multi-Objective Optimal Resource Allocation Using Particle Swarm Optimization in Cognitive Radio." In *2018 IEEE Seventh International Conference on Communications and Electronics (ICCE)*, pp. 44-48. IEEE, 2018.
[11] Karmakar, Nemai Chandra, ed. *Handbook of smart antennas for RFID systems*. John Wiley & Sons, 2011.
[12] Rieser, Christian J., Thomas W. Rondeau, Charles W. Bostian, and Timothy M. Gallagher. "Cognitive radio testbed: further details and testing of a distributed genetic algorithm based cognitive engine for programmable radios." In *IEEE MILCOM 2004. Military Communications Conference, 2004.*, vol. 3, pp. 1437-1443. IEEE, 2004.
[13] Yang, X. S. "Nature-Inspired Metaheuristic Algorithms United Kingdom." (2011): 148.
[14] Yang, Xin-She. "Firefly algorithm, Levy flights and global optimization." In *Research and development in intelligent systems XXVI*, pp. 209-218. Springer, London, 2010.
[15] Au, Kelvin Kar Kin, and Jianglei Ma. "Systems and Methods for Software Configurable Air Interface Adaptation." U.S. Patent Application 14/941,380, filed May 19, 2016.

Biography

Himanshu Sharma received his B.Tech in ECE from Kurukshetra University, Kurukshetra, M.E in ECE from Thapar Institute of Engineering and Technology, Patiala, Punjab and Ph.D. from Maharishi Markandeshwar (Deemed to be University), Mullana. He has more than 12 years of teaching experience. Currently he is serving as Associate Professor in Department of ECE at J. B. Institute of Engineering & Technology, Hyderabad. He has published more than 45 research papers in various reputed conferences and journals and supervised more than 20 M.Tech. Thesis. His area of research is Wireless Communication and Networks and currently he is working on Security mechanisms for Cognitive Radio Networks.

COMPLEX DYNAMICS OF FIRST SUPERIOR BARNSLEY FRACTAL

Sunil Shukla

Department of Computer Applications, UIM, Uttaranchal University, Dehradun.

mailtosunilshukla@gmail.com

Abstract.

Superior fractals have expanded the range of fractal applications, such as chaotic cryptography, in the creation of spectacular and lifelike computer visuals, fractal antenna, particle dynamics, social research, in the design of file compression systems, the construction of the internet's networks, and even in the diagnosis of some diseases. Indeed, this modest start has revealed a whole new dimension of fractal theory and its applications. This paper introduces new fractals using Barnsley functions and superior iterates. Barnsley has introduced three fractal functions to generate new escape time fractals which gives a new approach to the study of IFS (Iterated function systems). In addition to, this study also represents the complex dynamics of First Barnsley Fractals using superior iterates. In the existing literature, Julia sets and their generalizations have been developed using one-step feedback process (Picard iteration). In this paper two step feedback process (Mann iterates and superior iterates) has been introduced to study of Julia sets and obtained superior Julia sets.

Key words: Complex dynamics, Barnsley Fractals, Picard iteration, Mann iterates and superior iterates

1. INTRODUCTION

Michal F. Barnsley has given a new approach to fractal by using iterative function system [1-2] and [4]. In his book “Fractals Everywhere” [2] author has presented the concept of Barnsley fractal of type First, Second and Third. The images of Barnsley fractals are different from the classical fractals. This paper has presented a new approach to the Barnsley Fractals using superior iterates [5], [6], [7] and [8], named as Superior Barnsley Fractals. Further this paper has discussed the first Barnsley fractal and their behavior using superior iterates.

The two-step feedback method (Mann iterates and superior iterates) was used to examine Julia sets in this paper, and superior Julia sets were obtained. The IFS patterns, curves, spokes, and spirals symmetric along the y axis can be seen by zooming in on the superior Barnsley fracrtal.

2. SUPERIOR ORBIT

Let A be a subset of real or complex numbers and $f : A \to A$. For $x_0 \in A$, construct a sequence $\{x_n\}$ in A in the following manner

$$x_1 = s_1 f(x_0) + (1 - s_1) x_0$$
$$x_2 = s_2 f(x_1) + (1 - s_2) x_1$$
$$\vdots$$
$$x_n = s_n f(x_{n-1}) + (1 - s_n) x_{n-1}$$

where $0 < s_n \leq 1$ and $\{s_n\}$ is convergent to a non-zero number.

The above mention sequence $\{x_n\}$ is known as the Mann sequence of iterates or superior sequence of iterates. Let z_0 be an arbitrarily element of C that can be chosen at random. Construct a $\{z_n\}$ sequence of points C in the following manner:

$$z_n = sf(z_{n-1}) + (1 - s) z_{n-1,} n = 1, 2, 3...,$$

where f is a function over a subset of C and s is closed interval parameter lies in $[0,1]$. The above mention sequence $\{z_n\}$, represented by $SO(f, z_0, s)$, is the superior orbit for the complex-valued function f with an initial choice z_0 and parameter s. This may be symbolized by $SO(f, x_0, s_n)$. Notice that $SO(f, x_0, s_n)$ with $s_n = 1$ is $O(f, x_0)$. When $s_n = 1$, the superior orbit decreases to the normal Picard orbit, according to this observation.

The Barnsley Fractal functions are defined as:

(A) First Barnsley Fractal:

$z_{n+1} = c\ z_n - c$ ***where real (z) ≥ 0***

$= \boldsymbol{c\ z_n + c}$ otherwise

(B) Second Barnsley Fractal:

$z_{n+1} = c\ z_n - c$ where imag(***z***) ≥ 0

$= c\ z_n + c$ otherwise

(C) Third Barnsley Fractal:

Here the value of c is (1.0, 0)

$z_{n+1} = z_n^2 - c$ if real(***z***) ≥ 0

$= z_n^2 - c + px$ otherwise

here ***px***= real(*c*) * real (z_n)

3. ESCAPE CRITERION FOR BARNSLEY FRACTALS

After looking to the fractal formulas it has been observed that all the three Barnsley fractal formulas are of the form $\boldsymbol{z}^2+\boldsymbol{c}$, hence the escape criterion is defined as follows:

Let $B_c(\boldsymbol{z})$ be the Fractal Formula, then

$$
\begin{aligned}
&| B_c(\boldsymbol{z})| = |(\boldsymbol{c}\, \boldsymbol{z}_n + \boldsymbol{c}) | \\
&\geq |\boldsymbol{c}|\, |\boldsymbol{z}| - |\boldsymbol{c}| \qquad \text{since} |\boldsymbol{z}| \geq |\boldsymbol{c}| \\
&\geq |\boldsymbol{c}|\, \{|\boldsymbol{z}| - 1\} \\
&> |\boldsymbol{z}|\, (|\boldsymbol{z}| - 1) \\
&\text{since } |\boldsymbol{z}| > 2, \text{ there is} \\
&> (1 + \lambda^n)\, |\boldsymbol{z}|
\end{aligned}
$$

It has been found that the escape criterion is the same as of quadratic polynomial.Raniand Kumar [6] has developed the phenomenon of Superior orbits for the Madelbrot and Julia Sets and named as Superior Mandelbrot and superior Julia Set respectively using Mann Iterates. After applying the same phenomenon it can be redefine three Barnsley Fractals named as Superior Barnsley Fractal. Fractal can be considered with the fusion technique [9-14] to improve the visualization in such manner.

4. SUPERIOR BARNSLEY FRACTALS

After applying the superior iterates the three Barnsley Fractal Functions (A), (B) and (C) can be rewritten as:

$$
\begin{array}{lll}
\text{(A1)} \quad \boldsymbol{z}_{n+1} & = \boldsymbol{s}(\boldsymbol{c}\boldsymbol{z}_n - \boldsymbol{c}) + (1 - \boldsymbol{s})\, \boldsymbol{z}_n & \text{where real } (\boldsymbol{z}) \geq 0 \\
& = \boldsymbol{s}(\boldsymbol{c}\boldsymbol{z}_n + \boldsymbol{c}) + (1 - \boldsymbol{s})\, \boldsymbol{z}_n & \text{otherwise} \\
\text{(B1)} \quad z_{n+1} & = \boldsymbol{s}(\boldsymbol{c}\boldsymbol{z}_n - \boldsymbol{c}) + (1 - \boldsymbol{s})\, \boldsymbol{z}_n & \text{where imag } (\boldsymbol{z}) \geq 0 \\
& = \boldsymbol{s}(\boldsymbol{c}\boldsymbol{z}_n + \boldsymbol{c}) + (1 - \boldsymbol{s})\, \boldsymbol{z}_n & \text{otherwise} \\
\text{(C1)} \quad z_{n+1} & = \boldsymbol{s}(\boldsymbol{z}_n^2 - \boldsymbol{c}) + (1 - s)\, \boldsymbol{z} & \text{where real } (\boldsymbol{z}) \geq 0 \\
& = \boldsymbol{s}(\boldsymbol{z}_n^2 - \boldsymbol{c} + \boldsymbol{p}\boldsymbol{x}) + (1 - \boldsymbol{s})\, \boldsymbol{z}_n & \text{otherwise} \\
& \text{where } px = \text{real} (\boldsymbol{c}) * \text{imag} (\boldsymbol{z}_n) &
\end{array}
$$

This paper analysis the First Barnsley Fractal.

5. ANALYSIS OF FIRST SUPERIOR BARNSLEY FRACTAL

It has been observed that the first superior Barnsley fractal as a sequence of beautiful spiral for $\boldsymbol{f}\,(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{s})$ = (1.025, -0.093, 1.0) see Figure. 1. For $\boldsymbol{f}\,(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{s})$= (1.263, -0.989, 0.1) and $\boldsymbol{f}\,(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{s})$= (1.025, 1.229, 0.5) the symmetry has been observed around $\boldsymbol{y}$ axis see Figure. 1, 3 and 5. It has been observed the chaotic nature of points for first superior Barnsley fractal see Figure. 2, 4 and 6 where the orbit value is changes periodically rather than conversing to any fixed point.

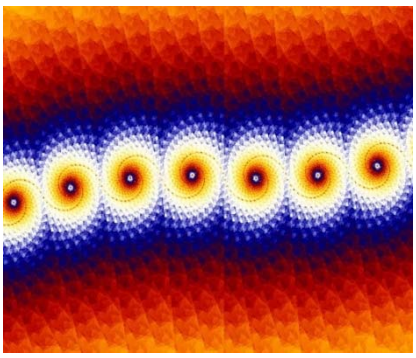


Figure. 1. Superior Barnsley Fractal for *f (a, b, s)*= (1.025, -0.093, 1.0)

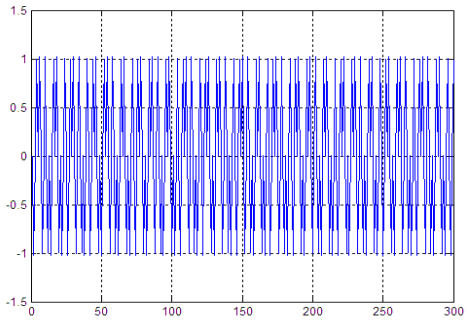


Figure. 2. Orbit of Superior Barnsley Fractal for *f* (*a, b, s*)= (1.025, -0.093, 1.0)

Table 1. For *f (a, b, s)*= (1.025, -0.093, 1.0)

Number of iteration i	| F(z) |
288	0.024404
289	0.99999
290	1.48E-05
291	1.025
292	0.025609
293	0.99875
294	0.001281
295	1.0237
296	0.024279
297	1.0001
298	0.00011668
299	1.0249

(Intentionally, several intermediary iterations have been skipped)

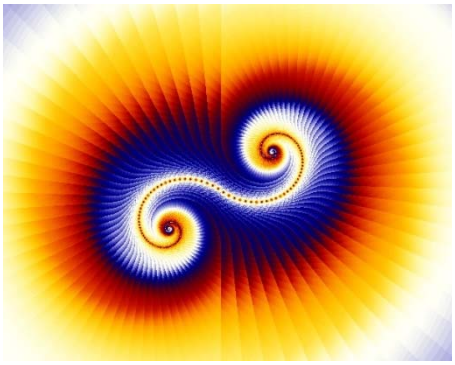


Figure. 3. Superior Barnsley Fractal for *f (a, b, s)*= (1.263, -0.989, 0.1)

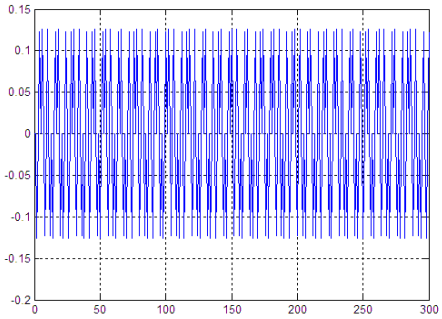


Figure. 4. Orbit of Superior Barnsley Fractal for *f (a, b, s)*= (1.263, -0.989, 0.1)

Table 2. For *f (a, b, s)*= (1.263, -0.989, 0.1)

Number of iteration i	| F(z) |
288	0.003153
289	0.12306
290	6.78E-07
291	0.1263
292	0.003321
293	0.12289
294	0.00017627
295	0.12612
296	0.003136
297	0.12308
298	1.85E-05
299	0.12628

(Intentionally, several intermediary iterations have been skipped)

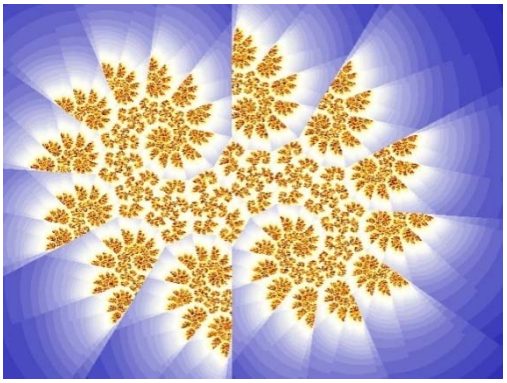


Figure. 5. Superior Barnsley Fractal for *f (a, b, s)*= (1.025, 1.229, 0.5)

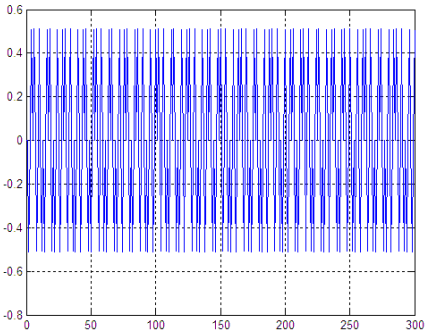


Figure. 6. Orbit of Superior Barnsley Fractal for *f (a, b, s)*= (1.025, 1.229, 0.5)

Table 3. For *f (a, b, s)*= (1.025, 1.229, 0.5)

Number of iteration i	| F(z) |
288	0.006249
289	0.50617
290	2.89E-09
291	0.5125
292	0.0064062
293	0.50601
294	0.00016115
295	0.51234
296	0.006241
297	0.50618
298	8.21E-06
299	0.51249

(Intentionally, several intermediary iterations have been skipped)

6. CONCLUSION

This paper has been presented the three Barsnley Fractals using superior iterates. The results of the Barnsley fractals and superior Barnsley fractals are illustrated in Figure. 1 to 6. On zooming the superior Barnsleyfracrtal we find the IFS patterns, ***curves***, ***spokes*** and ***spirals*** symmetric along ***y*** axis. Further by using the Mann iterates we are able to get the superior Barnsley fractals for the values of ***c*** more than –2 to +2. The work in this study can be expanded to find uses of Superior Orbits in optimising the noise and perturbations of the Mandelbrot set, as well as some novel applications and analysis using Fractal theory in fields such as medicine, engineering, and the arts.

7. REFERENCES

[1] Barcellos, A. and Barnsley, Michael F., 1990. Reviews: Fractals Everywhere. Amer. Math. Monthly , No. 3, pp. 266-268.

[2] Barnsley, Michael F., 1993.Fractals Everywhere.Academic Press, INC, New York.

[3] Becker, K.H. and Dorfler, M., 1990.Dynamical Systems and Fractals. Cambridge University Press, New York, USA.

[4] Demkom, S., Hodges, L. and Naylor, B., 1985. Construction of fractal objects with iteratedfunction systems. Proceedings of the 12th annual conference on Computer graphics and interactive techniques, pp. 271 – 278.

[5] H, Akhlaq. N, Nanda. C, MovvaSitaram. and S., Mohammad. 2022. Fractals: An Eclectic Survey, Part-I. Fractal and Fractional, 6, 89.

[6] Kumar, Manish.and Rani, Mamta., 2005. A new approach to superior Julia sets. J. nature. Phys. Sci, pp. 148-155.

[7] Negi, A., 2006. Fractal Generation and Applications, Ph.D Thesis, Department of Mathematics, GurukulaKangriVishwavidyalaya, Hardwar, 2006.

[8] Rani, M., 2002.Iterative Procedures in Fractal and Chaos.Ph.D Thesis, Department of Computer Science. GurukulaKangriVishwavidyalaya, Hardwar, 2002.

[9] Sharma, S., Kumar, T., Dhaundiyal, R., Mishra, A. K., Duklan, N., & Maithani, A. (2019). Improved method for image security based on chaotic-shuffle and chaotic-diffusion algorithms. International Journal of Electrical and Computer Engineering, 9(1), 273–280. https://doi.org/10.11591/ijece.v9i1.pp.273-280

[10] Singh, P., Shree, R., & Diwakar, M. (2021). A new SAR image despeckling using correlation based fusion and method noise thresholding. Journal of King Saud University - Computer and Information Sciences, 33(3), 313–328. https://doi.org/10.1016/j.jksuci.2018.03.009

[11] Diwakar, M., & Kumar, M. (2018a). A review on CT image noise and its denoising. Biomedical Signal Processing and Control, 42, 73–88. https://doi.org/10.1016/j.bspc.2018.01.010

[12] Joshi, K., Joshi, N. K., Diwakar, M., Tripathi, A. N., & Gupta, H. (2019). Multi-focus image fusion using non-local mean filtering and stationary wavelet transform. International Journal of Innovative Technology and Exploring Engineering, 9(1), 344–350. https://doi.org/10.35940/ijitee.A4123.119119

[13] Kumar, A., Singh, D., & Punia, P. (2016). Implementation of image dehazing technique using image fusion. International Journal of Control Theory and

Applications, 9(20), 307–315. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85006371817&partnerID=40&md5=17ff92c722005ff83c9c27c371a0f2ed

[14] Joshi, K., Diwakar, M., Joshi, N. K., & Lamba, S. (2021). A concise review on latest methods of image fusion. Recent Advances in Computer Science and Communications, 14(7), 2046–2056. https://doi.org/10.2174/2213275912666200214113414

Biographies

Dr. Sunil Shukla is having a teaching and research experience of approximately 15 Years. He has earned his Ph.D in 2013 from Uttarakhand Technical University, Dehradun. He has supervised M. Phil scholars at “DeshBhagat University”, MandiGobindgarh (Punjab). Many of his papers have been published in national and international journals. He has also attended a number of conferences, workshops and Faculty Development Programs.

Simulation and Performance Analysis of PDSCH in 5G NR

Santosh M Nejakar[1], Suprith P G[2], Mohammed Riyaz Ahmed[3] , Santosh Herur[4]

[1]*Associate Professor, Dept. of Electrical and ElecronicsEgineering, Jain Institute of Technology, Davangere, Email Id: nejakarganu@gmail.com*

[2]*Research Scholar, School Electronics and Communication Egineering, Reva University, Bangalore, India Email Id: pgsuprith@gmail.com*

[3]*Associate Professor, School Electronics and Communication Egineering, Reva University, Bangalore, India Email Id:riyaz@reva.edu.in*

[4]*Associate Professor, Dept. of Electrical and ElecronicsEgineering, Jain Institute of Technology, Davangere, Emai Id:santoshmherur@gmail.com*

Abstract

5G is the new upcoming digital wireless communication network, which has a high-speed data transmission cellular network also called New Radio (NR). It is industry associated 3rd generation partnership project (3GPP) that defines any system using "5G NR". Beyond the speed, it will allow many new mobile capabilities to be realized-high data capacity, low latency, and IoT capability. The significant exploration issue for 5G is to plan productive and reliable Physical Downlink Shared Channels (PDSCH). In this paper, the design, simulation, and its results are discussed, based on different parameters like subcarrier spacing and modulation schemes such as 64QAM and 256 QAM techniques using MATLAB. The overall discussion and its conclusion will benefit to development of a better 5G communication system.

Keywords: Multiple-Input-Multiple-Output (MIMO), Physical Downlink Shared Channel (PUSCH), Physical Downlink Shared Channels (PDSCH), New Radio, Paging Channel (PCH)

1. INTRODUCTION

Digital 5G cellular network uses millimeter waves, which provide a short range of communication. The antenna used for communication is of smaller size (Several Centimeters). To enhance the data rate, Multiple-Input-Multiple-Output (MIMO) concept is used; every cell comprises of a few recieving wires speaking with the remote gadget, consequently numerous bitstreams of information will be sent at the same time in equal.The 5G systems consist of mainly three channels such as transport, logical, and physical channels, these interns are subdivided into downlink and uplink channels. The downlink consists of Paging Channel (PCH), Broadcast Channel (BCH), Downlink Shared Channel (DL-SCH), Physical layer Downlink Shared Channel (PDSCH), Physical Downlink Control Channel (PDCCH). Uplink channels consist of Uplink layer Shared Channel (UL-SCH), Physical Downlink Shared Channel (PUSCH), Physical Downlink Control Channel (PUCCH)[1]. 5G wireless mobile network employments the orthogonal frequency division multiple access (OFDMA) technique, which is one of the popular modulation schemes in digital modulation and has capable of converting the wide-band recurrence specific channel into a bunch of numerous fading sub-channels[2]. These channels have the capability of acceptance from optimum receivers, indeed, even on account of MIMO transmission with sensible complications[3]. PDSCH has most of the utility to support numerous MIMO transmissions, Hybrid Automatic Repeat Request (HARQ), MAC layer scheduling, and many additional functions. This paper describes and analyzes the 5G physical channel (Downlink) based on the 3GPP criterion. The proposition carries out a standard-consistent of 5G downlink with its significant depiction being Error-checking, MIMO transmission, and Adaptive Modulation and Coding (AMC). It will in general be used to process the presentation of PDSCH and bear the expense of the important reference for the useful arrangement of the 5G system.

2. THEPHYSICAL DOWNLINK SHARED CONTROL CHANNEL DESIGN

The physical channel (PDSCH) holds encoded client data and paging information to that of User Equipment (UE) on a dynamic as well as an opportunistic basis. In fig. 1, it shows how the transmission block is passed to DL-SCH, which gives an output of 1 or 2 codewords, and information of every codeword is encrypted as well as regulated to create a lump of complex-esteemed difference representation. The portrayals are moved toward upto 4 MIMO layers. A PDSCH can have two codewords to help up to 8-layer transmission[5].The layers are associated with radio wire ports in a prerequisite straightforward way, consequently, the bar shaping or MIMO precoding process is conveyed to organize execution and straightforward to the UE. For each radio wire port (layers) utilized for the broadcast of the PDSCH, The RBs are allocated with images[6].

A. Downlink Shared Channel (DL-SCH)

DL-SCH channel holds client data as well as other pieces of information such as different types of System Information Block (SIB). The coding chain consists of Rate Matching, Code block (CB) Segmentation, cyclic redundancy code (CRC), LTEC, and CB. In the above fig. 2, the Transport block is forwarded to CRC which first appends error detection code, followed by section the transport block becomes codeblocks further the code blocks into CRC attachment achieved. Each code block is freely a Low-Thickness Equality Actually take a look at Code (LTEC). LTEC is a channel encoding technique that clears errors of the channel by defining parity bits for a selection of the data bits. Later, LPDC coded blocks are independently rate matched [8].

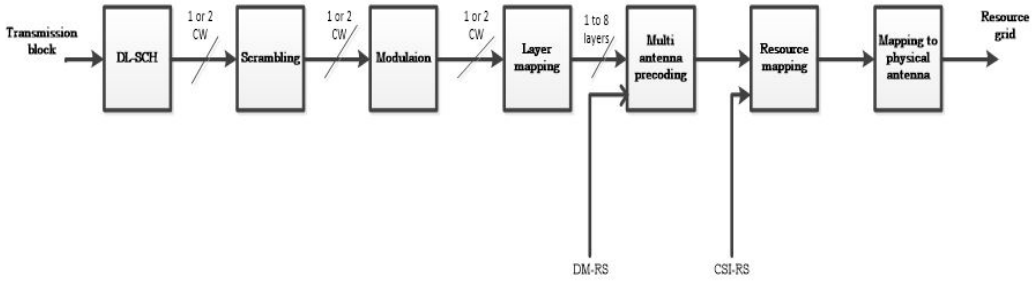


Figure1. Processing physical layer chain of PDSCH

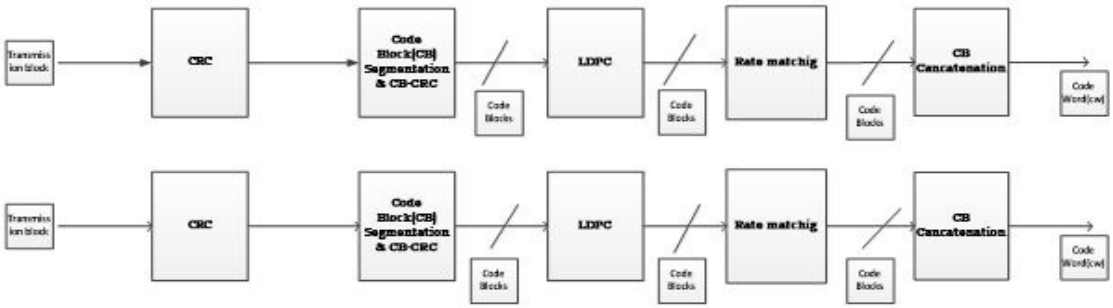


Figure2. Physical layer processing chain of downlink shared channel

The fundamental reason for the rate matching module is to match the number of pieces that can be communicated in the manage the cost of assignment to the number of pieces in the transport square and it likewise draw in sub-block interleaving, bit assortment, and pruning[9]. Ultimately, code block concatenation is done to produce a codeword, and a maximum of 2 codewords can be broadcasted at the same time on the PDSCH. Actual channels relate to a bunch of time-recurrence assets utilized for broadcasting specific vehicle channel information, control, or marker data.Each transport channel data and control, indicator information is mapped to its respective physical channel and it is further classified as PBCH, PDSCH, and PDCCH.

A. PDSCHScramblingandModulation

A maximum of two codewords can be broadcasted in a subframe and for each codeword; the data bits are scrambled with a various variety of scrambling patterns. The scrambling design is introduced toward the beginning of each subframe and relies upon RNTI, NCellID, NSubframe, and the codeword record. The regulation program form structures PDSCH are PDSCH are 'QPSK', '16QAM', '64QAM', or '256QAM', and these specify the modulation category of the codeword and the number of bits used per modulation symbol.

B. PDSCHsurface draw

The multi-layered changed images are furthermore intended to single or many layers as per the transmission rules used.The ports used for the single-layer are generally 0, 5, 7, or 8 and usually, for the broadcast diversity, the only single codeword is permitted. For the most part, the number of layers such as 2 or 4 should be equivalent to various antenna ports that are utilized for the broadcast of the actual channel[10]. For spatial multiplexing, codeword (1or2) can be broadcasted on up to 8 layers. The number of communication antenna ports larger than or equivalent to a number of layers to be used for the broadcast of the physical channel.

C. PDSCH Multi-Antenna Precoding and Resource Mapping

To map more layers, the precoding technique is accomplished. Using matrix multiplication with the pre-corner the numerous layers go through precoding antenna ports [13-14]. The exceptional instance of precoding is planning one layer to numerous antenna which empowers beamforming for line-of-sight,

transmission this would probably mean focusing on specific one more instance of precoding is planning a few layers to various receiving wires this more broad case is here and there alluded to as spatial multiplexing one critical part of precoding in 5G is that the related demodulation reference transmissions or DM-RS should go through the equivalent precoding thus the UE shouldn't be made mindful of precoder as the impact of a precoder is remembered for channel assessment this is the reason the specific precoder the G node B is to utilize isn't determined in the standard the pre-code indicated in the standard the pre-coder out is additionally connected to actual asset impedes either straightforwardly or by implication [15-16]. PDSCH at first planned to Virtual Asset Blocks (VAB) one planned to lattice PDSH images stay away from areas saved for different purposes, this incorporates all actual sign DMRS, CSI-RS. Planning of VAB to actual asset blocks are interleaved or Non-interleaved planning. [17] Non-Interleaved Mapping comprises straightforwardly planning each virtual square to a similar situation in the actual asset network. Interleaving planning gives recurrence variety by dispersing virtual squares over the entire data transfer capacity part [18-19]. The Interleaving granularity is 2 or 4 assets obstructs this plan how about we relegate back-to-back virtual asset square to a PDSCH design that is not difficult to flag just with an asset square and number of assets block while as yet getting recurrence variety. In this paper, we measure the PDSCH throughput execution boundary of a 5G NR and understand the vehicle and actual correspondence channels [20].

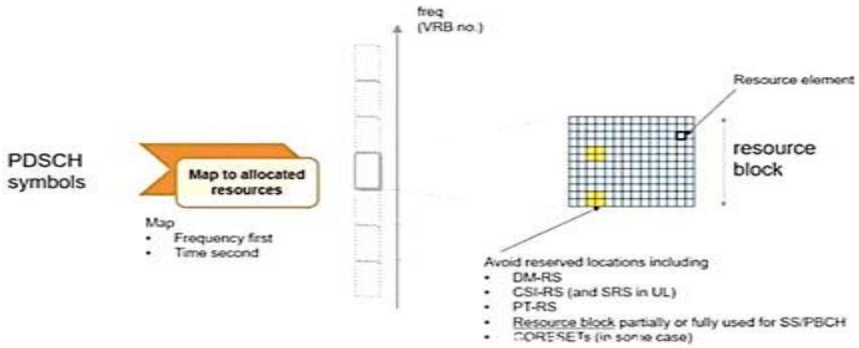


Figure3.Mapping toVRB

TableI.Parameter settings for different subcarrier Spacing/mapping techniques in the simulation

ParametersUsed	***Values***
Signal to Noise Ratio(dB)Range	-5,0,5
Number of 10ms Frames	2
Modulation Technique	64QAMand 256QAM
Number of UE Receive Antenna	2
Subcarrier dispersing in terms of KHz	15,30,60,120&240
Number of PDSCH Layers	2
PDSCH Transmission Antennas	8
Number of Codewords	2
PDSCH symbol allocation in each slot	0:13
Propagation Channel Type	CDL

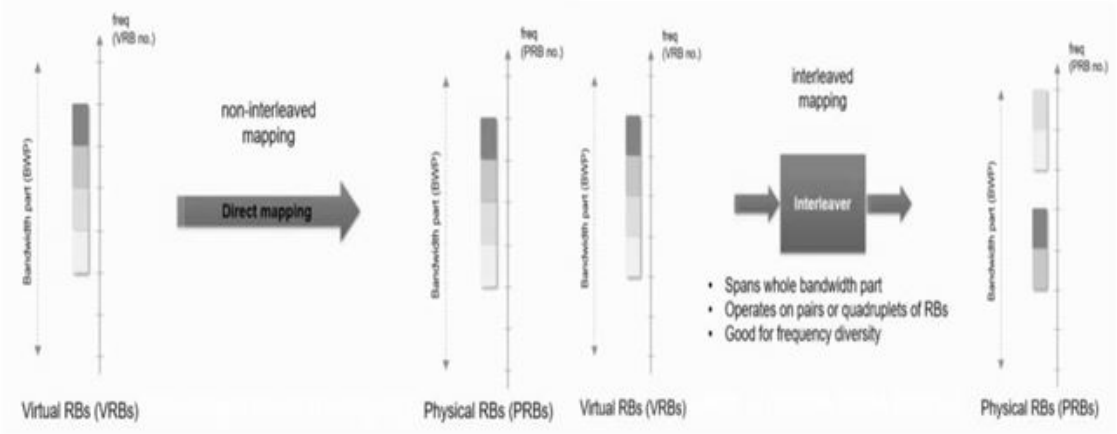


Figure4. Mapping from Virtual RBs to Physical RBs

TableII.Subcarrier Spacing and Throughput for differen tmapping techniques at SNR (5dB)

Subcarrier Spacing(kHz)	***MappingTechnique***	***Throughput (%)***
15	64QAM/256QAM	20/20
30	64QAM/256QAM	40/40
60	64QAM/256QAM	40/40
120	64QAM/256QAM	50/50
240	64QAM/256QAM	100/50

TableIII. Frame Structure

Subcarrier Spacing (μ)	Number of OFDM Symbols per Slot (N_{symb}^{slot})	Number of Slots per Subframe ($N_{slot}^{subframe,\mu}$)	Number of Slots per Frame ($N_{slot}^{frame,\mu}$)
0 *15 kHz*	14 *1 ms*	1 *1 slot x 1 ms = 1 ms*	10 *10 ms*
1 *30 kHz*	14 *500 μs*	2 *2 slots x 500 μs = 1 ms*	20 *10 ms*
2 *60 kHz (normal CP)*	14 *250 μs*	4 *4 slots x 250 μs = 1 ms*	40 *10 ms*
2 *60 kHz (extended CP)*	12 *250 μs*	4 *4 slots x 250 μs = 1 ms*	40 *10 ms*
3 *120 kHz*	14 *125 μs*	8 *8 slots x 125 μs = 1 ms*	80 *10 ms*
4 *240 kHz*	14 *62.5 μs*	16 *16 slots x 62.5 μs = 1 ms*	160 *10 ms*

3. SIMULATION RESULTS

This section shows the experiment result of a 5G NR link and its performance parameters, as expressed according to the standard of 3GPP NR. The outcome of the simulation describes the highest achievable throughput of the link provided by the existing resources for data communication. The graph defines the throughput over Signal to Noise ratio (SNR) across all antennas. Subcarrier dividing of 240 kHz with 64 QAM adjustment plot produces 100 percent of throughput, though 15 kHz with 64QAM/256QAM gives an exceptionally low throughput of 20%.

In 5G remote innovations, we have variable subcarrier separating, more prominent the worth more the quantity of openings per outline, lesser the symbol duration (up to 4μs), and maximum bandwidth (up to 400MHz), minimum scheduling interval (up to 0.06ms) andprovidesbetter performance. (Upto 400MHz), minimum scheduling interval (up to0.06ms) and provides better performance.

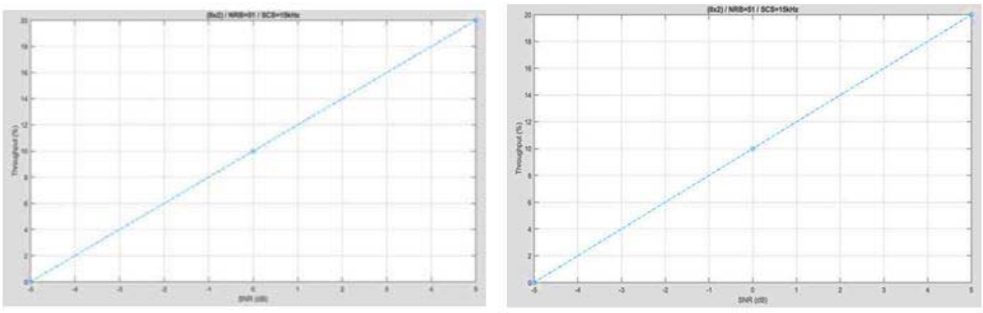


Figure.5(a) and (b):Throughput (%) versus SNR(dB) with 64QAM, 256QAM modulation scheme and subcarrier spacing of 15kHz

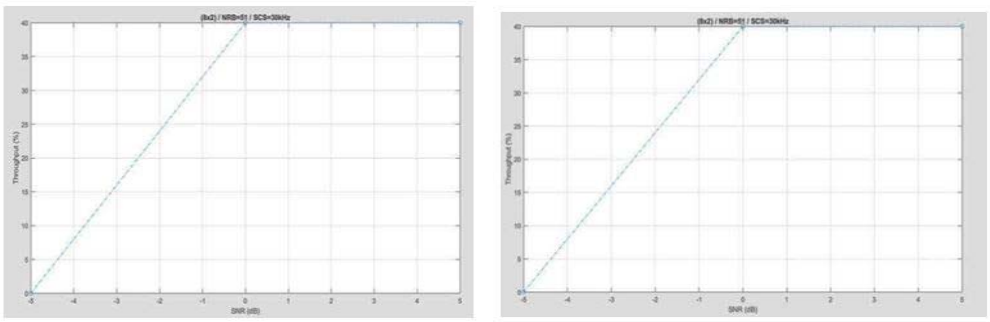


Figure.5(c) and (d):Throughput (%) versus SNR(dB) with 64QAM, 256QAM modulation scheme and subcarrier spacing of 30kHz.

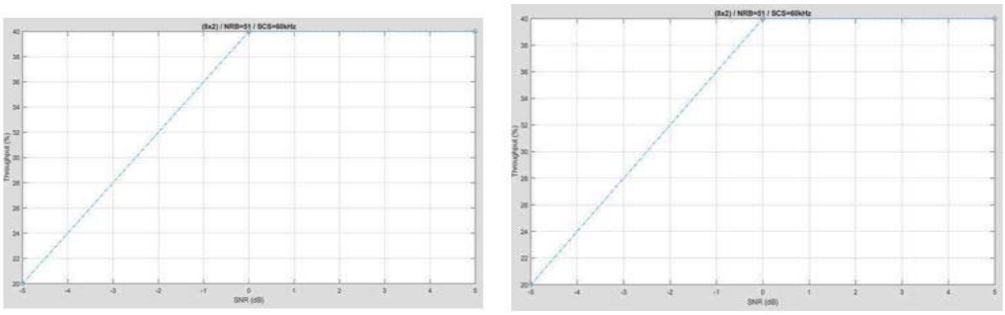


Figure.5(e) and (f):Throughput (%) versus SNR(dB) with 64QAM, 256QAM modulation scheme and subcarrier spacing of 60kHz.

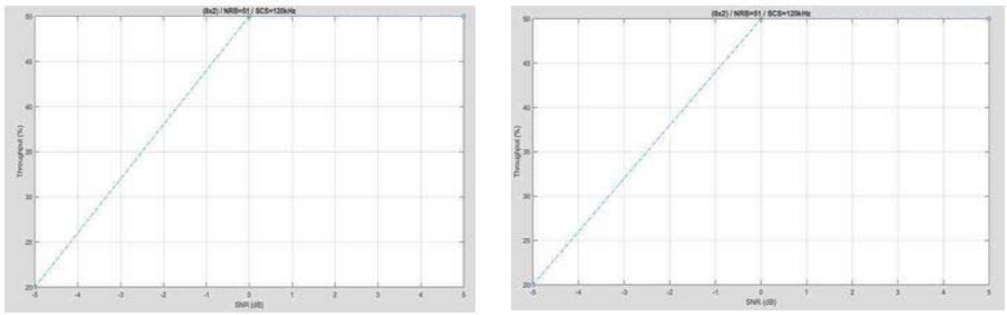


Figure.5 (g) and (h): Throughput (%) versus SNR (dB) with 64QAM, 256QAM modulation scheme and subcarrier spacing of 120 kHz.

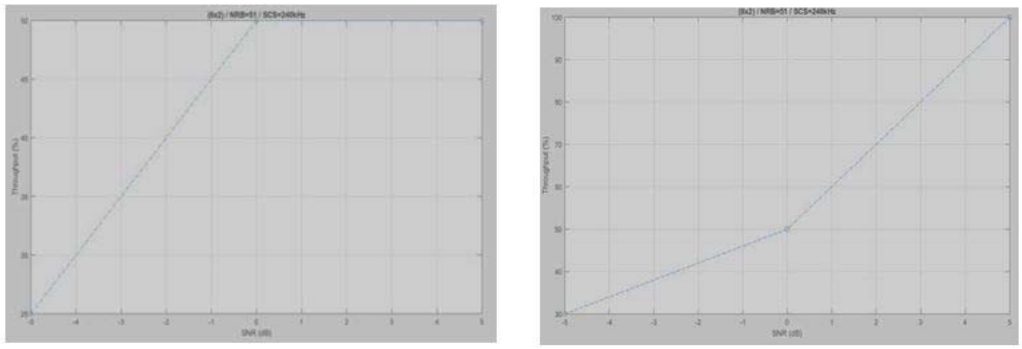


Figure.5(i) and (j):Throughput (%) versus SNR(dB) with 64QAM, 256QAM modulation scheme and subcarrier spacing of 240kHz

4. CONCLUSION

The next generation, digital 5G cellular network uses millimeter waves, which provides a short range of communication. The data rate is enhanced by using the MIMO technique; each cell consists of several antennas communicating with the wireless gadget, in this way different bitstreams of information will be communicated at the same time in equal. The above paper has given a framework of the key of 3GPP NR condition characterizing the condition of art in 5G remote framework, with an emphasis on the actual layer. The simulation result concludes that the throughput performance of a 5G NR link, as expressed by the 3GPP NR criterion is high in the case of subcarrier spacing of 250 kHz, 64QAM modulation with 8×2 antenna feature.

5. REFERENCES

[1] Jing Zhuand HaitaoLi,"On the Performance of LTE Physical Downlink Shared Channel", IEEE International Conferenceon Computer Science and Network Technology,pp.983-986, 2011.

[2] H. Chen, D. Mi, M. Fuentes, D. Vargas, E. Garro, J. L. Carcel, B. Mouhouche, P. Xiao, and R. Tafazolli,"Pioneering Studieson LTE eMBMS:Towards 5G Point-to-Multipoint Transmissions,"IEEEThe Sensor Array and Multichanne (SAM) Workshop,2018.

[3] W.Guo ,M.Fuentes,L. Christodoulou, and B.Mouhouche,"Roads to Multimedia Broadcast Multicast Servicesin 5G New Radio,"in IEEE International Symposiumon Broadband Multimedia Systems and Broadcasting,June2018.

[4] 3GPP TS 36.212 v15.1.0, "Evolved Universal Terrestrial Radio Access (E-UTRA); Multiplexing and channel coding,"April2018.

[5] Santosh M Nejakar, Dr. P G Benakop and Sharanabasappa R R" A Survey of the Challenges and Opportunities towardsNext Generation 5G Technology" Grenze International Journal of Engineering and Technology, Vol. 3, Issue 3, pp. 165-174,June2017.

[6] International Telecommunications Union(ITU),Draft New ReportITU-RM.,"Guidelines for evaluation of radio interface technologies for IMT-2020,"Oct.2017.

[7] Santosh M Nejakar, Dr. Prabhu and G Benakop "Orthogonal Frequency Division Multiplexing Modulation Scheme for4G/5GCellularNetwork" Euro. J.Adv.Engg.Tech., vol.2,Issue3,pp.46-50, 20.

[8] Li, Q. C., Niu, H., Papathanassiou, A. T., & Wu, G. (2014). 5G network capacity: Key elements and technologies. IEEE Vehicular Technology Magazine, 9(1), 71-78.

[9] Islam, S. R., Avazov, N., Dobre, O. A., & Kwak, K. S. (2016). Power-domain non-orthogonal multiple accesses (NOMA) in 5G systems: Potentials and challenges. IEEE Communications Surveys & Tutorials, 19(2), 721-742.

[10] Harish Reddy M and Dr. B. Rebekka,"Power allocation policies for QoS satisfaction in IoT using NOMA", IEEE Conference Record: # 42666; IEEE Xplore, ISBN: 978-1-5386-3570-4, 2018.

[11] Haq, N. A., & Sarvagya, M. (2018, February). Analysis on Channel Parameters and Signal Processing methods at mm-wave for 5G networks. In 2018 Second International Conference on Advances in Electronics, Computers and Communications (ICAECC) (pp. 1-6). IEEE.

[12] Nalband, A. H., Sarvagya, M., & Ahmed, M. R. (2020). Power Optimization and Antenna Selection Techniques for Hybrid Beam forming in MMWAVE Massive MIMO Systems. Power, 29(9s), 277-287.

[13] Saito, Y., Benjebbour, A., Kishiyama, Y., & Nakamura, T. (2013, September). System-level performance evaluation of downlink non-orthogonal multiple access (NOMA). In 2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC) (pp. 611-615). IEEE.

[14] Al-Abbasi, Z. Q., & So, D. K. (2017). Resource allocation in non-orthogonal and hybrid multiple access system with proportional rate constraint. IEEE Transactions on Wireless Communications, 16(10), 6309-6320.

[15] Islam, S. M., Zeng, M., & Dobre, O. A. (2017). NOMA in 5G systems: Exciting possibilities for enhancing spectral efficiency. ArXiv preprint arXiv: 1706.08215

[16] Liang, X., Gong, X., Wu, Y., Ng, D. W. K., & Hong, T. (2018, December). Analysis of outage probabilities for cooperative NOMA users with imperfect CSI. In 2018 IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC) (pp. 1617-1623). IEEE.

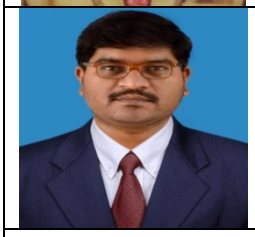
[17] Al-Abbasi, Z. Q., & So, D. K. (2017). Resource allocation in non-orthogonal and hybrid multiple access system with proportional rate constraint. IEEE Transactions on Wireless Communications, 16(10), 6309-6320.

[18] Aldababsa, M., Toka, M., Gökçeli, S., Kurt, G. K., & Kucur, O. (2018). A tutorial on nonorthogonal multiple access for 5G and beyond. Wireless communications and mobile computing, 2018.

[19] Ding, Z., & Poor, H. V. (2013). Cooperative energy harvesting networks with spatially random users. IEEE Signal Processing Letters, 20(12), 1211-1214.

[20] Xu, P., Yuan, Y., Ding, Z., Dai, X., & Schober, R. On the Outage Performance of Non-Orthogonal Multiple Access with One-Bit Feedback. IEEE Transactions on Wireless Communications, 1-15.

Biographies

	**Dr. Santosh M Nejakar** received B.E degree in E&C Engineering from VTU, Belgaum in 2008, M.Tech in Digital Electronics and Communication Systems from VTU, Belgaum in 2013 and also awarded with Ph.D. by VTU, Belagavi in 2021. He is currently working as Associate Professor and Head of the Department in the Electrical and Electronics Engineering Jain Institute of Technology, Davangere, Karnataka, India. Totally 14 papers of him are published in reputed journals. He has filed 21 patents out of which 10 patent have received its grant. His area of interest includes Wireless Communication, VLSI and Embedded systems. G Networks and Blockchain Techology.
	**Suprith PG** accepted his four year certification in Electronics and Communication Engineering and expert's in Digital Electronics and Communication System from Visveswaraya Technological University, Belgaum, India, in 2009 and 2013, separately. Presently, he is chasing after a Ph.D. from REVA Unversity, Bengaluru, India. His exploration advantages incorporate Non-Orthogonal Multiple Access (NOMA), millimeter-wave NOMA, and Wireless Sensor Networks. He is a coach for the E-Yantra discussion Robotics at IIT Bombay. He is an individual from ISTE.
	**Dr.** M R Ahmed (S'13–M'15–SM'17) received the B.E. and M. Tech. degree in electronics and communication engineering and computer networking from Visveswaraya Technological University, Belgaum, India, 2007 and 2010, respectively. He received the Ph.D. degree in electronics and communication engineering from the Jain University, Bengaluru, India, in 2016. Since 2011, he has been with REVA University, Bengaluru, India, where he is currently an Associate Professor and Assistant Director of School of Multidisciplinary Studies. He is guide of IEEE EMBS understudy part and tutor to IEEE ComSoc understudy section at REVA University. He has filled in as TPC individual from IEEE ICAECC meeting, and specialized analyst for 7 IEEE diaries and exchanges alongside various accomplished diaries and gathering procedures. His exploration advantages incorporate the computational mental neuroscience, WSNs and 5G. He has distributed in excess of 75 exploration papers in refereed diaries, global gatherings and has filled in as welcomed speaker. Riyaz is the counselor and the advisor to various Higher Educational Institutes for their change into an Entrepreneurial University for Knowledge Economy.
	**Dr. Santosh Herur** gotten B.E degree in E&C Engineering from VTU, Belgaum in 2009, and M.Tech in Digital Electronics from VTU, Belgaum in 2012 and also awarded with Ph.D. by VTU, Belagavi in 2021. He is at present filling in as Associate Professor and Head of the Department in the Electronics and Communication Engineering Department, Jain Institute of Technology, Davangere, and Karnataka, India. Totally 11 papers of him are published in reputed journals. Awarded with Best Research Paper for one of his paper presented in an International Conference. He has filed two patents out of which 1 patent have received its grant. His area of interest incorporates Image Processing, VLSI and Embedded frameworks. He has ISTE life-time enrollment.

A Review on the Integration of Blockchain and IoT

Naman Jain, Namita Tiwari*, Meenu Chawla*

*Research Scholar, *Professor*
Maulana Azad National Institute of Technology, Bhopal, India
namannj45@gmail.com,
namitatiwari21@rediffmail.com, chawlam@manit.ac.in

Abstract

Blockchain innovation has been accelerated since the development of Bitcoin, the first and the biggest of the digital forms of money. Distributed ledger technology and blockchains are generating a lot of interest and spawning a lot of projects across a wide range of industries.The Internet of things (IoT) is a quickly developing innovation that improves people's association with one another and with their surroundings.It is also a key player behind the advancements in addition to many other fields. In order to meet their design requirements, IoT uses various types of sensors, smart meters,RFIDs, and actuators resulting in networks of growing size and complexity.This comes at the cost of raising scalability, security, authenticity, and reliability issues. This is where the blockchain fits and solves many IoT-related issues. Blockchain distinct features of transparency, traceability, reliability and security nominate it to play a vital role in tackling the problem in IoT networks and resolving their issues.This paper explains the main problems facing Iot-based systems and the role of blockchains in addressing them and also surveys the current work in the literature on researches discussing the integration of IoT with blockchain.

Keywords: Blockchain, Internet of Things (IoT), Challenges, Integration.

River Journal, 1–6.

1 Introduction

Technology for data and communication is rapidly evolving [1]. The advancement of semiconductor devices and communication technologies has enabled a vast number of devices to communicate over the web/internet. These devices usually have sensors that can find information from the surrounding environment.The expression "Internet of Things" (IoT) was first presented by Kevin Ashton in 1999 [2]. IoT is the place where the physical and virtual universes meet.It is the major empowering agent of the carefully associated world where information gathered from gadgets and sensors are exchanged through the web to such an extent that items can be distantly observed and controlled without human intercession [2]. Blockchain, a type of distributed ledger technology, has been dubbed "the next big thing" by the popular press. In simple terms, a blockchain is a data structure that enables the creation and sharing of a tamper-proof digital record of transactions. It is very hard to change or remove blocks of information that are stored on the blockchain record [13]. The record consists of cryptographically connected blocks of transactions, which form a blockchain. It is very difficult to remove or alter information from blockchain ledger. The network authorises blocks using cryptographic methods. Each block also contains the hash value of the preceding block ("parent"), timestamps, and a noncerandom number, which is a unusual number used to validate the hash. A block can be added to the chain if the majority of the system's nodes agree on the authenticity of the transactions in it via consensus [6].

Currently, blockchain technology is categorized mainly into three categories: public blockchain, private blockchain, and consortium blockchain [5]. A private blockchain allows only those nodes from a certain organisation to join; it does not allow any nodes. It is also known as a integrated network since it is completely controlled by a one organisation. All records in a public blockchain are transparent to the whole public, and anybody may participate in the agreement cycle; hubs are not limited. In consortium blockchain, a gathering of preselected nodes can take part in an agreement cycle, consortium blockchain is made by a couple of associations is to some extent decentralized since just a few pieces of hubs would be chosen to decide the agreement.

2 Challenges in IoT System

It is well understood that IoT is based on the combination of numerous technologies such as communication and information technologies, sensors, and computing in addition to data analytic in collaboration to establish smart systems. The integration of such technologies results in increasing the complexity of IoT networks, especially for expanded ones. The surprising improvement of the IoT framework has opened innovative possibilities in different fields. In any case, the IoT actually has a few issues that stand like a divider notwithstanding the guaranteed spreading of IoT objects.The lack of trust and assurance is one of these concerns. The present IoT centralised approach employs a centralised third-party agency with complete control over data collecting and processing from different IoT items, with no constraints on how the information obtained is utilised. As a result, the central authority functions as a gemstone box for IoT users, which is a persuading circumstance for the majority of IoT device owners. The detailed challenges of IoT system is described in the Table 1.Blockchain is a capable to provide solution to the problem facing by IoT. Since, currently IoT system is large scale centralized system in which large number of devices interact through servers performing different tasks such as data storing, analyzing and authenticating.Table 1

3 Additional Blockchain Factors to Consider

Additional considerations must be considered when selecting whether to use a blockchain, and it must be determined if these aspects restrict one's capacity to apply a blockchain or a certain category of blockchain:

Visibility of data: Permission-ed type of blockchain systems will or will not make blockchain data available to the general public. Data is allowed to access to the node within the network.The blockchain may be seen and contributed to by anybody with access to a permissionless blockchain network. This raises a number of concerns that must be addressed. Is it necessary for everyone to have access to the application's data? Is it possible that having public data is harmful?

Records of Transaction : Some Blockchain system allow to available the public history of digital transaction available, from origin to all transaction. This factor may be advantageous for certain solutions but not for others.

Table 1 Blockchain deals with IoT challenges

Challenges	How Blockchain Resolve this Challenges
Failure at a single point	Blockchain technology enables distributed and independent interaction between system nodes, eliminating the need for a central server to oversee and manage interaction and operation between them. As a result, if one node goes down or fails, the entire network remains unaffected, overcoming the central piece of failure problem associated with the monolithic iot ecosystem.
Privacy	In blockchain system, cryptography is used which ensure that only authorized participants can use and perform data transaction.
Security	Blockchain provides significantly higher security by leveraging cryptography or the public key infrastructure, which provides more protection against numerous threats. Blocks are connected using a hash algorithm to dorm a chain where each block contains the hash of the last block embedded in its header. If any hackers want to alter the information of any block, they need to evaluate the hash along with the hashes of all successive block which is almost impossible task. Any adjustment to data is only permitted with the permission of the majority of network contributors, ensuring data integrity [11].
Cost	Blockchain is a decentralized and distrubted structure where a third agency or third-party is not necessary to supervise user communications So, no need to spend money on for installing a server with high software [13].
Scalability	Since the number of IoT devices are increasingly day by day so, the centralised paradigm is incapable of scaling and operating efficiently. The distributed and decentralized nature of the blockchain will manage the increase in IoT devices in an efficient way.
Flexibility	Through a variety of open source, Blockchain creates a flexible ecosystem for IoT devices of all kinds.
Susceptibility to Manipulation	In blockchain, manipulation, alteration, or update is only permitted with the permission of all participating parties. The blockchain system stops the manipulation or change of information to ensure data integrity.

Input of false data: Because a blockchain is a distributed ledger, some individuals may contribute fraudulent data to imitate data from legitimate sources (such as sensor data).The validation of information that enters a blockchain platform is challenging to automate. Wherever practical, smart contract application may include extra validattion to assist verify data.

Compliance: A system's usage of blockchain technology does not exclude it from adhering to all applicable rules and regulations. For example, there are several compliance issues relating to legislation and rules relating to PII or GDPR that state that some data should not be stored on the blockchain.

4 Literature Review

Using blockchain technology in conjunction with IoT is one of the most well-known topics that has drew the attention of a number of academics looking for ways to solve the challenges associated with centralised IoT design. According to recent research, blockchain technology is the best option for dealing with privacy and security issues in the IoT system. For instance, Zibin Zheng [5] presented an overview of blockchian architecture, technical challenges of blockchain like scalability, privacy leakage, energy consumption. Also, proposed how this scalability challenge of Blockchain is removed by storage optimization of blockchain and redesigning blockchain. Privacy leakage can be preserve by the public key and private key. The advantages and difficulties of combining blockchain with IoT system were also covered in this article. Nir Kshetri discussed whether blockchain is strengthening the internet of things(IoT), the answer of this question is "maybe", because [13] Blockchain overcome all the challenges of IOT - Cost and capacity constraint, Susceptibility to manipulation, Inadequate architecture, Downtime and service unavailability on cloud servers. Kiktenko [12] The proposed approach for constructing an IoT device blockchain comprises determining the type of IoT device, as well as the use cases and applications that will be implemented, the design of storage, security concerns, and the requisite blockchain characteristics. Gao, Zhimin—& Xu [3] Proposed a concept of Block-id which can be used for identity management, an unique system for managing people's identities that makes use of biometric authentication and trusted computing technologies. IoT gadgets are mostly used to deal with our everyday concerns and to work with our lives by detecting and collecting various types of data about our actual surroundings, which is then utilised to create new digital administrations. The Internet of Things has made massive development on a global basis, with billions of instruments sold and utilised globally to date across a variety of different markets [4].

5 Conclusion

Blockchain and IoT are two fantastic technologies, and combining them would produce excellent results in every imaginable sector. The paper discusses the development of the technologies, their integration, and their applications. The challenges IoT is dealing with centralized structure and blockchain role in resolving IoT issues. It presents the description of both IoT and Blockchain. This paper surveyed the work on IoT-blockchian integration and describe a summary of their work.

References

[1] Pavithran, D., Shaalan, K., Al-Karaki, J. & Gawanmeh, A. Towards building a blockchain framework for IoT. *Cluster Computing*. 23, 2089-2103 (2020)

[2] Balaji, S., Nathani, K. & Santhakumar, R. IoT technology, applications and challenges: a contemporary survey. *Wireless Personal Communications*. 108, 363-388 (2019)

[3] Gao, Z., Xu, L., Turner, G., Patel, B., Diallo, N., Chen, L. & Shi, W. Blockchain-based identity management with mobile device. *Proceedings Of The 1st Workshop On Cryptocurrencies And Blockchains For Distributed Systems*. pp. 66-70 (2018)

[4] Atlam, H., Walters, R. & Wills, G. Internet of things: state-of-the-art, challenges, applications, and open issues. *International Journal Of Intelligent Computing Research (IJICR)*. 9, 928-938 (2018)

[5] Buterin, V. On public and private blockchains (2015). *URL: Https://blog. Ethereum. Org/2015/08/07/on-public-and-private-blockchains*. (2019)

[6] Banerjee, M., Lee, J. & Choo, K. A blockchain future for internet of things security: a position paper. *Digital Communications And Networks*. 4, 149-160 (2018)

[7] Nofer, M., Gomber, P., Hinz, O. & Schiereck, D. Blockchain. *Business Information Systems Engineering*. 59, 183-187 (2017)

[8] Thakore, R., Vaghashiya, R., Patel, C. & Doshi, N. Blockchain-based IoT: A survey. *Procedia Computer Science*. 155 pp. 704-709 (2019)

[9] Atlam, H., Azad, M., Alzahrani, A. & Wills, G. A Review of Blockchain in Internet of Things and AI. *Big Data And Cognitive Computing*. 4, 28 (2020)

[10] Fernández-Caramés, T. & Fraga-Lamas, P. A Review on the Use of Blockchain for the Internet of Things. *Ieee Access*. 6 pp. 32979-33001 (2018)

[11] Atlam, H. & Wills, G. An efficient security risk estimation technique for Risk-based access control model for IoT. *Internet Of Things*. 6 pp. 100052 (2019)

[12] Kiktenko, E., Pozhar, N., Anufriev, M., Trushechkin, A., Yunusov, R., Kurochkin, Y., Lvovsky, A. & Fedorov, A. Quantum-secured blockchain. *Quantum Science And Technology*. 3, 035004 (2018)

[13] Kshetri, N. Can blockchain strengthen the internet of things?. *IT Professional*. 19, 68-72 (2017)

Consensus and control of Micro Grid using Graph Theoretical Approach

Gaurav Singh Negi[1], Mukul Kumar Gupta[2], Nitin Kumar Saxena[3]
[1]Assistant Professor, UIT, Uttaranchal University, Dehradun, India
[2]Assistant Professor, School of Engineering, UPES, Dehradun, India
[3]Professor, Krishna Institute of Engineering and technology, Ghaziabad, India
E-mail[1]: technonegi@gmail.com, Email[2]: mukulvjti@gmail.com
Email[3]: nitinsaxena.iitd@gmail.com

Abstract.

The paper introduces the micro grid survey and the renewable energy sources connected to micro grid, which forms a graph structure. With the help of graph theoretical approach how synchronization and consensus can be achieved which is very useful for any system operation. With the consensus, one can easily achieved the synchronization, which leads to less time delay. Various control strategies studied and compared for microgrid as a system. The four renewable energy sources are connected in microgrid which forms a multi agent system in which micro grid is works a master whereas four renewable energy sources acts as a slave and this is case of multi agents system. The objective of this survey is to attract the researcher in field of control and stability of microgrid and increase the chances of implementation of intelligent control methods. In the paper different renewable energy sources connected to microgrid as individual agents and works as a whole to meet the power requirements efficiently. The goal is to assess and classify the layout control techniques and assessment algorithms for the microgrid structures to maintain stability, reliability, load variations in standalone operation mode.

Keywords. Microgrid, Renewable Energy Sources, Multi-Agent System, Consensus Algorithm.

1. INTRODUCTION

For a variety of reasons, electricity generation is shifting away from traditional energy sources and toward renewable energy sources. This is owing to weather alter and the rapid exhaustion of traditional power resource such as gas, coal and oil, all of which are seeing exponential increases in their energy use [1]. The MicroGrids (MGs) which is integration of distributed energy resources are not only cost-effective, flexible and consistent however, gives ecological advantages as weigh against to the accessible main gird network. Annual load growth creates a gap between demand and traditional power output, as we can observe. [2].The MGs' rapid installation and on-demand requirements make them a viable alternative to today's energy-related issues, filling the gap between demand and generation. In most MGs, renewable energy sources (RES), like as solar photovoltaic cells, wind energy, fuel cells, and micro turbines, as well as diesel generators, are employed. [3].So as we see that many of the present MGs are linked to stand-alone or grid-connected systems. These two forms of microgrid connections are often used, with standalone MGs taking precedence above economic benefits in order to provide a steady power supply to consumers. [4].Grid-connected MGs are mostly acceptable to bear the incorporation of distributed generation units and, specially, of RES in distribution networks [5].Infiltration of these renewable energy resources is growing at constant speed. Initially, the Low Voltage networks are going to transform from passive to active. After that, amplify the complexity by using power electronics converters. That’s why there is two solutions for function i.e. islanded and grid-connected modes [6]. Later, such MGs' control architecture may be used to build smart grid (SG) features. A brief conversation on critical controllers for MGs are presented in [7-11], offers a short overview on decentralized control techniques. Only some distributed control schemes have been mentioned in [12]. While in [13, 14], a few applications on hierarchical control schemes are mentioned. a few overview

mechanism also are to be had on MGs as a selected control feature viz. electricity allocation [15], voltage and frequency control [16].Many of those overview literatures discusses the concepts in the back of those control strategies, however, the use of MAS in intelligent controlling of micro grid is not further extended. Also unbiased control of power equipment's operations the use of multi- agent systems (MAS) has been furnished to address several limits [17]. MAS are combination of numerous intelligent agents that linked together to resolve bigger troubles that not be covered or complete by the individual agents [18].

2. MICROGRID SYSTEM AND ITS ARCHITECTURE

Microgrid consists of different energy resources working in small scales where economic management becomes complicated. The MGs System is shown in fig 1. which shows are as connected with four renewable energy resources and these sources acts like individual agents of Multi Agent System. The main function of MAS is for MGs to interact with main grid to provide better operation. MGs are made up of a variety of DERs and users that may be represented as autonomous agents. Each agent has a specific amount of intelligence based on its kind and structure, which it uses to achieve pre-determined goals. MAS is a good choice for establishing a intelligent energy management system in MGs since it is strongly connected to distributed intelligence. It presents a framework for determining the optimum agent action by combining artificial intelligence and mathematics techniques. There are three common features shared by all agents: initiative, reactivity, and social capacity [19].

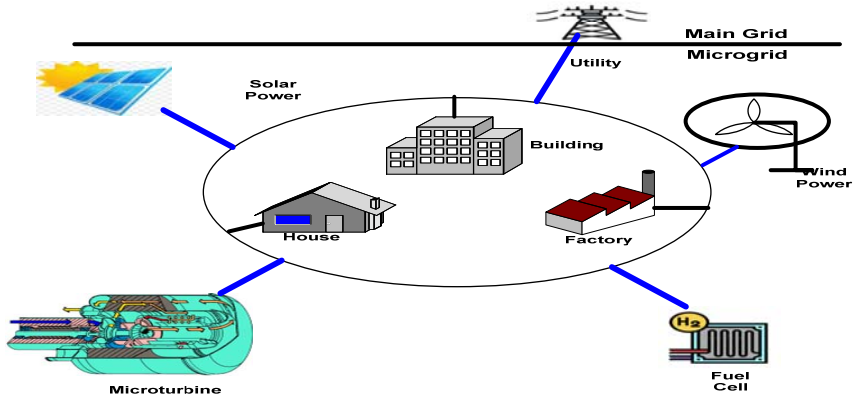


Figure 1: Microgrid System

Fig. 2 shows DC MGs containing nonconventional energy sources along with battery storage. Then, using power electronic converters, common energy resources are combined to form a dc bus. One of MG's top priorities is capable for function by a variety of RES and satisfy power requirement in the situation of a power outage. To accomplish the intended result, subsystems might communicate with one another. [20-21].

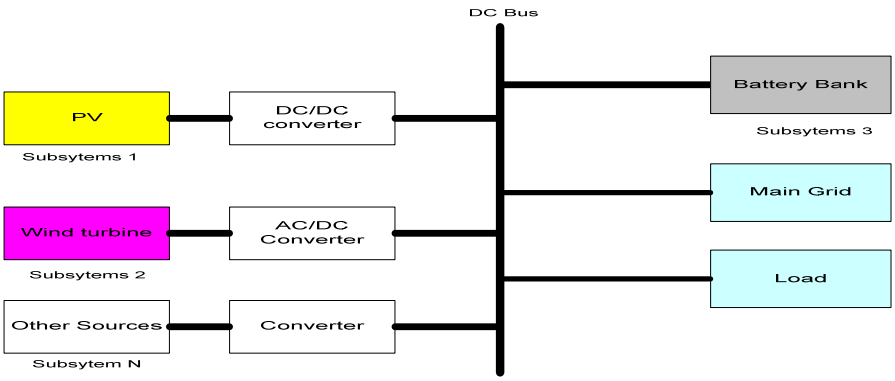


Figure 2: DC Microgrid

2.1. Microgrid architecture and control requirements

The main idea behind MGs is to connect a small number of DG devices in order to manage them efficiently without creating a complex community. A hierarchical manage method, a factor of not point of common coupling (PCC), distributed controls the usage of nearby data, and a selected location are the major additives that allow DG devices to be incorporated in a scientific manner to make sure consistent system performance. The MGs idea and distribution community are depicted in Fig. 3.

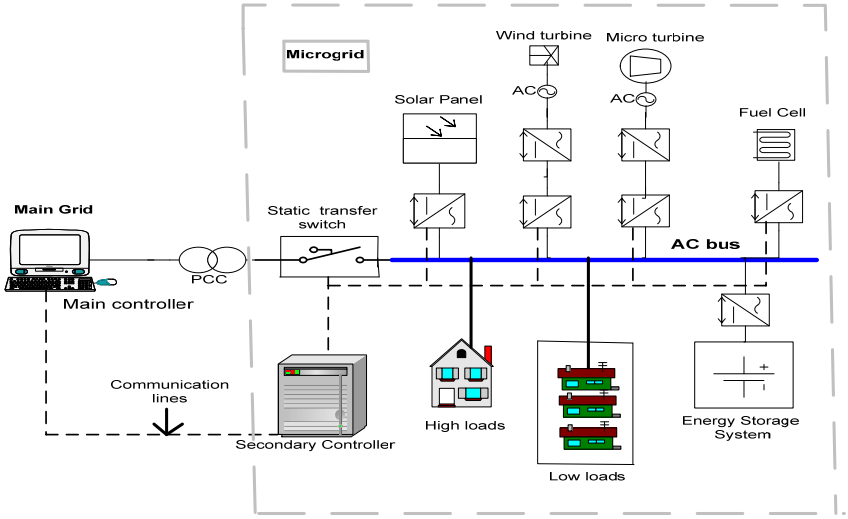


Figure 3: Modern distribution network.

2.2 Control Strategies: scenarios

Control strategies for MGs system with renewable energy resources can be divided into two main (a) Centralized control (b) Decentralized control which is shown in fig 4. Which can be further separated into sub- types. The centralized control is generally used in small MGs whereas decentralized control is used in high rated MGs having large numbers of component's due to centralized control is sluggish.

Centralized Control-Using a centralized control mechanism, MGs systems may be monitored and regulated separately. Secondary control approaches, such as power quality regulation, are based on renewable generation and load, with transitory control goals. On these classifications centralized control the load sharing ratio of the energy storage system can be adjusted on the MGs system level efficiency and / or the power capacity estimate.

Decentralized Control- Local measurements are used in the decentralized control technique, which decides operation at the component level using a built-in algorithm. Given the MG design, the multi-agent system (MAS) is the most effective approach to implement the decentralized control mechanism. The decentralized control approach of the distributed energy storage system relies on the adjustment of the primary MGs descent control, depending on the level of the MGs system.Both centralized and decentralized control techniques offer advantages and disadvantages that vary based on the MG's operational situation. So another control technique, distributed multi-agent control, is presented here, and it is a study topic in control system development. It analyses a system with autonomous agents that are linked via a decentralized communication network. The following is a summary of the MGs control functional requirements for efficient and reliable operation.

- The load sharing between each component of MGs mainly renewable sources is efficient.
- The transition between island and grid connected modes is smooth i.e. proper synchronization.
- Optimization of DERs units to reduce MGs running cost for economic dispatch.

The typology of control strategies for MGs is shown in figure below and comparison of control methods is presented in tabular form.

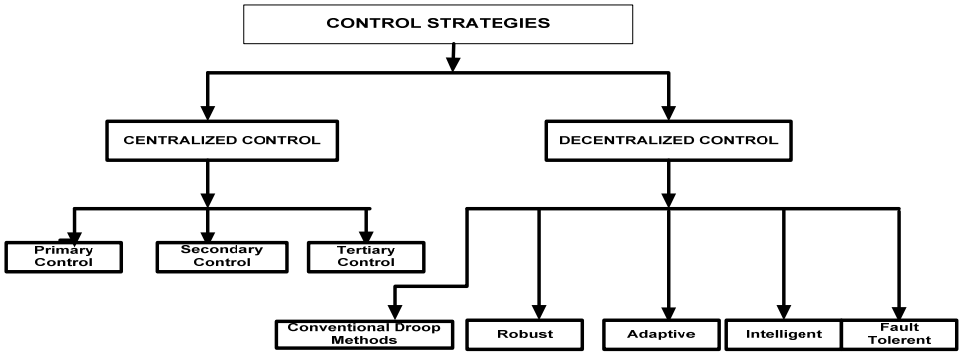


Figure 4: Main Control Strategies

Table 1: Comparison of control methods

Factors	Centralized	Decentralized
DG possession	Self-possessor	Many possessor
Goals	Determined tasks	undecided tasks
In service employees	accessible	In accessible
Flexibility	fewer	large
Extension	difficult	simple
Communication requirements	Much larger	small
Establishment expenditure	Large	small
Additional significant operation	doable	unfeasible

3. MULTI-AGENT SYSTEM

Multi Agent System (MAS) is a group of several agents in the system which is having a common objective and make intelligent decision with respect to energy management system with the rapid development of computing and allied areas like advanced sensor, actuators and devices in past decade, use of MAS is growing exponentially. In recent time, MAS have been widely investigated because of their application in spacecraft formation, social network, energy network, smart grid, micro grid power system, computer games, flight formation, robotics, transportation etc.MAS have many features like multi-tasking, cost effectiveness, distribution, fault tolerance, flexibility etc [22-24]. Multi agent system will be incorporated with the physical and chemical properties [25-30].

The use of Multi agent system (MAS) is mainly brings the intelligence in the operation of micro grid. With the lack of flexibility of SCADA cannot be applicable for high penetration of distributed energy system so this problem can be overcome by MAS. In MAS complex control problem divided in too many subparts and individual agent can solve these individual issues.In general the consensus algorithm is given by the following equation

$$\dot{x}_i(t) = \sum_{j \grave{o} N(i)} a_{ij}\left(x_j(t) - x_i(t)\right) \quad i = 1,2,\ldots\ldots n, \tag{1}$$

Where a_{ij} is the (i, j) entry of the adjacency matrix of the associated graph at time t, and $x_i(t)$ is the information state of the i th agent. Graph theory is most suitable for MAS as main agent is treated as master whereas remaining agent as Slave. All slaves are connected to each other, which share information among them and follow the master command. Here considering four different renewable energy sources are attached with Micro gird whose aim is to maintain the synchronization so that there will be least amount of energy loss. The effect of consensus is shown in figure.

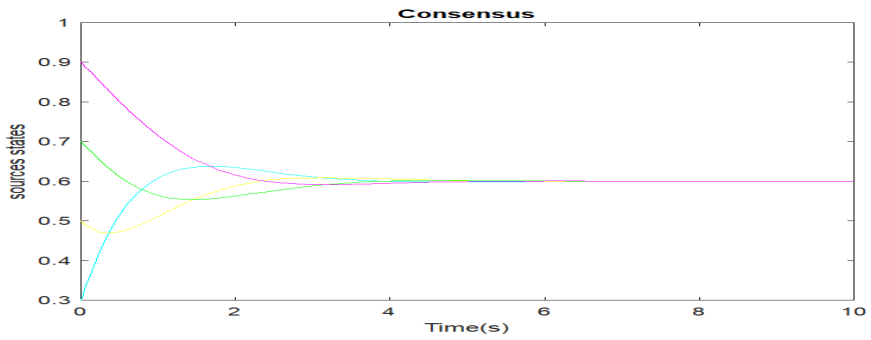


Figure 5: Consensus of Multi Agent System

Figure 5 shows multi agent technology interaction with MGs agents. The multi-agent method introduces various applications of the MGs, such as market operation, distributed management system and so on. Several agents were proposed, and simulated multiple micro-grid market scenarios and introduced the implementation of MAS in micro-grid operation. MAS has the following applications.The advantage of consensus is that if consensus is achieved then there will be minimum loss of energy and time. Here considering four renewable energy sources connected to MGs consensus is achieved.

- MGs control considering renewable energy, voltage control and thermal limit.
- Interaction between entities to achieve effective market operations.
- Decisions can be made with ready-made data or information.
- Expansion of existing systems / functions

4. CONCLUSION AND FUTURE SCOPE

This paper review the works related to MGs based on Graph Theory approach using MAS. Comparison of centralized and decentralized control of MGs discussed. Implementation of MAS on MGs of consensus control using graph theory approach is being suggested. Various applications of MAS on MGs have been discussed considering the electricity at islands modes of MGs and controlling through MAS. In addition, this document provides a method and enlightenment for using a multi-agent system to operate and control the MGs. The challenges faced by the MAS are also discussed, which will help researchers in this field to update and focus on exploring MAS in MGs applications. In future with more number of renewable energy sources connected to MGs problem can be considered.

5. REFERENCES

[1] Lasseter R. Microgrids and distributed generation. J Energy Eng 2007;133:144–9 ,

[2] Annaswamy, A. M (Ed.). (2013). IEEE vision for smart grid controls: 2030 and beyond IEEEPressPublications .

[3] Hatziargyriou, N. (2014). Microgrid: architecture and control. Wiley & Sons Publications, IEEE Press ISBN: 9781118720684..

[4] Hossain E, Kabalci E, Bayindir R, Perez R. Microgridtestbeds around the world: state of art. Energy Convers Manage 2014;86:132–53.

[5] Bracco S, Delfino F, Pampararo F, Robba M, Rossi M. The University of Genoa smart polygenerationmicrogrid test-bed facility: the overall system, the

technologiesand the research challenges. Renew Sustain Energy Rev 2013;18:442–59.

[6] Majumder, R. (2010). Modeling, stability analysis and control of microgridPh.D thesis Australia: Faculty of Build and Environment Engineering., Queensland University of Technology.

[7] Martin-Martinez, F., Sanchez-Miralles, A., &Rivier, M. (2016). A literature review of microgrids: A functional layer based classification. Renewable and Sustainable Energy Reviews, 62, 1133– 1153.

[8] Dragicevic, T., Lu, X., Vasquez, J. C., & Guerrero, J. M. (2016). DC microgrids - Part I: Areview of control strategies and stabilization techniques. IEEE Transactions on Power Electronics, 31 (7), 4876– 4891.

[9] Kaur, A., Kaushal, J., &Basak, P. (2015). A review on microgrid central controller.Renewable and Sustainable Energy Reviews, 55, 338– 345.

[10] Eid, B. M., Rahim, N. A., Selvaraj, J., &Khateb, A. (2016). "Control methods and objectives for electronically coupled distributed energy resources in microgrids: A review. IEEE Systems Journal, 10 (2), 446– 458.

[11] Guerrero, J. M., Chandorkar, M., Lee, T., &Loh, P. C. (2013). Advanced control architectures for intelligent microgrids—Part I: Decentralized and hierarchical control. IEEE Transactions on Industrial Electronics, 60 (4), 1254– 1262.

[12] Yazdanian, M., &Mehrizi-Sani, A. (2014). Distributed control techniques in microgrids.IEEE Transactions on Smart Grid, 5 (6), 2901– 2909.

[13] Bidram, A., &Davoudi, A. (2012). Hierarchical structure of microgrids control system. IEEE Transactions on Smart Grid, 3 (4), 1963– 1976.

[14] Canizares, C. A. (2014). Trends in Microgrid Control. IEEE Transactions on Smart Grid, 5(4), 1905– 1919.

[15] Han, H., Hou, X., Yang, J., Wu, J., Su, M., & Guerrero, J. M. (2016a). Review of powersharing control strategies for islanding operation of ac microgrids. IEEE Transactions on Smart Grid, 7 (1), 200– 215.

[16] Malik, S. M., Ai, X., Sun, Y., Zhengqi, C., &Shupeng, Z. (2016). Voltage and frequency control strategies of hybrid ac/dc microgrid: A review. IET Generation, Transmission & Distribution, 11 (2), 303– 313.

[17] Roche,Robin,Blunier,Benjamin,Miraoui,Abdellatif,Hilaire,Vincent,Koukam, Abder,2010.Multi-agentsystemsforgridenergymanagement:ashortreview. In: 36th Annual Conference on IEEE Industrial Electronics Society, IECON 2010. IEEE, Glendale ,pp.3341–3346.

[18] Weiss, Gerhard, 1999. Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence. The MIT Press.

[19] Logenthiran T, Srinivasan D, Khambadkone AM. Multi-agent system for energyresource scheduling of integrated microgrids in a distributed system. ElectrPowerSyst Res 2011;81:138–48. https://doi.org/10.1016/j.epsr.2010.07.019.

[20] A. Alzahrani, P. Shamsi and M. Ferdowsi, "Analysis and design of bipolar Dickson DC-DC converter," 2017 IEEE Power and Energy Conference at Illinois (PECI), Champaign, IL, 2017, pp. 1-6. doi: 10.1109/PECI.2017.7935733

[21] M. S. Mahmoud, M. Saif Ur Rahman and F. M. A. L. -Sunni, "Review of microgrid architectures – a system of systems perspective," in IET Renewable Power Generation, vol. 9, no. 8, pp. 1064-1078, 11 2015. doi: 10.1049/iet-rpg.2014.0171

[22] Kiran, P., Chandrakala, K. V., &Nambiar, T. N. P. (2017, June). Multi-agent based systems on micro grid—A review. In 2017 international conference on intelligent computing and control (I2C2) (pp. 1-6). IEEE.

[23] Zarma, T. A., Galadima, A. A., Modibbo, A., & Hussein, S. U. (2020, August). Review of Multi-Agent Micro-Grid Systems. In 2020 IEEE PES/IAS PowerAfrica (pp. 1-5). IEEE.

[24] Ren, W., & Cao, Y. (2010). Distributed Coordination of Multi-agent Networks. Annals of Physics (Vol. 54). https://doi.org/10.1017/CBO9781107415324.004

[25] Kumar, A., & Verma, A. (2016). Missing numbers in graceful graphs. International Journal of Control Theory and Applications, 9(21), 133–136. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85008259815&partnerID=40&md5=b85175173bb98fecc1b0ec98e2185b98

[26] Kumar, D., Chibber, V. K., & Singh, A. (2018). Physical and chemical properties of Mahua and Sal seed oils. In Advances in Intelligent Systems and Computing (Vol. 624, pp. 1391–1400). https://doi.org/10.1007/978-981-10-5903-2_146

[27] Kumar, M., Chandramauli, A., & Ashutosh. (2018). Partial replacement of fine aggregates of fire bricks with fine aggregates in concrete. International Journal of Civil Engineering and Technology, 9(3), 961–968. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85045069805&partnerID=40&md5=65b0161451cfa3352689f31d23a5b191

[28] Kumar, R., & Memoria, M. (2020). Proposed Selection Technique of Evolutionary Algorithm and its implementation for Combinatorial Problems. Proceedings - 2020 International Conference on Advances in Computing, Communication and Materials, ICACCM 2020, 2020-January, 408–412. https://doi.org/10.1109/ICACCM50413.2020.9213063

[29] Kumar, R., Memoria, M., & Chandel, A. (2020). Performance Analysis of proposed Mutation Operator of Genetic Algorithm under Scheduling Problem. Proceedings of International Conference on Intelligent Engineering and Management, ICIEM 2020, 193–197. https://doi.org/10.1109/ICIEM48762.2020.9160215

[30] Mallampalli, S. S., & Goyal, S. (2020). The Effect of Teaching Collocations to Kurdish Tertiary Students Through Mobile Learning. International Journal of Interactive Mobile Technologies, 14(18), 213–220. https://doi.org/10.3991/ijim.v14i18.16801

Algorithm for food item recognition using VGG-16 and InceptionV3 CNN

Pulkit Jain* and Paras Chawla

Department of Electronics and Communication Engineering, Chandigarh University Mohali, India
E-Mail: pulkit.mech@cumail.in; drparaschawla.ece@cumail.in
** Corresponding Author*

Abstract.

Computer Vision is impacting the retail sector by bringing revolutionary changes in shopping experience for customers. It has not only raised the competition among several retailers but after its combination with Artificial Intelligence (AI) it has automated the manual processes thereby saving costs and time. Deep Learning is also one area along with computer vision which has done wonders in the area of many open research problems like food item recognition using Machine Learning algorithms. Convolution Neural Network (CNN) is of the approaches which can fulfil the pre-requisites for solving image recognition research problems. The objective of this paper is to utilise the proposed algorithm for automated identification of food items. The algorithm has been tested successfully using VGG-16 and InceptionV3 CNN models. The results of food item recognition show nearly 96.37% in VGG-16 and 98.4 % accuracy in InceptionV3 respectively. At the end real time testing results on food items and barcode items are also presented.

Keywords. Convolution Neural Network, VGG-16, InceptionV3, Machine Learning.

1. INTRODUCTION

An increase in number of health related problems, chronic diseases like obesity and cancer have raised alarms and created an urgent need for keeping a check on diet intake. A very common and an open research issue in the area of health and nutrition are to design appropriate mechanism for measuring accurate diet intake. To have a check on the food consumed per day, individuals generally keep a record of meals taken every day. But such an exercise is carried out manually utilising text based description but it is a tedious and dreary task. To beat this issue, there have been endeavours like image recognition using camera scanning. But due to wide variety of classes of food items available specially fruits and vegetables, the recognition accuracy becomes a major challenge. In this regard, many researchers have utilised image based recognition techniques. A smart phone based system was designed for food item recognition and logging using a small dataset which could

replace traditional methods of recording food items [1]. Similar work was carried out which achieved an accuracy of nearly 62.5 % for recognition of Japanese using machine learning approach [2]. Other related works talked about the usage of image retrieval, and related image processing techniques for not only food recording but also recording the nutritional content of food items [3]. Deep learning is one of the latest area which is also being used to address this open research problem [4]. It basically refers to collection of certain algorithms to solve such complex issues. The most particular trademark is that the distinguishing and important features of images are extricated automatically during the training of the models. CNN [5] is one of the approaches which can fulfil the pre-requisites of deep-learning methodology for solving image recognition problems. CNN is currently a best in class procedure for image recognition tasks especially after validation of its excellent performance in Visual Recognition Challenge. The entire paper is organized as below: Section II talks about the pre-processing steps involved in the algorithm used for food item recognition. Section III describes the main steps of algorithm used followed by description of hardware setup designed. The architecture of the two CNN models used is discussed in Section IV. The results obtained after testing are discussed in Section V.

2. ALGORITHM-PRE PROCESSING STEPS

Simple distortions like crop, flip and scale as pre-processing operations on images during model training can play a vital role in improving the results of the algorithm. These operations are simply analogous to natural variations and background noises present in the real world scenario. Hence, these make the models obtained after training very efficient. The first step starts with conversion of image data type to float32 type. This conversion followed by the cropping and scaling operation is depicted in Figure2.1.

```
decoded_image_as_float = tf.image.convert_image_dtype(decoded_image,
                                                      tf.float32)
decoded_image_4d = tf.expand_dims(decoded_image_as_float, 0)
margin_scale = 1.0 + (random_crop / 100.0)
resize_scale = 1.0 + (random_scale / 100.0)
margin_scale_value = tf.constant(margin_scale)
resize_scale_value = tf.random_uniform(shape=[],
                                       minval=1.0,
                                       maxval=resize_scale)
scale_value = tf.multiply(margin_scale_value, resize_scale_value)
precrop_width = tf.multiply(scale_value, input_width)
precrop_height = tf.multiply(scale_value, input_height)
precrop_shape = tf.stack([precrop_height, precrop_width])
precrop_shape_as_int = tf.cast(precrop_shape, dtype=tf.int32)
precropped_image = tf.image.resize_bilinear(decoded_image_4d,
                                            precrop_shape_as_int)
precropped_image_3d = tf.squeeze(precropped_image, axis=[0])
cropped_image = tf.random_crop(precropped_image_3d,
                               [input_height, input_width, input_depth])
```

Figure 2.1. Image Conversion, Cropping and Scale Operation

The crop operation starts with random placing of a bounding box in full sized image. Then crop parameter value decides the size of the box with respect to given input image. When zero, no crop i.e. the output and input image size are identical. When 0.5 then box size is half the dimensions (height x width) of input image. Next operation almost similar to crop is ‘scale’ with one difference that the bounding box is oriented at the centre initially. When scale % is zero bounding box and input image are of same size. When 50 % then box will be in any random range in between half and full dimensions of the image. For flip operation, ‘random_flip_left_right’ function available in tensor flow library is used as depicted in Figure2.2(a) It flips the image generally along second dimension i.e. width or will pass the image with same dimensions as input. In case of batch images, each image undergoes flip operation randomly independently of other images.

```
if flip_left_right:
  flipped_image = tf.image.random_flip_left_right(cropped_image)
else:
  flipped_image = cropped_image
```

(a)

```
brightness_min = 1.0 - (random_brightness / 100.0)
brightness_max = 1.0 + (random_brightness / 100.0)
brightness_value = tf.random_uniform(shape=[],
                                     minval=brightness_min,
                                     maxval=brightness_max)
brightened_image = tf.multiply(flipped_image, brightness_value)
distort_result = tf.expand_dims(brightened_image, 0, name='DistortResult')
```

(b)

Figure 2.2. Code snippet depicting a) flip left right b) brightness adjustment operation

The brightness can also be adjusted by varying brightness variable as depicted in Figure2.2(b) The steps involved in the algorithm for the same are: a) User inputs % value in brightness variable b) Calculation of minimum and maximum value c) Automatic selection of a random value between minima and maxima d) Multiplication of this value with input image resulting in output image with variable brightness.

3. ALGORITHM USED AND HARDWARE SETUP

The entire algorithm for training of CNN models comprises of seven main steps as shown in Figure3.1(a). The first phase involves calculation of bottleneck values after complete analysis of all input images. 'Bottleneck' is the name given to the layer just before the final output layer which mainly carries out the classification task. It outputs a set of values for the classifier to properly distinguish between the various classes of fruits and vegetables.

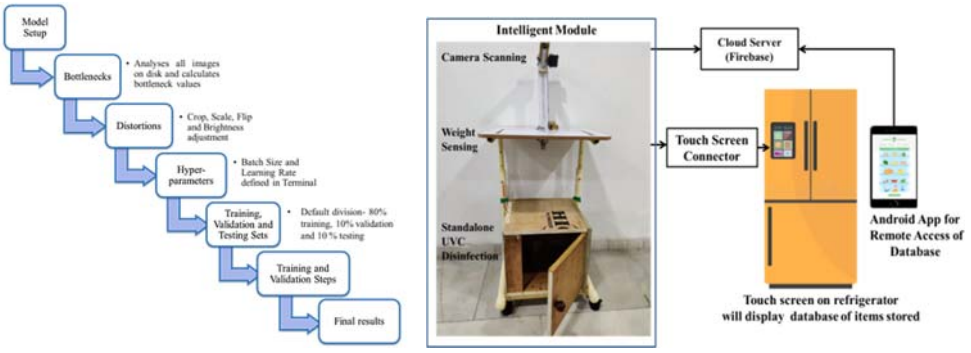


Figure 3.1 (a) Steps of Algorithm (b) Hardware Setup for food item identification

Next step is adding the distortions to improve the quality of results and getting the model trained for all worst case scenarios. Learning rate and batch size are certain hyper-parameters which play an important role to improve overall precision of model. Next step is division of the dataset into training (80%) and testing sets (20%), to avoid the problem of overfitting. The hardware setup [6] developed for applying the algorithm for food item identification comprises of intelligent module for camera scanning and weight sensing as depicted in Figure3.1(b).

4. VGG-16 AND INCEPTIONV3 CNN

VGG-16 [7] is one of the most preferred CNN architectures in the recent past. It has 16 convolution layers and has much more complex architecture than initial versions like LeNET. Only limitation is large number of parameters i.e. nearly 138 million which at times becomes difficult to handle. The flowchart depicted in the Figure4.1(a) shows

uniform structure of VGG-16. VGG-16 CNN model has been trained on standard dataset i.e. FRUITS 360 [8]. The programming language and platform used for training is Python and Google Colab.

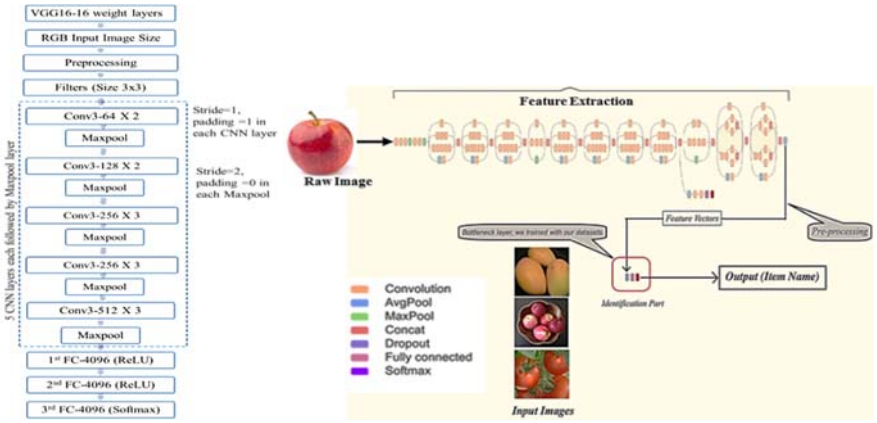


Figure 4.1 Flowchart depicting layers in (a) VGG-16 and (b) Inception-V3 CNN

InceptionV3 [9] is a 48 layer-network, which is pre-trained on ImageNet database for over 1000 classes as shown in Figure4.1(b). This model provides reduction in computational cost by 28% using factorization into smaller convolutions i.e. 5x5 convolution into two 3x3 CNN blocks. Moreover, the factorization into asymmetric convolutions i.e. decomposing 3x3 into 1x3 and 3x1 CNN blocks help in reduction of parameters by 33%.

5. TESTING AND RESULTS

The accuracy graph (orange line-validation, blue-line-training) and cross entropy loss graphs obtained after training of VGG-16 and Inception-3 CNN model is shown in Figure 5.1 and Figure5.2. The numbers of epochs are 10 and 4000 in both models respectively.

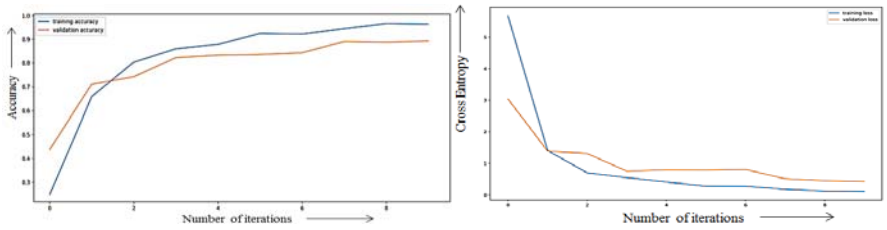


Figure 5.1 (a) Accuracy and (b) Cross Entropy of VGG-16 model v/s epochs

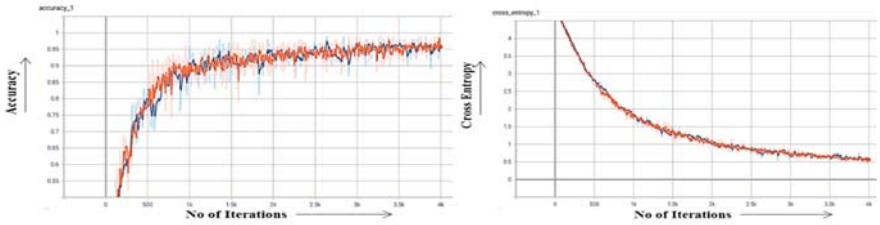


Figure 5.2 (a) Accuracy (b) Cross Entropy of InceptionV3 model v/s epochs

The model training results on test images and real time testing results are depicted in Figure5.3. The top five results are also depicted alongside which shows correct identification of different food items with high accuracy values. Table I depicts the comparison between training accuracy results of previous and the present work.

Figure 5.3 Model testing results on test and real-time images

TABLE 1 TRAINING ACCURACY RESULTS COMPARISON WITH PREVIOUS WORKS

CNN Model	Dataset	Training Accuracy (in %)
VGG-16-Previous Work [7]	FOOD-101	94.02%
InceptionV3-Previous Work [10]	FRUITS360	96.5%
VGG-16	FRUITS360	96.37%
InceptionV3	FRUITS360	98.4%

The testing with Quick Response (QR) and Barcode items are also depicted in Figure5.4.

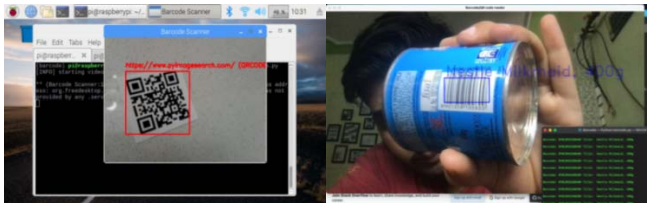


Figure 5.4 QR code and Bar code successful scan results

6. CONCLUSION

The testing of designed algorithm resulted in an excellent accuracy of 96.3% using VGG-16 and 98.4% using InceptionV3 CNN. The successful real time testing results carried using the designed hardware setup justified the performance of the algorithm proposed.

7. REFERENCES

[1] F. Zhu, et. al., 'The use of mobile devices in aiding dietary assessment and evaluation', IEEE J. Sel. Top. Signal Process., vol. 4, no. 4, pp. 756–766, 2010.

[2] H. Hoashi, T. Joutou, K. Yanai, 'Image recognition of 85 food categories by feature fusion', IEEE Int. Symp. Multimedia, pp. 296–301, 2009.

[3] K. Aizawa, Y. Maruyama, H. Li, C. Morikawa, G. C. De Silva, ' Food balance estimation by using personal dietary tendencies in a multimedia food log,' IEEE Trans. Multimed., vol. 15, no. 8, pp. 2176–2185, 2013.

[4] Q. V. Le, 'Building high-level features using large scale unsupervised learning', ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc., pp. 8595–8598, 2013.

[5] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, 'Gradient-based learning applied to document recognition,' Proc. IEEE, vol. 86, no. 11, 1998.

[6] P. Jain, P. Chawla, 'Smart Module Design for Refrigerators based on Inception-V3 CNN Architecture,' Second Int. Conf. on Electon. and Sustainable Comm. Syst., pp. 1852-1859, 2021.

[7] S. Yadav, Alpana, S. Chand, 'Automated Food image Classification using Deep Learning approach,' Int. Conf. Adv. Comput. Commun. Syst., pp. 542–545, 2021.

[8] H. B. Unal, E. Vural, B. K. Savas, Y. Becerikli, 'Fruit Recognition and Classification with Deep Learning Support on Embedded System,' Innov. Intell. Syst. Appl. Conf. ASYU 2020.

[9] C. Szegedy, et al., 'Going deeper with convolutions,' IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, vol. 07-12-June-2015.

[10] Z. Huang, Y. Cao, T. Wang, 'Transfer learning with efficient convolutional neural networks for fruit recognition,' Proc. IEEE 3rd Inf. Technol. Networking, Electron. Autom. Control Conf., pp. 358–362, 2019.

Biographies

Pulkit Jain is currently a research scholar in ECE Department, Chandigarh University. His research areas include image processing, computer vision, and embedded systems.

Paras Chawla is currently working as an Associate Dean Academic Affairs and Professor, ECE Chandigarh University. He received the "Coventor Scholarship award" from MANCEF, New Mexico-USA. He has published more than 70 papers in various reputed National and International Journals/Conferences.

Performance Analysis of PMSM Using a DTC-based ANFIS Controller for Electric Vehicle

Marulasiddappa H B[1], Pushparajesh Viswanathan[2]

[1]*Research Scholar, JAIN - A Deemed to be University, Bengaluru, Karnataka, India*
Email*: marul.bethu@gmail.com
ORCID: 0000-0001-7558-3579
[2]*Associate Professor, JAIN - A Deemed to be University, Bengaluru, Karnataka, India*
v.pushparajesh@jainuniversity.ac.in

Abstract: Now a day's environment pollution is the major concern in designing vehicles. Until now internal combustion engine (ICE) vehicles takes a major part in vehicle manufacturing. Day by day, the use of ICE are deteriorating because of pollution and less fuel availability. In the present scenario, the electric vehicle (EV) plays a major role in the place of an ICE vehicle. The performance of EVs can be improved by the proper selection of electric motors. Initially, EV preferred induction motors for traction purposes, but due to complexity in controlling the induction motor, the permanent magnet synchronous motor (PMSM) is replacing the induction motor in EV due to its advantages. Direct torque control (DTC) is one of the known techniques for PMSM drive in EV to control the torque and speed. However, the presence of torque ripple is the main drawback of this technique. Many control strategies are followed to reduce the torque ripples in PMSM. Adaptive neuro-fuzzy inference system (ANFIS) controller strategy is explained to reduce torque ripples and settling time. Here, the performance parameters like torque, speed and settling time are compared between the conventional proportional-integral (PI) controller and the ANFIS controller.
Keywords: Direct torque control **(**DTC), Electric vehicle (EV), Torque ripple, PMSM

1. Introduction

In the early days, the DC motor was the better choice for electric vehicles because of their simpler control method and its characteristics and hence is most preferable for EV motor. The presence of a commutator and brushes makes the motor more complex for an electric vehicle. To overcome complexity in the structure of the DC motor, induction motors (IM) are used in EV applications [1]. The absence of a commutator and brushes makes it simple in construction. Analysis of IM becomes complex due to its nonlinear characteristics [2]. To overcome IM drawbacks, permanent magnet synchronous motor (PMSM) finds development in EV. PMSM has a higher current density and is more efficient than an induction motor. Hence, it is very much suitable for EV applications. Comparison is made in Table 1 on electric motors with respect to different parameters.

Table 1 Comparison of different motors for electric vehicle [3]

Parameter/motor	PMSM	SRM	IM	DC
Efficiency	High	Low	Low	Low
Power density	High	Low	Low	Low
Controllability	Medium	Low	High	High
Reliability	Medium	High	High	Low

It is better to choose a suitable motor for an electric vehicle depending on its requirements. Brushless DC motors are a better choice for 2-wheelers. PMSM or induction motors are a good choice for four-wheeler vehicles. However, when considering the efficiency and power density, PMSM is a better choice for an electric vehicle. For EV applications, controlling torque over a wide range is necessary. Complexity in controlling a PMSM drive is one of the drawbacks. Scalar and vector controls are the basic control methods for AC drives. Scalar control gives a good steady-state performance. To get highly accurate and good dynamic performance vector control methods are used. The proportional-integral (PI) controller is a commonly applied controller rather than a PID controller, but it leads to instability in the system. Field Oriented Control (FOC) is used to improve steady-state performance and to reduce torque ripples. FOC improves the dynamic response [4], [5] and increases its performance. However, this method is sensitive to small changes in the parameter when temperature changes. Many control strategies are proposed to

improve the performance of PMSM motor. These authors [6] established a new technique called direct torque control (DTC) in place of FOC drawbacks. DTC has the capability to control torque and flux directly by increasing the efficiency of the PMSM drive. The simplicity of DTC has one of the major advantages. The conventional DTC produces more torque and flux ripples and parameters like torque and flux are measurable efficiently [7], [8]. This paper [9] uses to look up table-based vector control for switched reluctance motor to improve the torque response. This paper [10] explains a new strategy called a combination of PI and iterative learning control (PI-ILC) for a repetitive task in a control system. This control technique is used to reduce the speed signal ripples if this motor has repetitive tasks.

This paper [11] establishes a hybrid intelligent controller, which reduces the torque, ripples and improves performance. The author [12] delivered new hybrid technique for dual phase PMSM where he achieved good dynamic response. In this proposed technique [13], inter turn fault produces ripple in the torque, hence torque injection is applied to minimize it. Section II explains the modeling of the PMSM drive. Section III describes the ANFIS controller, section IV explains the results, and finally, section V gives the conclusion.

2. PMSM Mathematical modeling

PMSM modeling is explained by taking rotor as reference. The following equations [9], [10] are expressed for PMSM with respect to rotor reference.

$$\begin{bmatrix} v_{qs}^r \\ v_{ds}^r \end{bmatrix} = \begin{bmatrix} R_s + L_q p & \omega_r L_q \\ -\omega_r L_q & R_s + L_d p \end{bmatrix} \begin{bmatrix} i_{qs}^r \\ i_{ds}^r \end{bmatrix} + \begin{bmatrix} \omega_r \Psi_{af} \\ 0 \end{bmatrix} \quad (1)$$

$$\Psi_{af} = L_q * i_{fr} \quad (2)$$

Where, superscript 'r' represents rotor reference frame, v_{ds}^r = d-axis voltage , i_{ds}^r = d-axis current, v_{qs}^r = q-axis voltage, i_{qs}^r= q-axis current, P = is number of poles, Ψ_{af} = permanent magnet flux linkage, L_m= mutual inductance, L_d = d-axis inductance, R_s = stator resistance, ω_r = rotor speed, L_q = q-axis inductance.

The torque equation of a PMSM is given as:

$$T_e = \frac{3}{2} * \frac{P}{2} (\Psi_{af} * i_{qs}^r + (L_d - L_q) * i_{qs}^r * i_{ds}^r \quad (3)$$

The resultant flux linkage is:

$$\Psi_{res} = \sqrt{((\Psi_{af} + L_d * i_{ds}^r)2 + (L_d * i_{qs}^r 2)} \quad (4)$$

Proposed methodology

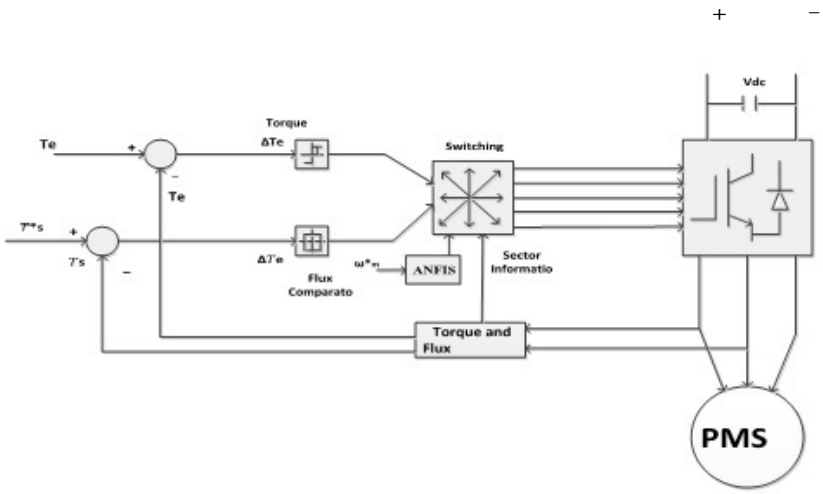


Figure. 1 Block diagram of proposed controller

Figure 1 shows the block diagram of proposed controller. Generally, flux estimation is done by using current modeling, voltage modeling, or by using both. Proper control of the drive system is necessary to determine the proper flux in direct torque control based PMSM. For low-frequency operations, the determination of flux is based on current modeling. To understand this stator current and rotor mechanical position is necessary. Rotor parameters cause more errors in the speed estimation of PMSM. To avoid these errors DTC proposes voltage model to calculate torque and flux by using equations (5)-(7). In the DTC position,sensors are absent and only stator resistance is considered.

$$\Psi_{s\alpha} = \int (V_{s\alpha} - Rs\, i_{s\alpha})dt \tag{5}$$

$$\Psi_{s\beta} = \int (V_{s\beta} - Rs\, i_{s\beta})dt \tag{6}$$

The torque is given by:

$$T_e = \frac{3}{2} * p(\Psi_{s\alpha}\, i_{s\beta} - \Psi_{s\beta}\, i_{s\alpha}) \tag{7}$$

The torque has an input -1, 0, +1 corresponding to flux values 1 and 0. The vectors are selected appropriately whether errors are outside or inside the hysteresis bands. The selected vectors should not make flux and torque errors leave the hysteresis bands. There are six switches and eight voltage vectors. The selection of an appropriate voltage vector improves the overall performance and reduces ripples in the torque and flux.
The following Table 2 gives the lookup table for the PMSM drive.

Table 2 Look-Up Table

Flux error	Torque error	Q1	Q2	Q3	Q4	Q5	Q6
1	1	V2(110)	V3(010)	V4(011)	V5(001)	V6(100)	V1(100)
	0	V0(000)	V7(111)	V0(000)	V7(111)	V0(000)	V7(111)
	-1	V6(100)	V1(100)	V2(110)	V3(010)	V4(011)	V5(001)
0	1	V3(010)	V4(011)	V5(001)	V6(100)	V1(100)	V2(110)
	0	V7(111)	V0(000)	V7(111)	V0(000)	V7(111)	V0(000)
	-1	V5(001)	V6(100)	V1(100)	V2(110)	V3(010)	V4(011)

3. Results

There are two cases studied. Case 1 is a PI controller-based speed control and the second case is an ANFIS-based speed control. The following Figure. 2 gives the characteristic curve of speed in PI controller in DTC where the speed settles at 0.25 secs. The following Figure 3 gives the torque characteristics of PI controller in DTC; here, the maximum peak of the torque goes to 67 Nm. Then, the voltage curve is shown in Figure 4 and the current characteristics are in Figure.5.

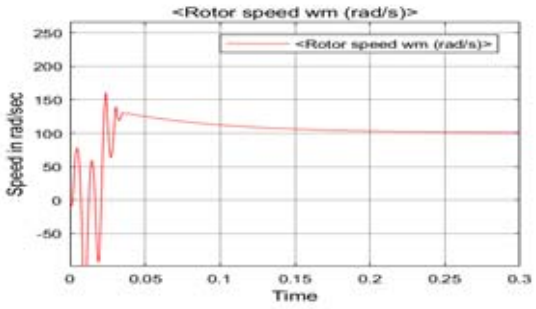


Figure. 2 Speed waveform of PI control method

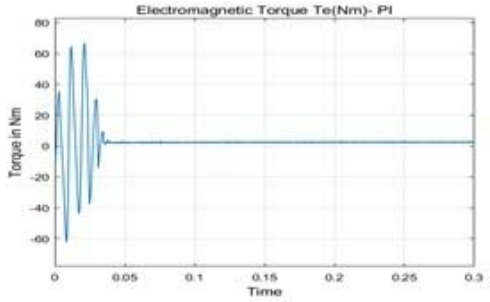


Figure. 3 Torque waveform of PI control method

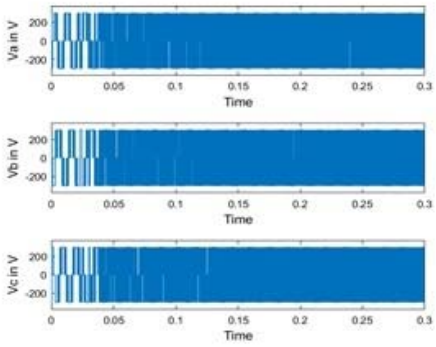


Figure. 4 Voltage curve of PI control method

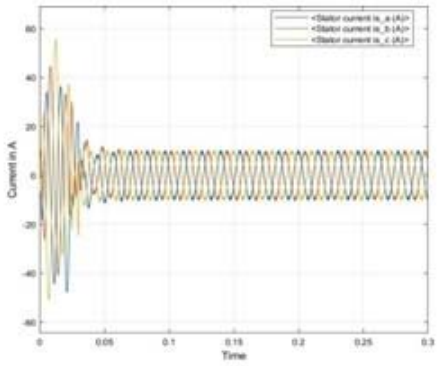


Figure. 5 Current curve of PI control method

Speed behavior of the ANFIS is shown in Figure 6. Using DTC. It shows that the settling time of speed is 0.07 secs, which is faster than the PI controller. The torque behavior of the ANFIS controller is shown in Figure 7 using DTC, here the maximum torque goes to 39 Nm. Figure. 8 shows the current characteristics of the DTC based ANFIS controller. Figure 9 shows the comparison of the speed curve between PI and ANFIS controller in DTC. Then Figure 10 shows the comparison of torque curve between PI and ANFIS controller in DTC. In this comparison, the speed settling is faster in ANFIS andthe torque ripple minimization is happening in ANFIS.

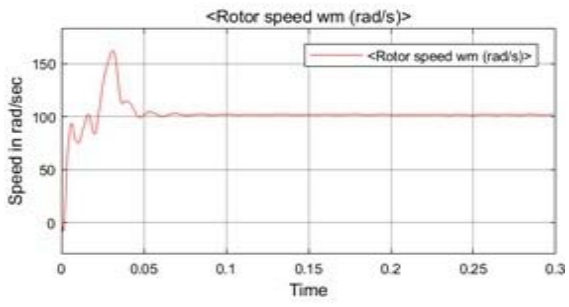


Figure 6 Speed characteristics of ANFIS controller in DTC

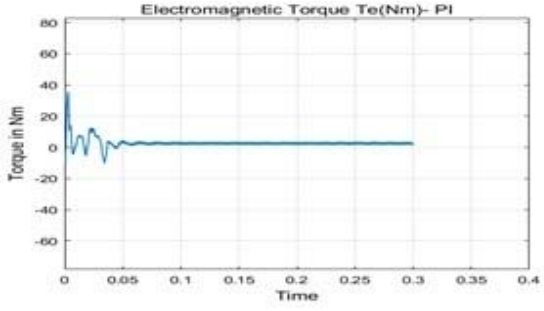


Figure 7 Torque characteristics of ANFIS controller in . DTC

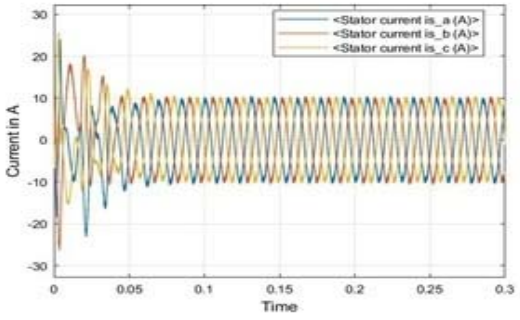


Figure 8 Current characteristics of ANFIS controller in DTC

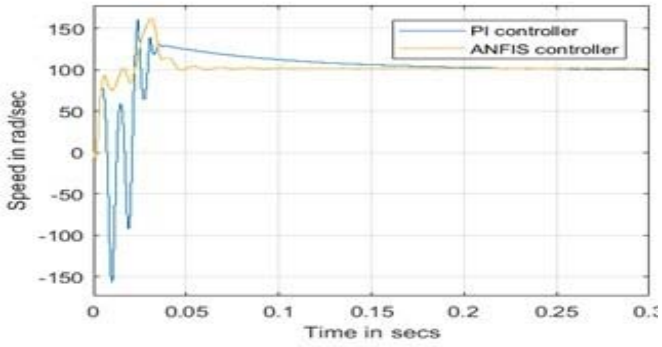


Figure 9 Speed waveforms comparison between two controller in DTC

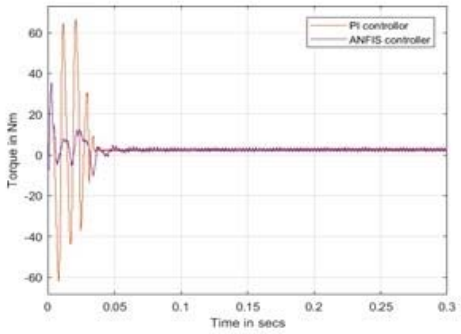


Figure 10 Torque curve comparison between PI and ANFIS controller DTC

Table 3 Comparison of torque values of PI controller and ANFIScontroller

Technique	Torque In N-M			Parameters		
	Maximum value	**Minimum value**	**MEAN**	**σ_τ**	**T_r**	**t_s sec**
PI Controller	65.2	28.570	42.410	8.130	0.720	0.0720
ANFIS	65.64	40.17	54.48	6.53	0.47	0.048

Where, σ_τ = total torque, t_s = Settling time, T_r = torque ripple co-efficient, Max = torque max value, Min = minimum value of torque, Mean = mean value of total torque.

4. Conclusion

This work illustrates the ANFIS-based direct torque controlled PMSM drive system and simulation results are shown for both controllers. With the use of a proper lookup table, torque error is minimized within the described hysteresis band. Table 3 compares the ANFIS controller with the PI controller for different parameters. The torque ripple coefficient of the ANFIS controller is 0.47 which is low compared to 0.72 of the PI controller. Settling time is low in the ANFIS controller than PI controller and hence torque ripples are considerably low in comparison to the PI control method. It can be concluded from the results that, operational parameters of the ANFIS is better than conventional PI control method.

5. References

[1] Tabbache, B., Kheloui, A., & Benbouzid, M. E. H. (2010, September). Design and control of the induction motor propulsion of an electric vehicle. In *2010 IEEE Vehicle Power and Propulsion Conference* (pp. 1-6). IEEE.

[2] Gdaim S, Mtibaa A and Mimouni M "Design and experimental implementation of DTC of an induction machine based on fuzzy logic control on FPGA" IEEE Trans. Fuzzy Sys., vol. 23, no. 3, pp. 644-655, Jun. 2015.

[3] Hashemnia, N., & Asaei, B. (2008, September). Comparative study of using different electric motors in the electric vehicles. In *2008 18th International Conference on Electrical Machines* (pp. 1-5). IEEE.

[4] El Ouanjli, N., Derouich, A., El Ghzizal, A., Chebabhi, A., & Taoussi, M. (2017, November). A comparative study between FOC and DTC control of the Doubly Fed Induction Motor (DFIM). In *2017 International Conference on Electrical and Information Technologies (ICEIT)* (pp. 1-6). IEEE.

[5] Mehazzem, F., Nemmour, A. L., & Reama, A. (2017). Real time implementation of backstepping-multiscalar control to induction motor fed by voltage source inverter. *International Journal of Hydrogen Energy*, *42*(28), 17965-17975.

[6] Takahashi, I., & Noguchi, T. (1986). A new quick-response and high-efficiency control strategy of an induction motor. *IEEE Transactions on Industry applications*, (5), 820-827.

[7] Reza, C. M. F. S., Islam, M. D., & Mekhilef, S. (2014). A review of reliable and energy efficient direct torquecontrolled induction motor drives. *Renewable and Sustainable Energy Reviews*, *37*, 919-932.

[8] Naik, N. V., Panda, A., & Singh, S. P. (2015). A three-level fuzzy-2 DTC of induction motor drive using SVPWM. *IEEE Transactions on Industrial Electronics*, *63*(3), 1467-1479.

[9] Pushparajesh, V., Balamurugan, M., & Ramaiah, N. S. (2019). Artificial Neural Network Based Direct Torque Control of Four Phase Switched Reluctance Motor. *Available at SSRN 3371369*.

[10] Nicola, M., Nicola, C. I., & SacerdoŢIanu, D. (2020, May). Sensorless Control of PMSM using DTC Strategy Based on PI-ILC Law and MRAS Observer. In *2020 International Conference on Development and Application Systems (DAS)* (pp. 38-43). IEEE.

[11] Pushparajesh, V., Nandish, B. M., & Marulasiddappa, H. B. (2021). Hybrid intelligent controller based torque ripple minimization in switched reluctance motor drive. *Bulletin of Electrical Engineering and Informatics*, *10*(3), 1193-1203.

[12] Wang, X., Wang, Z., & Xu, Z. (2018). A hybrid direct torque control scheme for dual three-phase PMSM drives with improved operation performance. *IEEE Transactions on Power Electronics*, *34*(2), 1622-1634

[13] Hang, J., Ding, S., Ren, X., Hu, Q., Huang, Y., Hua, W., & Wang, Q. (2021). Integration of Inter-Turn Fault Diagnosis and Torque Ripple Minimization Control for Direct-Torque-Controlled SPMSM Drive System. *IEEE Transactions on Power Electronics*.

BIOGRAPHIES OF AUTHORS

V. Pushparajesh has been serving as Associate Professor in the in the Department of Electrical and Electronics Engineering ,School of Engineering and Technology ,Jain Deemed to be University since 2017.He has completed his Ph.D from Anna University in Electrical Engineering specialized in Power Electronics and Special Electrical Drives. He has received the International Best Research Award for the year 2018- 2019 instituted by SDF International, London, UK. Adding to his credit he also received national award like Best Academic Researcher, Outstanding faculty award, Best faculty award and Best Placement coordinator from various Research Organization. He has Many National and Internationals journals in his research account.

Marulasiddappa H B is a research scholar from Jain Deemed to be university, Bengaluru Karnataka, India. Research scholar completed his under graduation from Kuvempu university in 2005 in the stream Electrical and Electronics Engineering, and post-graduation in power electronics from VTU in 2011, Belagavi, Karnataka, India. He has 15 years of teaching experience. He published many papers in international conferences and journals.

.

Realtime Mask and Face Detection Using Tensor Flow, Keras and OpenCV

Vaibhav Yadav[1], Ayush Gupta[2], Sanjiv Kumar[3], Prabhishek Singh[4], Manoj Diwakar[5], , Kapil Joshi[6]
[1,2,3,4]Amity School of Engineering and Technology, Amity University Uttar Pradesh, Noida, India
[5]Graphic Era deemed to be University, Dehradun, Uttarakhand, India
[6]Department of CSE, UIT, Uttaranchal University, Dehradun, India
vaibhavyadav477@gmail.com[1], Ayush16gupta20@gmail.com[2], skumar8@amity.edu[3], prabhisheksingh88@gmail.com[4], manoj.diwakar@gmail.com[5], [6]kapilengg0509@gmail.com

Abstract— Due to the COVID19 pandemic, our day-to-dayliveshavebeenchangeddrasticallyandglobaltradeandmovements are severely disrupted. To wear a proper sanitizedprotectivefacemaskhasnowbecomeanecessityforthegeneralpublic.Nowadays,manyof thepublicandprivatecompanieswillrequesttheiremployeestousefacemaskswhile doing their job in the workplace. Face maskdetection system has now become an integral part of our society.ThisResearchPaperfocuses our mind onamoresimplifiedand a strict approach to protectourselvesfromCovid19 pandemic.Byusingmachinelearningtechnologieslike TensorFlow with Keras and OpenCV, and ScikitLearn, wecandetectthat whether a person is wearing his/herfacemask or not.Thismethodcorrectlydetectsthehuman face and then recognizes whether that face is wearing amask or not and even detects whether the mask is over the noseor not. As a surveillance system, the software can detect that the face is wearing a mask or not even in motion. The method isbuilt to achieve accuracy upto 95.77% in two different data sets.WeareusingtheSequentialformofConvolutionalNeuralNetworkmodelinwhichoptimizedvaluesareachievedtodetectthe correctpresenceof themask.

Keywords: Keras,OpenCV,TensorFlow,FacialRecognition,Real-TimeMaskandFaceDetection,CNNarchitecture.

1. Introduction:

According to (WHO) World Health Organization, in the year2019, millions of people were infected by coronavirus diseaseand almost a million deaths were reported due to COVID19.SymptomslikeLossofTasteandBreathingproblemslikeshortness of breath were reported all around the world and it was seen that elderly people with lung disease may havesevere complications. Droplets of cough of the infected personcarrying the virus can reach to the neighboring people, this make the virus communicable.

Our people should know the importance of wearing face mask as a safeguard against

COVID19, toregularize the riskof getting infected froman infected person during the "pre-symptomatic" phase.WHOfocuses that medical masksandventilators are a priority formedicalpersonnel,because of whichmaskscreeninghasnowbecomeanintegraltaskintoday'ssociety.Thismaskdetectionsystem basicallydetectsthepositionofyour face and then finally detects the presence of the mask on yourface. The thing is almost similar to general object recognition to recognize objectclasses.

2. Dataset:

Twodatasetswereusedtoexperimentwiththecurrentmethod:dataset 1 contains 1,376 (one thousand and thirty-six) images,whichincludes690(sixhundredandninety)imagesofrandom people who are wearingthe protective face masks andthe rest 686 (six hundred and eighty) images of people who are not wearing face mask. The first figure primarily includes thefront face pose withthe same typeofmaskwhichisjust thewhitecoloronly.

Kaggle's record 2 consists of 853 (eight hundred and fifty-three) imagesandtheirfacesareerasedwithor withoutmasks.In the figure (Fig-2) there are also face collections that aredesignedforhead-turning,bowing,andbowingwithmorethanonefacepresent inthe sameframe.

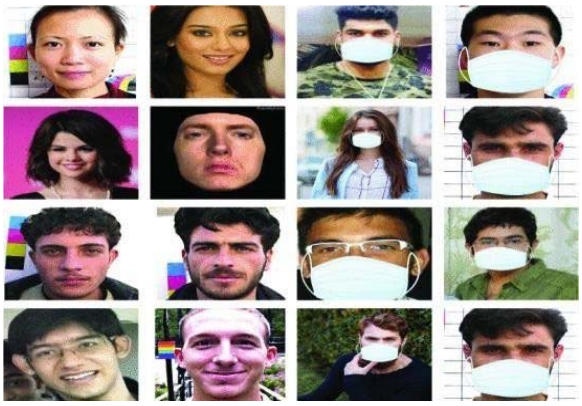


Figure 1: Samples including faces with and without white face masks.

Figure 2: Samplesincludingfaceswithandwithout colored face masks.

3. IncorporatedPackages:

A. TensorFlow:

TensorFlow is defined as an interface that expresses machine learningalgorithms, that includes some special topics like sentimentanalysisthatisusedtocaptureexpressions,speechrecognition for voice, geographic information extraction, visionby computer, tax etc.

B. Keras:

Kerasprovidesbasicbuildingblocksforbuildingandtransporting AA arrays at high iteration rates. It not only has thescalabilitybutitalsohasthemodifiedcross-platformcapabilities. Therefore, the core data structures and APIs whichare included in Keras are the multiple layers and differentmodels of the idea that we have worked upon. The vector isconvertedtoabinaryclassarrayandthemodelisgenerated.

C. OpenCV:

OpenCVis an open-source machine learning tool andsoftware libraryusedtodistinguishandrecognizefacesandobjects,trackprogressive modules, eye movements to track, track cameraactions, red-eye of flash images, search for comparison imagesfromanimagedatabase.

4. TheProposed TechnicalMethod:

The technical methodconsists ofaparticularclassifier that was structured by the algorithmandCNNmodelthatcomprisesofthetwo2Dconvolutional layers which are used to connectthelayers ofdense neurons. This mask recognition system algorithm is asfollows:

A. DataProcessing:

Data preprocessing refers to the process of converting data froma specific format into a kind of format that is more robust anduser-friendly. The data can be of any form like an image or in avideo format or even in plain text. This includes a compositionmodel and analyses the relationship available between differententities and sub properties. The suggested method interacts withOpenCV and NumPy tools that interacts with images and data in videoformat.

a) ConvertingRGBimagetograyimage:

Nowadays, our current imagerecognitionsystems areabletowork with grayscale format of images without even moving into theprocess of converting color images tograyscale images. The main reason is that grayscalecolormethodhas very littleconsequencewhen working withrobustdescriptors.Sincegrayscalestreamlinesthealgorithmandreduces computational requirements, it is used to extract someparticulardescriptorsinsteadofworkingoncolor.

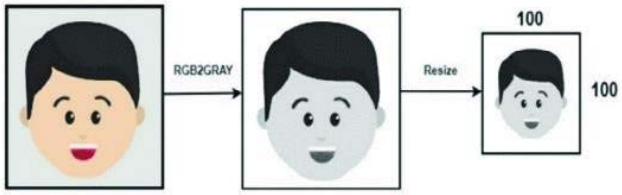


Figure 3: Conversion of RGB image into Grayscale image of 100 x 100 size.

b) ImageReshaping:

The image that we take as input is a three-dimensional system tensor inwhich each channel has a single prominent pixel. The maincondition is that the images must be similar in size to match the3Dfeatureoftensor.It canonlyacceptfinelytunedimages,whichcausesseveral difficultieswithdatacollectionandmodeldeployment.

B. Trainingof the Model:

We have included a layer of Mobile nets in this CNN architecture to improve the response time of the system. This may affect the accuracy of the model but the resulted accuracy is sufficient for our model. This particular modification has let us improve the software exponentially. Now the next step is the splitting of the dataset. This means themodel should be trained on a particular set and then should betestedonanotherdataset.Thelossofvalidationisbeingmonitored with a Model Checkpoint. The images that we areusing to train the system is included in this sequential model.Here, we use 20% of the data that has been already trained as the validation data. The model has been trained for almost 20 epochs, which nowmaintainsa balancebetweenprecisionandthe possibilityofoverfitting.

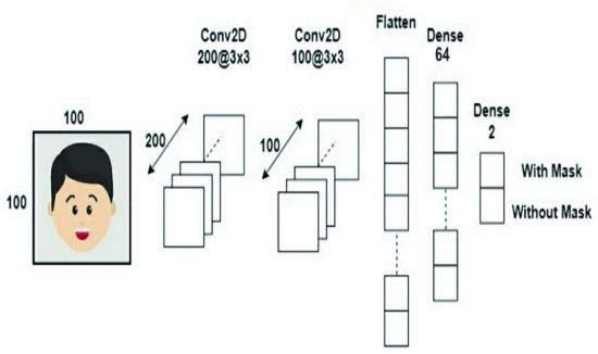


Figure 4: Convolution Neural Network (CNN) architecture

5. WHAT's UNIQUEINOUR RESEARCH PAPER?

A. Facial recognition with and without glasses

There were three independent variables: the glasses and gender of the person being stimulated, and theperson'sprocessingdepth. Theinitialtwoindependentvariableswerevariedthroughphotographs of men and women with or without glasses. Theprocessing depth was manipulated by having the test subjectsassess the characteristics (deep) or the appearance (superficial) of the stimulus subjects during the initial exposure phase [16].

B.ProperwayofwearingFaceMask:

Face mask is the key part in this pandemicsituation asthisistheonlysafeguardthatcanpreventthespreadof this infectious disease. Other errors such as wearing the mask below nose or wearing a loose mask are generally not taken intoaccount but they can also contribute to the spread of the virus.

6. Result and Analysis of the Model:

The face and mask recognition model has been trained, validated, and tested on dataset.The method achieves an accuracy of upto 95.77%. It shows how this optimizedprecision reducesthe costoferrors and can be considered to be amore versatile and detailedMaxPoolingisconsideredtobeoneofthereasons which offers a rudimentary translational. The optimized value of neurons is 64, which is not muchhigh.The larger numbers of neurons can result in poorer performanceofthe model propose.

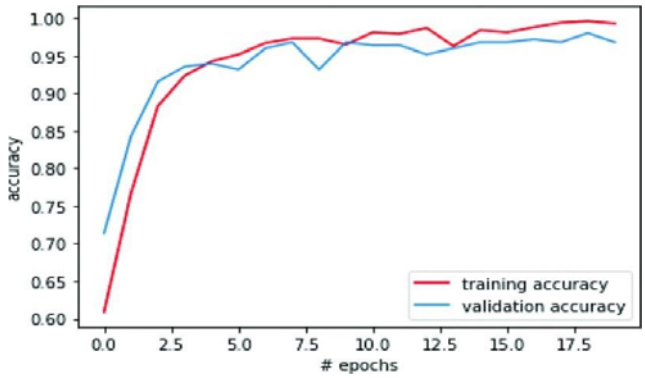


Figure 5: Graph showing - Epochsv/sAccuracyin the reference todataset.

7. Conclusion:

We explained how this technology can be helpfulforsituationslikeCovid 19.Wearing a mask could becomemandatory shortly giventhe Covid19 crisis. The model usedwill make a major contribution to the public health systems. The main motive ofthis model is to set a strict policy for the general public to wearprotective face mask in a proper way so that we can prevent thespread of COVID 19 on a large scale and even put a full stop to it.

8. References:

[1]."Coronavirusdisease2019(covid-19):situationreport205",2020.

[2]. "Coronavirus Disease 2019 (COVID-19) – Symptoms",*Centers for Disease Control and Prevention*, 2020, [online]Available: https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html.

[3]. "Coronavirus—HumanCoronavirusTypes—CDC",2020,[online]Available: Cdc.gov.

[4]. "Advice on the use of masks in the context of COVID-19:interim guidance",2020.

[5]. M.Jiang,X.Fan, andH.Yan,"Retina Mask: AFace Maskdetector",2020,[online]Available:arXiv.org.

[6]. B. Suvarnamukhi and M. Seshashayee, "Big Data Con-concepts and Techniques in Data Processing", *InternationalJournalofComputerSciencesandEngineering*,vol.6,no.10,pp. 712-714,2018.

[7]. Jiri Prinosil, Ondrej Maly, "Detecting Faces with FaceMasks", *TelecommunicationsandSignalProcessing(TSP)202144thInternationalConference on*,pp.259-262,2021.

[8]. Sushil Kumar Mishra, Nitish Kumar Ojha, "A NovelApproachforFaceMaskDetectionusingTensorFlowforCovid-19", *InventiveResearchinComputingApplications(CIRCA)2021ThirdInternationalConference on*,pp.1362-1366, 2021.

[9]. TimoAhonen,AbdenourHadidandMattiPietikaInen,"Facedescription with local binary patterns: Application to FaceRecognition", *IEEE transactions on pattern analysis andMachine Intelligence*,vol.28,no. 12, Dec 2006.

[10]. KZhang,Z.Zhang,Z.LiandY.Qiao,"JointFaceDetection and Allignment using Multi task CascadedConvolution Neural Networks.

[11] Sharma, T., Diwakar, M., Singh, P., Lamba, S., Kumar, P., & Joshi, K. (2021). Emotion Analysis for predicting the emotion labels using Machine Learning approaches. 2021 IEEE 8th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering, UPCON 2021. https://doi.org/10.1109/UPCON52273.2021.9667562

[12] Rebellow, A. M., & Suri, P. (2019). Role of demographic factors on decisionmaking styles of indian corporate executives-public and private sectors. Journal of Management Information and Decision Sciences, 22(3), 308–321. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85075777296&partnerID=40&md5=3a598c0e9f45cf3b8b87c1dae88fd387

[13] .Sajwan, V., & Ranjan, R. (2019). Classifying flowers images by using different classifiers in orange. International Journal of Engineering and Advanced Technology, 8(6 Special Issue 3), 1057–1061. https://doi.org/10.35940/ijeat.F1334.0986S319

[14] Shah, S. K., Negi, V., Singh, S., & Tomar, S. (2013). Design and devlopment of a W-band crossbar mixer at 94 GHz. 2013 IEEE International Conference on Signal Processing, Computing and Control, ISPCC 2013. https://doi.org/10.1109/ISPCC.2013.6663431

[15] Sharma, A. K. (2019). Simulation implementation of three phase VSI fed three phase induction motor drive with filter. International Journal of Engineering and Advanced Technology, 8(6), 1462–1468. https://doi.org/10.35940/ijeat.F8126.088619

[16] Joshi, K., Joshi, N. K., Diwakar, M., Tripathi, A. N., & Gupta, H. (2019). Multi-focus image fusion using non-local mean filtering and stationary wavelet transform. International Journal of Innovative Technology and Exploring Engineering, 9(1), 344–350. https://doi.org/10.35940/ijitee.A4123.119119

A Review on Attack Detection On Software Define Network

Jatin Nimade, Namita Tiwari**, Meenu Chawla**

*Research Scholar, **Professor*
Maulana Azad National Institute of Technology, Bhopal, India
Jatinnimade05@gmail.com,
namitatiwari21@rediffmail.com, chawlam@manit.ac.in

Abstract

To gain cost efficiency and network flexibility, most businesses are converting their traditional networks to Software-Defined Networks (SDN). However, we have seen recent security breaches and attacks against SDN which have shown that technology's security flaw. On a software-defined network, attack detection is an essential part of the infrastructure for security management(SDN). Advancement in machine learning has benefited a variety of fields with also including security.Anomaly-based intrusion detection systems are trained using machine learning approaches to detect even unknown threats. We can undertake a complete review of many various types of research related to ML-based Attack Detection systems using NSL-KDD dataset.

Keywords: NSL-KDD, IDS.

1 Introduction

Based on the detecting mechanism, the IDSs are divided into two categories. Anomaly-based and misuse-based detection are two of these types of detection. The usual type of traffic is modeled in anomaly-based detection frameworks. Any departure from the model is considered an attack. With innovative attacks, this method of detection offers

River Journal, 1–6.

an advantage. Misuse-based detection systems, on the other hand, we have coordinated the particular signatures for threats that have already been detected. There are various types of traffic who are associated with the preset assaults are regarded as normal. Based on the detecting mechanism, Two types of IDS-based detection are anomaly-based and misuse-based detection. The usual type of traffic is modeled in anomaly-based detection frameworks. Any departure from the model is considered an attack. With innovative attacks, this method of detection offers an advantage. Misuse-based detection systems, on the other hand, coordinate specific signatures to pre-detected threats.

Machine Learning (ML) technologies examine large data sets and predict the Variables of interest's future values using mathematical methods. Machine learning approaches are also used to train and assess the Intrusion Detection System(IDS) in the area of cyber security on dataset related to the security. An IDS that has been properly trained can identify unauthorized network activity as well as new assaults, such as zero-day attacks, forecast and detect.

We have examined many works with our related to Attack detection using the Network Security and Data Mining (NSLKDD) data set in this paper.The remainder of this work is divided into the following sections: NSLKDD data set is described in section 2 . We outline the various studies on ML-based IDS and their accuracy of different algorithm in section 3. and finally we have to conclude the paper in section 4.

2 NSL-KDD Data set

.

The NSL-KDD is a more comprehensive version of the KDD dataset that incorporates information from the KDD cup 99 data sets. The collection includes records of internet traffic detected by the simplified intrusion detection network and encountered by the genuine intrusion detection system. Each record in the dataset has 43 features, 41 of which are related to the input traffics and labels such as normal or assault, as well as the input, traffics' violence. In the dataset, there are four main types of assaults.: DoS, Probing, U2R, R2L.

Tavallaee et al. [13] conducted the usage of the KDDCUP'99 data set resulted in a poor anomaly detection evaluation, according to sta-

tistical analysis. They proposed the NSLKDD dataset, on the basis of original KDD dataset, to fix the issues.

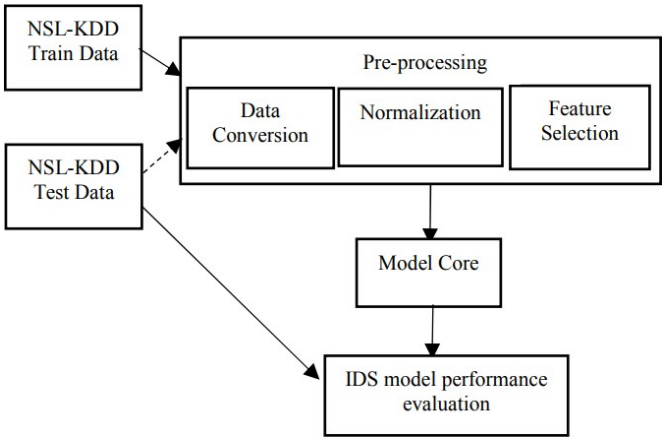


Figure 1 Flow of attack detection research efforts utilising the NSL-KDD dataset.

3 Related Work

In this section review and comparison of different Attack Detection in literature is done. Also tables(3.1) are included in this review which further describes and compares literature studied. 3.1 gives insights to various algorithm and some datsets used in different literature, datasets and compares different Attack Detection classification methods used and achieved accuracy on respective dataset.

3.1 Literature review

In the study, SHAHZEB HAIDER et al., [1] developed We used benchmark ensembles of deep learning and hybrid state-of-the-art methodologies to evaluate a framework for deep CNN ensembles that is both efficient and scalable for detecting the most common and sophisticated DDoS attacks in SDNs, and we put it through its paces on an SDN dataset based on flow. Both detection accuracy and computational complexity have improved with the suggested technique. this DDos attack detection achieved an accuracy level of 99.45

Table 1 Compression Of Different Paper

Ref	Algorithms	Detection	Data sets	Classes	Accuracy
[1]	RNN & RNN RNN & LSTM LSTM & LSTM CNN & CNN	DDos Detction	NSL-KDD	2	RNN-98.16% RL-98.75% LSTM-98.17% CNN-99.45%
[2]	DL-CNN	Attack Detection	NSL-KDD	2	98.43%
[3]	Ensemble voting Multi Tree DNN	Intrusion Detection	NSL-KDD	5	Ensemble voting-85.2% Multi Tree-84.23% DNN-81.61%
[4]	SAE -SVM SAE-SVM	Intrusion Detection	NSL-KDD	2 5	SAE-SVM-99.14% SAE-SVM-80.48%
[5]	RNN	Attack Detection	NSL-KDD	2	RNN-84.56%
[6]	Auto Encoder	Intrusion Detection	NSL-KDD	2	AN-90.61%
[7]	CM-KNN DPTCM-KNN	Anomaly Flow Detection	Real Time Traffic	2	TCM-KNN-92.03% DPTCM- KNN-97.88%
[8]	SVM and K-Means	Attack Detection	KDD-99	2	SVM & K-means - 92.86%
[9]	Multi DT's Algorithm and DNN, KNN Classifier	Attack Detection	NSL-KDD	2	Multi DT's- 85.2%
[10]	NB-SVM	Attack Detection	Real Organization Traffic	2	NB- SVM - 95.96%
[11]	LOA+CNN	DDos	NSL-KDD	2	LOA+CNN -98.2%
[12]	KNN+ELM	Intrusion Detection	NSL-KDD	5	KNN+ELM -84.29%
[13]	CNN	Intrusion Detection	NSL-KDD	5	CNN-79.48%

In study X. Gao et al., [4] presented DT, DNN, and KNN classifiers are used in ensemble voting adaptive algorithms, as well as a ensemble voting adaptive algorithms with DT, DNN, and KNN classifiers and multi-tree approach with multiple Decision Trees (DTs). This IDS is validated using the NSLKDD dataset. The accuracy achieved was 85.2 %.

In P. Pokhrel et al., [9] proposed an Attack Detection that relies It was tested on two firms' traffic after integrating NB and SVM.. On two separate kinds of traffic, this IDS obtained accuracy scores of 95 % and 96 %.

Majd et al., [5] proposed in order to get better the system's overall accuracy, On the basis of flow statistics, a 5-level hybrid categorization system was presented. They use k-Nearest Neighbor technique (kNN) for the first level, and Extreme Learning Machine(ELM) for the second

. In order to avert irreversible harm as a result of a cyberattack, This IDS Acchived accuracy level of 84.29 % on five different classes of the dataset

KEHE WU et al., [6]To handle the problem of an imbalanced data collection, CNN was used to automatically determine the ach class's weight coefficient is based on its numbers in the cost function by extracting traffic attributes from a raw data collection. This IDS Achieved an accuracy level of 79.48 % on five different classes of the dataset.

D. Arivndainabi et al., [7] presented an DDos attack Detection using a loin optimizer algorithm with CNN classifiers. The NSL-KDD dataset was used to validate DDos Attack detection, with an accuracy of 98.2% on two different classes of dataset.

XIANWEI GAO et al., [11] suggested a paradigm for adaptive ensemble learning The model's main idea is to use ensemble learning to combine the benefits of various techniques. To boost the detection effect. It has been demonstrated showing our ensemble model improves detection accuracy effectively when compared to other research articles. The proposed algorithm has an accuracy of 85.2% on five different classes of detaset.

JULIAN JANG-JACCARD et al., [14] present a new 5-layer AE-based model that is superior at detecting abnormal network traffic. Our proposed model's major components and architecture are the outcome of a thorough and thorough investigation into the impact of major AE model performance indicators on detection accuracy.The accuracy of our proposed 5-layer architectural model is unsurpassed. That is 90.16. %

4 Conclusion

In this Study the advent of ML new ideas for Attack Detection in Software Define Network(SDN) has been implemented by various researcher and different form of classification model has been developed.We almost present a review of over 13 papers providing alternative approaches for implementing Attack Detection in this work.For the development of an IDS, a variety of machine learning algorithms were used. With all of the potential combinations of feature selection techniques. The most extensively used datasets in this domain are the KDD99 and NSL-KDD datasets. Ensemble and hybrid classifiers out-

perform single classifiers in terms of efficiency, resulting in a a high rate of detection and forecast accuracy.

References

[1] Haider, S., Akhunzada, A., Mustafa, I., Patel, T., Fernandez, A., Choo, K. & Iqbal, J. A deep CNN ensemble framework for efficient DDoS attack detection in software defined networks. *Ieee Access*. 8 pp. 53972-53983 (2020)

[2] Sallam, Y., Ahmed, H., Saleeb, A., El-Bahnasawy, N. & Abd El-Samie, F. Implementation of Network Attack Detection using Convolutional Neural Network. *2021 International Conference On Electronic Engineering (ICEEM)*. pp. 1-6 (2021)

[3] Peng, H., Sun, Z., Zhao, X., Tan, S. & Sun, Z. A detection method for anomaly flow in software defined network. *IEEE Access*. 6 pp. 27809-27817 (2018)

[4] Gao, X., Shan, C., Hu, C., Niu, Z. & Liu, Z. An adaptive ensemble machine learning model for intrusion detection. *IEEE Access*. 7 pp. 82512-82521 (2019)

[5] Latah, M. & Toker, L. An efficient flow-based multi-level hybrid intrusion detection system for software-defined networks. *CCF Transactions On Networking*. 3, 261-271 (2020)

[6] Wu, K., Chen, Z. & Li, W. A novel intrusion detection model for a massive network using convolutional neural networks. *Ieee Access*. 6 pp. 50850-50859 (2018)

[7] Arivudainambi, D., KA, V. & Chakkaravarthy, S. LION IDS: A meta-heuristics approach to detect DDoS attacks against software-defined networks. *Neural Computing And Applications*. 31, 1491-1501 (2019)

[8] Al-Qatf, M., Lasheng, Y., Al-Habib, M. & Al-Sabahi, K. Deep learning approach combining sparse autoencoder with SVM for network intrusion detection. *IEEE Access*. 6 pp. 52843-52856 (2018)

[9] Pokhrel, R., Pokharel, P. & Timalsina, A. Anomaly-based–intrusion detection system using user profile generated from system logs. *International Journal Of Scientific And Research Publications (IJSRP)*. 9 (2019)

[10] Li, Z., Rios, A., Xu, G. & Trajković, L. Machine learning techniques for classifying network anomalies and intrusions. *2019 IEEE International Symposium On Circuits And Systems (ISCAS)*. pp. 1-5 (2019)

[11] Gao, X., Shan, C., Hu, C., Niu, Z. & Liu, Z. An adaptive ensemble machine learning model for intrusion detection. *IEEE Access*. 7 pp. 82512-82521 (2019)

[12] Ravale, U., Marathe, N. & Padiya, P. Feature selection based hybrid anomaly intrusion detection system using K means and RBF kernel function. *Procedia Computer Science*. 45 pp. 428-435 (2015)

[13] Tavallaee, M., Bagheri, E., Lu, W. & Ghorbani, A. A detailed analysis of the KDD CUP 99 data set. *2009 IEEE Symposium On Computational Intelligence For Security And Defense Applications*. pp. 1-6 (2009)

[14] Xu, W., Jang-Jaccard, J., Singh, A., Wei, Y. & Sabrina, F. Improving performance of autoencoder-based network anomaly detection on nsl-kdd dataset. *IEEE Access*. 9 pp. 140136-140146 (2021)

A Systematic Evaluation on Energy-Efficient Cloud Data Centers with Reduced SLAV

Suraj Singh Panwar[1], M.M.S.Rauthan[2], Arti Rana[3], Varun Barthwal[4]

[1,2,4]Department of Computer Science & Engineering, HNB Garhwal University

[3]Department of Computer Science & Engineering, Uttaranchal University

[1]surajpanwar1977@gmail.comm

[2]mms_rauthan@redifmail.com

[3]artirana11190@gmail.com

[4]varuncsed1@gmail.com

Abstract.

In cloud computing, cloud data centers are looking for ways to improve the services they offer to their customers. So, with the exponential rise of cloud computing applications, data centers are becoming more energy-intensive, capable of delivering performance assurance, and service level agreements. It offers instant services to computer resources, pay-as-you-go services through a global network of data center sites. Cloud servers are consuming a lot of electricity and leaving a lot of carbon footprint on the environment. The goal of cloud computing services is to deliver seamless services via the Internet utilizing virtualization technologies to meet the QoS (Quality of Service) demands of end-users. This review article focuses to identify the research done in Energy Consumption (EC), Quality of Services, and Service Level Agreement Violation (SLAV) by different techniques using statistical, machine learning, heuristic, and metaheuristic method. Host CPU utilization prediction, Underused or Overused detection, VM selection and migration, and VM placement were applied to manage the resources. Researchers have used different algorithms using the PlanetLab and Bitbrains workload and had performed experiments on different environments i.e., CloudSim, Matlab, Java, OpenStack, etc and the work done basis on performance metrics are evaluated.

Keywords. Cloud Computing, Energy Consumption, SLAV, Quality of Service, Data Center, Performance Assurance.

1. INTRODUCTION

The industry has used the term "Cloud Computing" to describe a powerful service and application that is integrated and packaged on the web [1]. It has developed as a novice flexible, efficient and powerful computational technology that offers customized, reliable, and dynamic computing environments to users. Remote data centers host, high-capacity servers and storage systems, which are used to host cloud applications. Rapid growth in requests for cloud-based facilities necessarily involves the development of bulky amounts of data centers, consuming huge amounts of electrical power [2].

Optimizing energy can be accomplished by uniting resources based on current utilization, well-organized virtual network structures, and the thermal position of computing equipment and nodes. An approach to cost-effective VM migration based on erratic electricity prices significantly decreases the energy costs associated with the cloud service operation.

A quick introduction to cloud computing has been given in section I. In Section II definition of cloud computing, Virtualization, Energy Consumption, service level agreement, VM migration, CloudSim, and Workload data is defined. Section III provides the importance of the paper. Section IV outlines the analysis, and evaluation of existing relevant work for the resource management, performance, and comparisons of approaches with their benchmark algorithm considering energy consumption, and SLAV. Conclusion of the paper is summarised in section V.

2. CLOUD COMPUTING

Cloud computing offers computational resources as well as on-demand resources, scalability, flexibility, and greater availability to consumers anywhere and at any time by utilizing virtualization technology [3]. It is distributed in nature and is a group of independent resources which are spread in remote locations. National Institute of Standards and Technology (NIST) has a widely recognized definition, “Cloud Computing is a model for providing ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and issued with minimal management effort or interaction from service providers" [4].

Cloud computing is based on three service models IaaS (Infrastructure as a Service), PaaS (Platform as a Platform), and SaaS (Software as a Service). Cloud computing delivers IT infrastructure and utilities depending on a service model. Based on the their deployment, clouds are categorized as: Public, Private, Community and Hybrid Cloud[5].

A. Virtualization

Virtualization technology allows many separate operating systems, software, and applications to operate on a single Host, allowing massive data centers to be handled more effectively, dividing hardware resources into logical units called virtual machines (VMs). VMware, Xen, and KVM are commonly used virtual environments in cloud DCs [6].

B. Energy Consumption, Service Level Agreement and VM Migration

Cloud computing may be a step forward in analyzing, recognizing and implementing global energy reductions in a system to achieve true energy-efficient computing services while lowering costs [7]. Uniting hardware and reducing repetition can help us to save energy. Services should be able to be moved to other locations, if necessary, as well as virtualized and managed within a data center. Machine-readable accounting of the demands and characteristics of applications, networks, servers, or even entire locations must be accessible to enable energy efficiency in the future [8]. The most significant challenges are task scheduling, load balancing, Quality of Service (QoS), resource management, energy efficiency, and service level agreement (SLA). An effective management is required to optimize power utilization and maximize performance. In

addition to energy consumption and SLAV, financial cost and CO2 emission from data centers, cooling systems also affects the environment significantly [9].

Virtual machine placement and migration have a significant influence on QoS and energy consumption. VMs are moved to different hosts, when existing host is either overburdened or underburdened based on power conservation considerations. VM allocation procedure will be in control of selecting suitable hosts for the designated VMs. Virtual Machine Migration is divided into three stages [10].

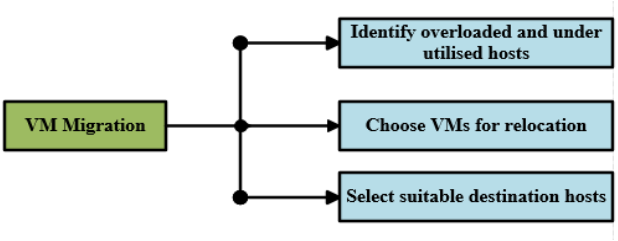


Figure 1: Stages of VM Migration

Computer resources are distributed to VMs constructed for users. To advance the QoS and management of power, VM scheduling has been done in such a way so that minimum numbers of Hosts are running. This method is known as Dynamic Consolidation of Virtual Machine (DCVM). Predicting host utilization is an ongoing research work and a variety of solutions have been proposed. A single Host can host several VMs, and VMs utilize the resources of hosts based on user requests. VM migration technique is suitable to control power consumption. To shrink the VM migrations, appropriate VM selection and VM placement methods must be created. When a VM from an overloaded host is in migration, then both source and destination host consume power without delivering any services, which should be minimized without affecting QoS and hence reducing energy utilization.

C. *CloudSim and Workload Data*

CloudSim [11] is the preferable simulation tool for research where workload data is used for input to test the algorithm. PlanetLab data given in table 1 and Bitbrains workload traces given in table 2 are used frequently for the simulation in experimental environment[12].

3. PURPOSE OF SURVEY

Cloud DCs are the backbone of cloud computing, consisting of networked computers, cables, power supplies, and other components that host and store data. Cloud Data centers demand a lot of energy, which results in significant operational expenses and carbon emissions. The main challenge is to establish a balance between system performance and energy consumption [13]. In this survey, balance between energy efficiency, SLA assurance, VM placement, VM selection and migrations are evaluated performed by different researchers [14].

4. RELATED WORK

Many researchers have applied different techniques for VM management and effective resource management approach to reduce energy utilization, and SLA Violation in cloud computing data centers. Details of their research work are described below.

Beloglazov and Buyya, (2010) [15] used Single Threshold (ST), Minimization of Migration (MM) procedures, and bin packing heuristics for VM consolidation. Random data and CloudSim simulator were used for VM allocation and live migration in an energy-aware resource scheduling system. In terms of energy conservation, the findings showed that dynamic VM consolidation with adaptive criteria is far superior to static thresholds. The MM algorithm outperformed ST, DVFS, and NPA by 23 %, 66 %, and 83 %, respectively.

Z. Cao et al. (2012) [16] proposed strategies for dynamically consolidating VMs (DCVM) to lower energy consumption and SLAV. Host overburden detection, VM selection, and allocation strategy were proposed by the authors. The authors employed standard deviation and mean CPU utilisation metrics to discover overloaded hosts. The Extension of Maximum Correlation (MCE) policy was utilised to select VMs for migration. Mean and variance computations were done for VM allocation on CloudSim utilising Planet Lab traces.

Farahnakian et al. (2013) [17] presented a DCVM in which the active number of hosts is minimised based on current and historical usage. The K-Nearest Neighbour (KNN) approach is used to forecast each Host's CPU use. The goal of their prediction approach is to detect Host overloading and underutilization to optimise dynamic VM consolidation.

A. Nadjar et al. (2015) [18] illustrates a decentralised scheduling strategy for DCVM fitted with an Auto-Regressive Integrated Moving Average (ARIMA) method to progress distribution of resources by predicting VM resource utilisation in future. In their model three types of controller Global Manager (GM) apply First Fit Decreasing (FFD), Cluster Manager (CM) performs Max Load VMP (ML-VMP) and Local Manager (LM) used ARIMA model.

Duggan et al. (2016) [19] presented an energy-efficient solution based on Reinforcement Learning (RL) to choose VMs for relocation from overburdened hosts using local regression approach. The Learning agent selects the best VM for migrating from an overloaded host while balancing energy consumption and migration. The experiments indicate an autonomous VM selection approach that can deduct energy cost and VM migration count.

M.A. Khoshkholghi et al. (2017) [20] projected a cost-effective energy-efficient supervision of VMs by developing an algorithm for overloading the detection of host using the iterative weighted linear regression (IWLR), which consider SLA constraints for data centers.

Wang and Tianfield (2018) [21] by introducing Space-Aware Best Fit Decreasing (SABFD) technique for VM placement, researchers focused on using efficient energy utilization by DCVM. They also created a policy called High CPU Utilization-based Migration VM Selection (HS) for VM selection. Using the CloudSim andPlanet Lab

workload, the results showed that DCVM included SABFD and HS produced the greatest results.

A. Aryania et al. (2018) [22] proposed an Ant Colony System (ACS) method for resolving the Virtual Machine Consolidation (VMC) challenge and reduces data center energy consumption. Simulation findings showed that Energy-aware Virtual Machine Consolidation (EVMC-ACS) enlarged the number of sleeping host by 16 % as compared to ACS-VMC. Furthermore, in various circumstances, the suggested algorithm minimizes the number of migrations by 89 %, the energy utilisation during relocation by 91 %, the amount of SLAV by 79 %, and total energy utilization by 25 %when compared to ACS-VMC.

F.F. Moges et al. (2019) [23] proposed the OpenStack Neat framework's VM placement algorithm to address the problem of consolidation by modifying bin-packing heuristics to account for host power efficiency. The proposed algorithms improve energy proficiency up to 67 % with reference to PABFD and MBFD. They also defined an innovative bin-packing method termed as medium-fit power-efficient decreasing (MFPED) which reduces SLAV and VM relocations by up to 78 % and 46 %, respectively, when compared to MBFD, depending on the cloud scenariousing CloudSim on PlanetLab and Bitbrains cloud traces.

Mehran Tarahomi et al. (2020) [24] proposed a useful micro-genetic approach for selecting appropriate VM destinations among physical hosts. In comparison to other ways, their simulations revealed that the micro-genetic method gives significant enhancements in relations of power consumption. They used CloudSim to test the suggested approach and relate it to reference algorithms (genetic and Power-Aware BFD VM provision algorithms) in a variety of scenarios and PlanetLab datasets that improved power consumption.

V. Barthwal et al. (2021) [25] projected a VM placement process built on ACO meta-heuristic Predicted Utilization (AntPu) for dynamically placing VMs in the data center to minimise Energy Utilization (EU) and SLAV. Using CloudSim and the PlanetLab dataset the projected method offers a significant enhancement in energy usage and service level agreement comparing with other alternatives.

Garg et al. (2021) [26] approached LATHR (Load Aware three-gear THReshold) and MBFD algorithms to reduce overall energy utilization improving service quality. When used with a dynamic workload and a flexible number of virtual machines (range from 1–290) on each host, produces promising results. The proposed technique was found to reduce SLA defilements (26 %, 55 %, and 39 %) as well as energy utilization (12 %, 17 %, and 6 %) when related to inter quartile range, median absolute deviation, and double threshold overload recognition strategies, respectively.

Most of the Researchers have used PlanetLab (table 1) or Bitbrains (table 2) workload traces for simulation in CloudSim, or other environments, whose number of VMs are given below.

Date	***No. of VMs***	***Date***	***No. of VMs***
03/03/2011	**1052**	*03/04/2011*	**1463**

Table 1: Number of VMs in PlanetLab workloads traces [24]

06/03/2011	**898**	*09/04/2011*	**1358**
09/03/2011	**1061**	*11/04/2011*	**1233**
22/03/2011	**1516**	*12/04/2011*	**1054**
25/03/2011	**1078**	*20/04/2011*	**1033**

Table 2: Number of VMs in Bitbrains workloads traces [24]

Date	*No. of VMs*	*Date*	*No. of VMs*
01/08/2013	**1238**	*06/08/2013*	**1231**
02/08/2013	**1237**	*07/08/2013*	**1218**
03/08/2013	**1234**	*08/08/2013*	**1209**
04/08/2013	**1233**	*09/08/2013*	**1207**
05/08/2013	**1232**	*10/08/2013*	**1205**

A brief description of the above literature review, algorithms developed using different methods, comparison of Energy Consumption and SLAV reduction with Benchmark algorithm, workload data, experimental environment, and performance metrics is given in Table 3. The authors have already discussed setting of implementation; the host, virtual machine, datasets, simulators, and other criteria for comparing the proposed method to their benchmark algorithm.

Table 3: Management of Resources using different Methods in Cloud Data Center

S. No.	*Author/ Year*	*Algorithm / Method*	*Benchmark Algo*	*Reduction in EC%*	*Reduction in SLAV%*	*Data set/ Workload*	*Experiment Environment*	*Work Done*	*Performance Metrics*
1.	A. Beloglazav et al. [15], 2010	MM Bin Packing	ST	23	1.1	Random data	CloudSim	VM Consolidation	Energy Savings, SLAV
2.	Z. Cao et al. [16], 2012	EV_MCE	LR_MMT_1.2	-13	70	PlanetLab	CloudSim	Host Overload Detection, VM Selection	Energy Consumption, QoS, SLAV, VM Migrations
3.	F. Farahnakian et al. [17], 2013	DC-KNN	THR	9.2	19.9	PlanetLab	CloudSim	Utilization Prediction	SLAV, Energy Consumption
4.	A. Nadjar et al. [18], 2015	ARIMA MSV_ML	MAD_MMT_2.5	5.4	90	PlanetLab	CloudSim	DCVM	Energy Savings, Migration count, SLA
5.	M. Duggan et al. [19], 2016	AI tech RL RLLM	Lr-Mmt	3	48	PlanetLab	CloudSim	VM Migration	Energy Consumption, VM Migration, SLAV

6.	M A Khoshkholghi et al. [20], 2017	IWLR	LR_MMT	26	85	PlanetLab	CloudSim	Utilization Prediction, VMConsolidation	Energy Utilization, VM Relocation, SLAV
7.	H. Wang et al. [21], 2018	SABFD HS	DVMC	72	69	PlanetLab	CloudSim	VMPlacement VM Migration	Energy Efficient
8.	A. Aryania et al. [22], 2018	EVMC-ACS	ACS-VMC	25	79	Random Workload	Java	VM Consolidation	Energy Consumption, SLAV
9.	F F Moges et al. [23], 2019	MFPED	MBFD	67	78	Planet Lab Bitbrains	CloudSim	VMPlacement	Energy Consumption, SLAV, VMMigrations
10.	M. Tarahomi et al. [24], 2020	Micro-GA	GA	7.68	2.44	PlanetLab	CloudSim	VMPlacement	Power Consumption
11.	V. Barthwal et al. [25], 2021	AntPu	PABFD	24.59	53.19	PlanetLab	CloudSim	VMPlacement PM Overloading	Energy Consumption, SLA Violations, QoS
12.	V. Garg et al. [26], 2021	LATHR - MBFD	IQR	12	26	HPG4 100 hosts 290 VM	Matlab	Overload Detection, VM Migration	Energy Consumption, QoS, IER, No. of Migration

5. CONCLUSION

Since cloud data centers consume a lot of power, produce a lot of invoices, and are very expensive. So, energy consumption, and SLA violations have become major concerns in the previous decade. As a result, energy conservation and QoS are focused and are the most important study fields in Cloud computing.Many academia's are focusing on data center infrastructures' energy utilization, SLAV, and QoS. This article looks at virtual machine and physical machine consolidation strategies using various methodologies in the context of energy-efficient VM migration and SLA assurance. The approaches discussed by different researchers using many methods focusses at global energy conservation, effective VM migration, SLA assurance, and resource management. When comparing with their benchmark method we have seen that energy consumption and SLAV decreases with considerable percentages which has been given in table 3. So, the future scope of reducing these parameters is still in research of cloud computing. Effective VM consolidation techniques reduce power consumption during service provisioning. They also proposed SLA and Energy Aware policy for real-time VM consolidations in Cloud data centers to optimize the power and guarantee the service performance. As a result, resource utilisation increases, while total energy consumption, and SLAV decreases in cloud data centers.

6. REFERENCES

[1] Weiss, Aaron, Computing in the Clouds, NetWorker, 11(4), pages- 16-25, ACM Press, New York, USA, Dec. 2007.

[2] Luiz Andre Barroso, Urs Holzle, and Parthasarathy Ranganathan, The Datacenter as a Computer: Designing Warehouse-Scale Machines, Third Edition, 2018.

[3] Chaurasia et al. Comprehensive survey on energy aware server consolidation techniques in cloud computing, The Journal of Supercomputing, 2021.

[4] Mell, P. and Grance, T., The NIST definition of cloud computing, 2011.
[5] Sotomayor, B., Montero, R.S., Llorente I.M., and Foster I., Virtual infrastructure management in private and hybrid clouds, IEEE Internet Computing, 13(5), Pages- 14-22, 2009.
[6] Varrette, S., Guzek M., Plugaru V., Besseron X. and Bouvry P.,Hpc performance and energy-efficiency of Xen, KVM and VMWare hypervisors, 25th International Symposium on Computer Architecture and High-Performance Computing, 2013.
[7] Gelenbe, E., Steps toward self-aware networks, Communications of the ACM, 52(7), Pages - 66-75, 2009.
[8] Berl, A., Gelenbe E., Girolama M., Giuliani G., Meer H., Dang M.Q. and Pentikousis K., Energy-efficient cloud computing, The computer journal, 2010, 53(7), Pages - 1045-1051.
[9] Buyya, R., J. Broberg, and A.M. Goscinski, Cloud computing: Principles and paradigms, Vol. 87, John Wiley & Sons, 2010.
[10] Ruan, X. and Chen. H., Performance-to-power ratio aware virtual machine (VM) allocation in energy-efficient clouds, IEEE International Conference on Cluster Computing, Pages - 264-273, 2015.
[11] Buyya, R., R. Ranjan, and R.N. Calheiros, Modeling and simulation of scalable Cloud computing environments and the CloudSim toolkit: Challenges and opportunities, in 2009 international conference on high-performance computing & simulation, 2009.
[12] Park, K., and V.S. Pai, CoMon: a mostly-scalable monitoring system for PlanetLab, ACM SIGOPS Operating Systems Review, 40(1), Pages - 65-74, 2006.
[13] Buyya, R., A. Beloglazov, and J. Abawajy, Energy-efficient management of data center resources for cloud computing: a vision, architectural elements, and open challenges, 2010.
[14] Kaur, T. and I. Chana, Energy efficiency techniques in cloud computing: A survey and taxonomy, ACM computing surveys (CSUR), 48(2), Pages - 1-46, 2015.
[15] Beloglazov, A. and Buyya R., Energy-efficient resource management in virtualized cloud data centers,10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing, 2010.
[16] Cao, Z. and Dong, S., Dynamic VM consolidation for energy-aware and SLA violation reduction in cloud computing, IEEE 13th International Conference on Parallel and Distributed Computing, Applications and Technologies, 2012.
[17] Farahnakian, F., et al., Energy aware consolidation algorithm based on k-nearest neighbor regression for cloud data centers, Department of IT, University of Turku, Finland, IEEE/ACM 6th International Conference on Utility and Cloud Computing, 2013.
[18] Nadjar, A., AbrishamiS., and DeldariH.,Hierarchical VM scheduling to improve energy and performance efficiency in IaaS Cloud data centers, 5th International Conference on Computer and Knowledge Engineering (ICCKE), 2015.
[19] Duggan, M., et al., A reinforcement learning approach for the scheduling of live migration from underutilised hosts. Memetic Computing, 9(4), Pages - 283-293, 2017.
[20] Khoshkholghi, M.A., et al., Energy-efficient algorithms for dynamic virtual machine consolidation in cloud data centers, IEEE Access, Vol-5, Pages - 10709-10722, 2017.
[21] Wang, H. and Tianfield, H., Energy-aware dynamic virtual machine consolidation for cloud datacenters, IEEE Access, Vol.-6, Pages - 15259-15273, 2018.

[22] Aryania, A., Aghdasi, H.S.and Khanli, L.M., Energy-aware virtual machine consolidation algorithm based on ant colony system, Journal of Grid Computing, 16(3), Pages - 477-491, 2018.

[23] Moges, F.F. and Abebe, S.L., Energy-aware VM placement algorithms for the OpenStack Neat consolidation framework, Journal of Cloud Computing, 8(1), Pages - 1-14, 2019.

[24] Tarahomi, M., Izadi, M. and Ghobaei-Arani, M., An efficient power-aware VM allocation mechanism in cloud data centers: a micro genetic-based approach, Cluster Computing, 24(2), Pages - 919-934, 2020.

[25] Barthwal, V. and Rauthan, M.M.S., AntPu: a meta-heuristic approach for energy-efficient and SLA aware management of virtual machines in cloud computing, Memetic Computing, 13(1), Pages - 91-110, 2021.

[26] Garg, V. and Jindal, B., Energy-efficient virtual machine migration approach with SLA conservation in cloud computing, Journal of Central South University, 28(3), Pages - 760-770, 2021.

Biographies

Suraj Singh Panwar received the B.Tech. degree in Computer Science and Engineering from Govind Ballabh Pant Engineering College Gurdhauri, Pauri in 2002, M.Tech. degree in Computer Science and Engineering from Uttarakhand Technical University, Dehradun in 2013, and currently pursuing Research in the Department of Computer Science and Engineering, H.N.B. Garhwal University, a Central University, Uttarakhand. He has worked as a Faculty of Computer Science at the Department of Computer Science and Engineering, School of Engineering and Technology, H.N.B. Garhwal University. His research areas include Machine Learning, Fractals, and Cloud Computing.

ManMohan Singh Rauthan is working as a Professor in Department of Computer Science and Engineering, H.N.B. Garhwal University, Srinagar Garhwal, Uttarakhand, India. He completed M.Sc. (Physics) from I.I.T., Delhi. He received M.Tech. Degree in Computer Science from Roorkee University. He did his Ph.D. in Information Technology from Kumaon University, Uttarakhand, India. He worked at Govt. Computer Center, Bhopal as a Programmer. He also worked at National

Informatics Centre, Delhi as Sr. Systems Analyst for more than 6 years. His research interest includes information retrieval and cloud computing.

Arti Ranais currently working as an Assistant Professor in Department of Computer Science & Engineering, Uttaranchal University. She received B.Tech Degree in Information Technology from Uttarakhand Technical University. She did her M.Tech in Computer Science & Engineering from Faculty of Technology, Uttarakhand Technical University. Currently she is pursuing PhD in Computer Science & Engineering from Uttarakhand Technical University. Her research areas include Machine Learning, Artificial Intelligence, Datawarehouse & data mining.

Varun Barthwal is a Research Scholar in the Department of Computer Science and Engineering at the H.N.B. Garhwal University, Uttarakhand, India. He completed B.E. Degree in Computer Science and Engineering from G B Pant Engineering College, Uttarakhand, India in 2007. He received M.E. degree in Computer Technology and Application from Delhi College of Engineering, New Delhi, India, in 2009. His research interests focus on energy efficiency in cloud computing.

ARTIFICIAL INTELLIGENCE ADVANCES IN DRUG DISCOVERY

[1]Devesh Maithani, [1]Abhishek Chauhan, [2]Himani Maheshwari, [2]Amarjeeet Rawat

[1]MCA Student, Department of Computer Applications, Uttaranchal University, Dehradun, Uttarakhand, India

[2]Assistant Professor, Department of Computer Applications, Uttaranchal University, Dehradun, Uttarakhand, India

Email: himanimaheshwari@uttaranchaluniversity.ac.in, amarjeet@uttaranchaluniversity.ac.in

ABSTRACT

In today's world, artificial intelligence has become the most important field of computer science. Machine learning (ML) is an artificial intelligence (AI) field in which we use a training dataset to forecast future outcomes. Machine learning is being used in a wide range of fields, banking, insurance, medical imaging, stock and also including healthcare. Without machine learning, drug discovery takes a long time, costs a lot of money, and has a lower success rate. In recent year's Decision tree, SVM, Naive Bayesian, ANN, and Random Forest (RF) algorithms are some of the Machine Learning applications or algorithms utilised in the field of discovery of drug.Machine learning has become a significant tool in the drug development field as a result of various applications or algorithms.

Keywords: Drug Discovery, Machine Learning, Artificial Intelligence, healthcare

1. INTRODUCTION

Healthcare is one of the applications where machine learning performs well and produces good results in a short period of time. Because high-throughput screening (HTS) procedures were widely used in the pharmaceutical sector in the 1990s, but it requires a significant amount of time and cost to screen and pick potential candidates, time is of the essence (hit compounds)[1].Machine learning may be the most effective method for discovering new pharmaceuticals, as it increases productivity and reduces the time spent searching for and implementing novel drugs. Drug discovery [2] in clinical trials is a lengthy process that takes 10-12 years to complete [3]. According to the Eastern Research [4] Group (ERG), a time window of 10-15 years is required for the development of a drug, with an average cost of US$ 2-3 billion and a success rate of 2.01% [5], so we can see why we need technology that provides the highest success rate in the shortest amount of time. Machine learning, a subset of AI (Artificial Intelligence), encompasses all capabilities, allowing us to create modules that outperform others.

To use machine learning to discover medications, one must first create a knowledge bridge between the drug and the ailment. Both traditional approaches and statistics can be used to build correlations between diseases and drugs. Identifying drug correlations using traditional methods is time-consuming, difficult, and costly.

Multidisciplinary machine learning integration approaches and procedures appear to be beneficial in drug development, thanks to advances in computer intelligence and data collecting. As a result, computational tools must be used to develop and study drug-disease relationships.

2. BACKGROUND

In the pharmaceutical industry, drug discovery, or the process of discovering novel medications, is critical. Finding novel medications is still an extremely expensive and time-consuming procedure at this point, involving Phases I, II, and III clinical trials [4]. Machine learning approaches in Artificial Intelligence (AI), particularly in-depth learning techniques, have lately become widely used and achieve high-quality performance in a variety of domains, including speech recognition, picture classification, bioinformatics, and others.The field of drug development is a most vital uses of modern AI approaches.i.e drug discovery. According to a survey, different patterns can be found in machine learning and drug discovery fields [4].

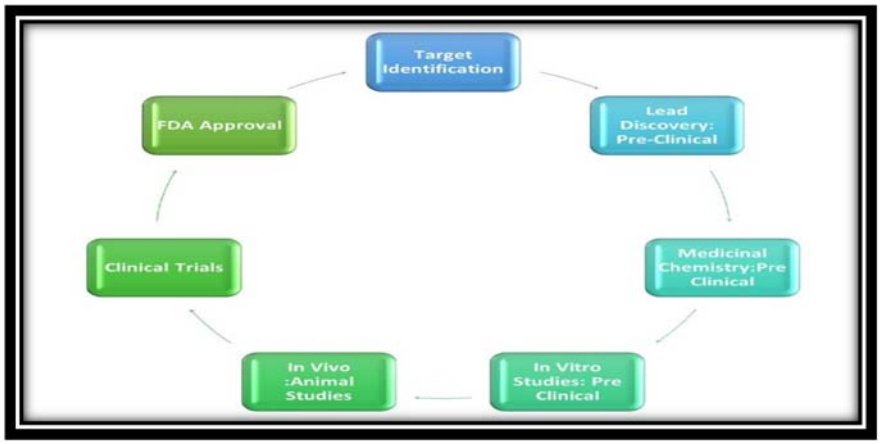


Fig 2.1: Phases in Drug Discovery

2.1. Target Identification: Identifying the target is the initial step in medication development. Understanding the objectives is more important than the medicine in this phase. Though the most likely strategy to develop a drug is to first detect the target and then concentrate on drug development, it is sometimes difficult for humans to identify every possible compound combination. This process can take up to two years for most medications.

2.2. Lead Discovery: pre-clinical: Exploring thousands of compounds that can disrupt the disease's intentions and reducing the number of possible combinations that can work on the target in the second step. This cycle usually takes 1 to 2 years.

2.3. Medicinal chemistry: Pre-clinical: The reduced chemicals are re-evaluated at this point in order to determine the interactions that caused the sickness. The 3D configuration of the compounds, as well as the targeted interactions, were used in further investigation. The results of the analysis are added to the target. This stage takes 1 to 2 years as well.

2.4. In vitro Studies: Pre-Clinical: The cell system is used to test compounds that have been screened up to this point. The phase in which petri dish investigations take place is known as in vitro studies. The effectiveness of the medicine is evaluated at this step by looking at the combinations that interfere with the target.

2.5. In vivo Studies:Animal Studies: Composites that pass the in Vitro-phase are taken and experimented on animals such as rats in this step. The results gained from these animal experiments are fairly representative when evaluated to the 2D in vitro-cell structure-models. The failure rate in its class is significant because to changes in the structure of the animal model in animals, and the results obtained in vitro may not be comparable to those obtained in vivo.

2.6. Clinical Trails: In this stage, a combination of promising traits was identified, and clinical trials were initiated. Human volunteers are used as test subjects at this level.

2.7. FDA Approval: Compounds that fall outside of all of these categories must be approved by the FDA. It is available on the market for general usage as soon as it receives FDA approval.

The rate of drug failure has risen in later stages of clinical testing, and this has always been a major concern. The clinical studies (1998- 2008)the failure rate of phase 2 & phase 3was 54 percent [7]. Concerns about safety account for 17% of failures, while ineffectiveness is another factor that accounts for the rest. In phase 2 and phase 3 treatment failure, side-effects and the death risk are also important factors [6 & 8]. The failure of a treatment and the time-consuming process, which takes extremely lengthy periods of time and incurs significant costs, can be annoying, specifically when trials weren't sufficiently successful. ML assists in this procedure by learning from previous data and experience, removing certain unknown aspects and reducing human effort, expense, and time.

3. METHODOLOGY

In supervised learning, a labelled dataset is used to train the model. That model or algorithm separates the data and appropriately predicts the outcomes. Many disciplines and organisations benefit from supervised learning to solve real-world challenges. There are two types of supervised learning: Regression and Classification. To accurately divide and classify data into specified groups, classification algorithm is used. It detects certain entities in a database and attempts to draw judgments about how those entities should be labelled. Regression is used to understand the relationship between dependent and independent variables. SVM, decision-trees, Random Forest, K-Nearest Neighbour (KNN), Naive Bayesian classifier, polynomial regression, linear regression, and logistical regression (LR) are some of the most extensively used supervised learning algorithms [18] and we can use Artificial Neural Network (ANN) [19].

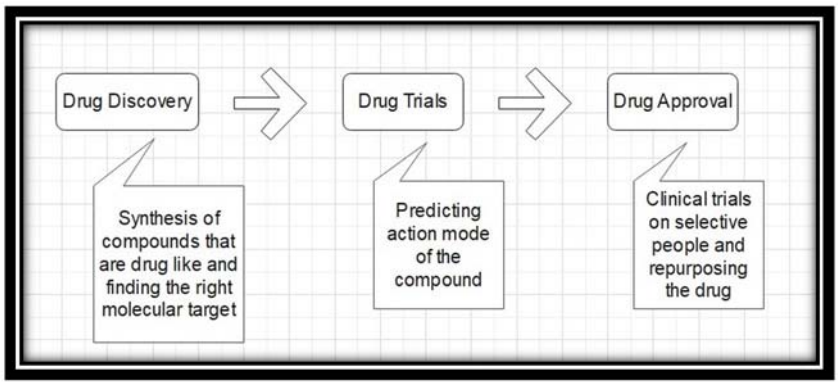


Fig 3.1: Machine learningflow in Drug Development

Unsupervised learning denotes, when there is no output variables. The basic goal of unsupervised-learning is to understand the dissemination of data in order to learn more about it (Fig 3.1). This can be divided into two categories: association and clustering. The clustering challenge requires you to collect data based on a specific pattern or behaviour. You wish to uncover rules that define a substantial percentage of your data in association. Decision tree, SVM, Naive Bayesian, ANN, and Random Forest (RF) algorithms are some of the current Machine Learning applications or algorithms utilised in the field of discovery of drug.

3.1. Support Vector Machine (SVM)

In drug discovery, SVM is used to divide classes of composites mostly based on function selector by creating a hyperplane. It generates an endless number of hyperplanes by using commonalities between lessons. It divides lessons, such as chemicals, into chemical function spaces based on established functions to convey linear knowledge.A first hyperplane is one that is obtained by maximising margin among lessons in N-dimensional space; it is represented by a hyperplane, which is used to categorise information factors using choice limitations. Regression models are important for establishing the relationship between ligand and medication by using queries to forecast.

3.2. Naive Bayesian (NB)

The NB algorithm has evolved into a valuable tool for categorization in predictive modelling. To categorise the capabilities of datasets, it is employed, and depending on the input features, aspect correlation, and dimensionality of the data, it may be one of the most successful solutions for the task. It is still unknown how well NB works with decision tree algorithms for textual content mining. These methods increase the precision of recovered data sets, which are often acquired from large, jumbled sources. In the drug development process, biomedical data classification is critical, especially in the goal discovery subgroup. This method has shown tremendous promise as a classification- tool for biomedical data, which is often congested with unrelated records and data, which is referred to as "noise." It can also be used to predict ligand-goal interactions, which is a significant advancement in lead discovery. This approach lately been included into a range of medication development procedures by researchers. In a study with possible hobby as estrogenic receptor antagonists in breast cancer, Pang et al. used NB designs and other approaches as classifiers for active and inactive medicines.

3.3. Random Forest (RF)

It is a broadly used set of rules for big datasets with several capabilities, since it simplifies the process of eliminating outliers, as well as classifying and designating datasets based solely on relative capabilities defined for the specific set of rules. It's commonly utilised for big inputs and variables, as well as accessibility based solely on statistics series from several databases. It can be used for a range of tasks, such as attributing absent statistics, dealing with outliers, and calculating classification attributes. The basic mathematical structure of RF is made up of an ensemble of uncorrelated choice trees, each of which is in charge of making a single prediction. The candidate who receives the most votes is deemed the most suitable. Although false- positives can occur in any statistical study, when compared to other methods, RF, along with NB and SVM has been recommended to make the fewest number of errors. Character mistakes are reduced when more than one choice timber is used because it assembles a variety of predictions rather than focusing solely on one. RFs are commonly used in drug discovery for character selection, classifiers, and regression.

3.4. Decision Tree: Decision trees are used to classify data and provide suggestions using a set of rules. In the pharmaceutical industry, decision trees are used to solve problems such as drug similarity prediction, developing combinatorial libraries, and generating chemical profiling data. Models based on decision trees are simple to validate and comprehend Drug absorption, permeability, and solubility, as well as metabolic stability, and penetration, are all predicted using decision trees. Because of the hierarchical nature of decision trees, large or substantial datasets are required. Small changes in the dataset can induce splits in the outcomes. The decision tree's effectiveness is also determined by the order in which contentious attributes are selected. Separation traits should be considerate of value or worth [7].

4. DISCUSSION:

In recent years, machine learning has been applied in a variety of fields, including healthcare, where a machine learning model predicts the optimum outcome. Prior to machine learning, the drug development process took far too long (6-10 years), and the success rate of new drugs was extremely low. It takes a lot less time now that machine learning algorithms have been developed, and the success rate is very high. Machine learning to identify novel drugs has three primary steps or phases (Drug Discovery, Drug Trials, and Drug Approval). In the subject of healthcare, machine learning has numerous algorithms, including decision trees, support vector machines, Naive Bayesian, and Random Forest (RF) algorithms. Machine learning has become a significant tool in the drug development field as a result of various applications or algorithms.

5. REFERENCES:

[1] M. R, “Critical review of the role of HTS in drug discovery,” pp. 11(7-8):277–9, 2006.

[2] A. B. A. B. a. J. E. A. Sertkaya, “Examination of clinical trial costs and barriers for drug development,” US Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation, pp. 1-92, 2014.

[3] R. Biswas and K. Haque, “Drug Discovery and Drug Identification using AI,” pp. 49-51, 2020.

[4] N. Stephenson, E. Shane and J. Rowland, “Survey of Machine Learning Techniques in Drug Discovery,” pp. 185-193, 2019.

[5] A. V. Singh and A. Negi, “Towards Better Drug Repositioning Using Joint Learning,” pp. 1-4, 2019.
[6] G. M. Schuhmacher O, “Hinder Changing R&D models in research based pharmaceutical companies,” J Transl Med, p. 14 (1):105, 2016.
[7] R. M. Manne, “Machine Learning Techniques in Drug Discovery and Development,” International Journal of Applied Research, vol. 7(4), pp. 21-28, 2021.
[8] C. D. L. J. W. B. F. J. K. A. Hwang TJ, “Failure of investigational drugs in late-stage clinical development and publication of trial results,” JAMA Internal Med, vol. 176(12), pp. 1826- 1833., 2016.
[9] B. G. Rajagopal and M. Arock, “Application of Machine Learning Techniques for study of drug interactions using clinical parameters for Creutzfeldt-Jakob disease,” 2020 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES), pp. 415-420, 2020.
[10] G. Shobana, “Drug Administration Route Classification using Machine Learning Models,” Proceedings of the Third International Conference on Intelligent Sustainable Systems [ICISS 2020] IEEE Xplore Part Number: CFP20M19-ART; ISBN: 978-1-7281-7089-3, pp. 654-659, 2020.
[11] A. Akay, “Deep Learning: Current and Emerging Applications in Medicine and Technology,” pp. 1-15, 2019.
[12] K. Babaria and S. Das, “Algorithms for Ligand based Virtual Screening in Drug Discovery,” in 2015 International Conference on Applied and Theoretical Computing and Communication Technology (iCATccT), 2015.
[13] A. N. Lima and E. A. Philot, “Use of machine learning approaches for novel drug discovery,” pp. 2-42, 2016.
[14] M. Alam, J. Li and J. Wang, “Drug Discovery Approaches using Quantum Machine Learning,” pp. 1-4, 2021.
[15] R. Gupta, D. Srivastava and S. Tiwari, “Artifcial intelligence to deep learning: machine intelligence approach for drug discovery,” pp. 1315-1360, 2021.
[16] S. S. Kadiyala, “APPLICATION OF MACHINE LEARNING in Drug Discovery,” pp. 1-61, 2018.
[17] S. Barrett and W. L. B., “Advances in the Application of Machine Learning Techniques in Drug Discovery, Design and Development,” pp. 99-100.
[18]H. Maheshwari, P. Goswami and R. Isha, “*A Comparative Study of Different Machine Learning Tools”,* International Journal of Computer Sciences and Engineering, Vol.7, Issue.4, pp.184-190, 2019.
[19] Rana, A., Rawat, A. S., Bijalwan, A., & Bahuguna, H. (2018). Application of Multi Layer (Perceptron) Artificial Neural Network in the Diagnosis System: A Systematic Review. Proceedings of the 2018 3rd IEEE International Conference on Research in Intelligent and Computing in Engineering, RICE 2018.

BIOGRAPHIES

Devesh Maithani pursuingMaster of Computer Application from Uttaranchal University Dehradun. He completed his bachelor's degree in Information Technology from Uttaranchal University in 2020. Currently he is doing internship from Cognizant as a Programmer Analyst Trainee.

Abhishek Chauhanpursuing Master of Computer Application from Uttaranchal University Dehradun. He completed his graduation in Bachelors in Computer Application from Uttaranchal University in 2020. Currently he is doing internship from de-facto InfoTech as a Software Developer Trainee.

Dr Himani Maheshwari is an Assistant Professor in the Department of Computer Applications, Uttaranchal University, Uttarakhand, India. Artificial Intelligence, Machine Learning, Data Science, Web GIS, DSS, and Computer Vision are among her key research interests. Dr. Himani Maheshwari completed her MTech from Uttarakhand Technical University and Ph.D. from IIT Roorkee. She has published more than 25 research papers (9 are in Scopus), 3 book chapters and also registered 4 software copyrights. She received Young Scientist Award in 2020 from STEM Research Society. She also qualified GATE and UGC NET exams.

Mr. Amarjeet Rawat is an Assistant Professor in the Department of Computer Applications, Uttaranchal University, Uttarakhand, India. Recommender System, Machine Learning, Data Science, Big Data, and Web Technologies are among his key research interests. Mr. Amarjeet Rawat completed her MTech from Graphic Era University and pursuing Ph.D. from Uttaranchal University, Dehradun. She has published various research papers and book chapters in reputed journals and conferences.

Identification of the deployment defects in Microservice hosted in advanced VCS and deployed on containerized cloud environment

Amarjeet Singh
710 Duncan Ave Apt#1409 Pittsburgh, Pennsylvania, 15237, USA amarteotia@gmail.com

Vinay Singh
6032 Blue Ridge Dr, Apt#A, Highlands Ranch, Colorado, 80130, USA
vsbuild7@gmail.com

Alok Aggarwal
School of Computer Science, University of Petroleum & Energy Studies Dehradun, India
alok.aggarwal@ddn.upes.ac.in

Shalini Aggarwal
Uttaranchal Intt. of Mgt., Uttaranchal University, Dehradun, Shalinia289@yahoo.com

Pratibha Pandey
Uttaranchal Intt. of Mgt., Uttaranchal University, Dehradun,
pratibhapandey8502@gmail.com

Manisha Khanduja
Uttaranchal Intt. Of Mgt., Uttaranchal University, Dehradun,
manisha.khanduja@gmail.com

Abstract.

Micro-services architecture has evolved as the popular software development model for the enterprise applications. Since enterprise applications are complex by nature and they require out of the box scalability and low latency, these hence Micro-services provides a significant contribution in accomplishing these objectives. Enterprises can achieve a long-term vision of an API-enabled, loosely coupled, highly scalable and flexible platform architecture with Micro-services in containerized cloud environment. It has been observed that there are very few works available on this topic. Software industry needs to emphasize in the area to find out the defects at initial level rather than seen errors in production environments. Few works focus on Micro-services while others on Kubernetes issues and challenges but there is no relation between the two has been found. That motivated us to go for the proposed work, which depicts to identify the defects in early stage for Micro-services deployed on Kubernetes. Few standard basic guidelines for the micro-service architecture, in terms of naming convention, automation, monitoring & warning, fault design, and design philosophy, are proposed.

Key words. Version Control System, Git, Subversion, micro-service, Kubernetes, container

I. INTRODUCTION

As companies embark on digital transformation, the enterprise architecture team aims to design a flexible and scalable architecture. The most important drivers for micro-service architecture are, Duplicate the System of Engagement (SoE) and the System of Records (SOR) using clearly defined API agreements. It provides flexibility to change tools and technologies without affecting experience and business performance. Build large, complex, high-fault tolerance systems that offer scalability and availability to meet the needs, use the lean model to create presentation and integration layers to create responsive and high-performance experiences. The platform should offer needs-based scalability and high availability. Adopt DOOPS principle for continuous integration, continuous delivery with faster release times. Each micro service has its own data model and manages its own data. Data is transferred between micro-services using "mute pipes" like light protocols such as Event Broker and / or REST. Smaller in scope that encompasses the same business functionality. The internal operation is a "black box" from which external programs can only access through the API [1][2]. The foremost common issues are the weakening of the picture quality of the container and attempting to use special pictures without indicating the registration data. Typically, particularly difficult if we are starting to work with Kubernetes or utilizing CI/CD for the primary time [3]. Kubernetes informational suggest that you simply transmit the arrangement using the setup outline or in mystery when the application begins. This information may incorporate database qualifications, API endpoints, or other setup banners [4]. A common botch of engineers is to make applications that don't have and don't have the reference properties of setup maps or arrangement maps/secrets. Whether we are propelling a new app on Cubernets or moving to an existing stage, apps frequently crash early. We have diverse situations and we have a partitioned cabinet set for each environment; one improvement set and two generation sets. We have our possess computer program within the advancement cluster (isolated application) and we utilize our claim computer program (single application, smaller scale administrations, Crown, etc.) within the advancement cluster. We as of now have around ten organizations per diminutive and thousands of module arrangements. With the growing number of holders within the cluster, we are confronting the issue of moderate arranging [5]. For this issue to happen, 70 conditions on a single hub were adequate. It took a few minutes in a push. Moderate planning is no issue in a generation environment where we are dealing with heavy workloads and CPU utilization and thus we don't require more than 70 modules per hub. Kubernetes offers two primary functions, to begin with first is called life sensors and second one is accessibility sensors. Fundamentally, the lifetime/uptime sensor will perform a few activity from time to time to confirm that our application is working accurately (such as sending an HTTP ask, opening a TCP association, or a command in your holder) To act). We confront a number of arrange issues such as tall delays and parcel dropping. In any case, the root of the issue is the Linux part, not the Cybernet itself. The execution of cybernetics largely depends on the execution of the Linux bit. You ought to unquestionably check the processor recurrence control mode. We

ran LXC sometime recently Kubernetes and a few of its administrations moved to Kubernetes without changing the code base. A curiously thing was that their execution on the K8s was second rate to that of the LXC. Indeed the foremost IT divisions appeared moo execution in utilizing the same CPU. On the off chance that the life test comes up short, it will crush our holder and make a modern one. On the off chance that the status test comes up short, this module cannot be gotten to as a benefit endpoint, which implies that activity will not be sent to this module until it is prepared. Within the current conveyance, Progressed Nginx servers point to Kubernetes endpoints by default. The logon controller listens to occasions and resets the grouping without interference. However, let's take a closer see at the method of steady upgrading. Sends and reports any startup prepare to the holder related with the Kubernetes module. At that point report it to the Kubernetes API. Since these forms run in parallel, a few administrations may as of now be down, but will proceed to be sent to the activity module when entering from the current upstream [6]-[8]. Figure 1 shows the Kubernetes eco system for Micro-services.

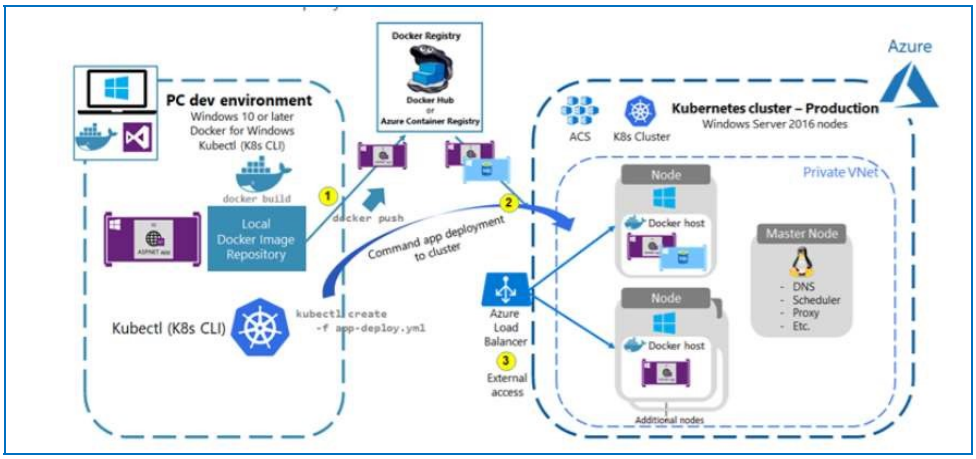


Figure 1. Kubernetes eco system for Micro-services

Few works focus on Micro-services while others on Kubernetes issues and challenges but there is no relation between the two has been found. That motivated us to go for the proposed work, which depicts to identify the defects in early stage for Micro-services deployed on Kubernetes [9-14]. Few standard basic guidelines for the micro-service architecture, in terms of naming convention, automation, monitoring & warning, fault design, and design philosophy, are proposed. Rest of the paper is organized as follows. Section 2 gives the proposed methodology. Results and discussion are given in section 3. have found it difficult and has been interpreted as per the project requirement. Actual granularity for a micro-servicestrategic distance [15] from this, make unchanging holders. In case of any surrenders or vulnerabilities, designers can revamp and redeploy holders.

Inaccessible administration is done through runtime APIs or by making farther shell sessions to the have on which the Micro-services are running [16].

II. METHODOLOGY

The word “micro” used in Micro-services world architects adequate. It took a few minutes in a push. Moderate planning is no issue in a generation environment where we are dealing with heavy workloads and CPU utilization and thus we don't require more than 70 modules per hub. Kubernetes offers two primary functions, to begin with first is called life sensors and second one is accessibility sensors. Fundamentally, the lifetime/uptime sensor will perform a few activity from time to time to confirm that our application is working accurately (such as sending an HTTP ask, opening a TCP association, or a command in your holder) To act). We confront a number of arrange issues such as tall delays and parcel dropping. In any case, the root of the issue is the Linux part, not the Cybernet itself. The execution of cybernetics largely depends on the execution of the Linux bit. You ought to unquestionably check the processor recurrence control mode. We ran LXC sometime recently Kubernetes and a few of its administrations moved to Kubernetes without changing the code base. A curiously thing was that their execution on the K8s was second rate to that of the LXC. Indeed the foremost IT divisions appeared moo execution in utilizing the same CPU. On the off chance that the life test comes up short, it will crush our holder and make a modern one. On the off chance that the status test comes up short, this module cannot be gotten to as a benefit endpoint, which implies that activity will not be sent to this module until it is prepared. Within the current conveyance, Progressed Nginx servers point to Kubernetes endpoints by default. The logon controller listens to occasions and resets the grouping without interference. However, let's take a closer see at the Rest of the paper is organized as follows. Section 2 gives the proposed methodology. Results and discussion are given in section 3. have found it difficult and has been interpreted as per the project requirement. Actual granularity for a micro-service component is required to be de-scoped. Therefore, the actual focus area is to improve the reusability for a micro-service component in a functional domain. Software project teams devoting their efforts to create the Micro-services are presented with many challenges. Close to 70 percent of the project teams require the refactoring and analyzing of the source code as a major factor in the development [9]-[16]. Designers tend to take off shell get to images so they can settle them in generation. Be that as it may, assailants regularly misuse this get to infuse malevolent code. To maintain a strategic distance from this, make unchanging holders. In case of any surrenders or vulnerabilities, designers can revamp and redeploy holders. Inaccessible administration is done through runtime APIs or by making farther shell sessions to the have on which the Micro-services are running[17-20].

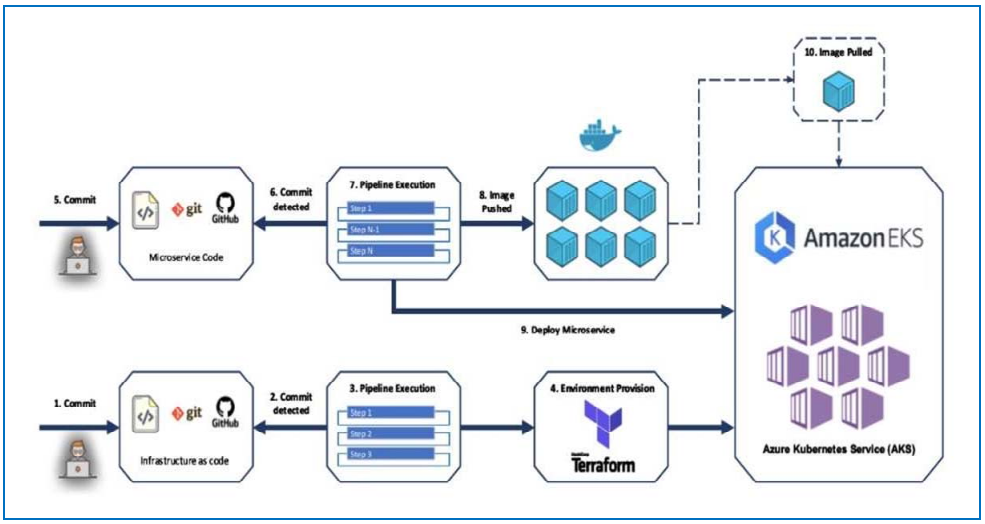


Figure 2. Kubernetes deployment Architecture over Amazon cloud

Therefore, there are numerous open-source bundles for designers with promptly accessible holders, counting Node.js, Apache Web Server and the Linux working framework. In any case, for security purposes, we would like to know where holders start, when they were upgraded, and in the event that they're free of any known vulnerabilities and malicious code. It also included the term of aggregates in interface [21].

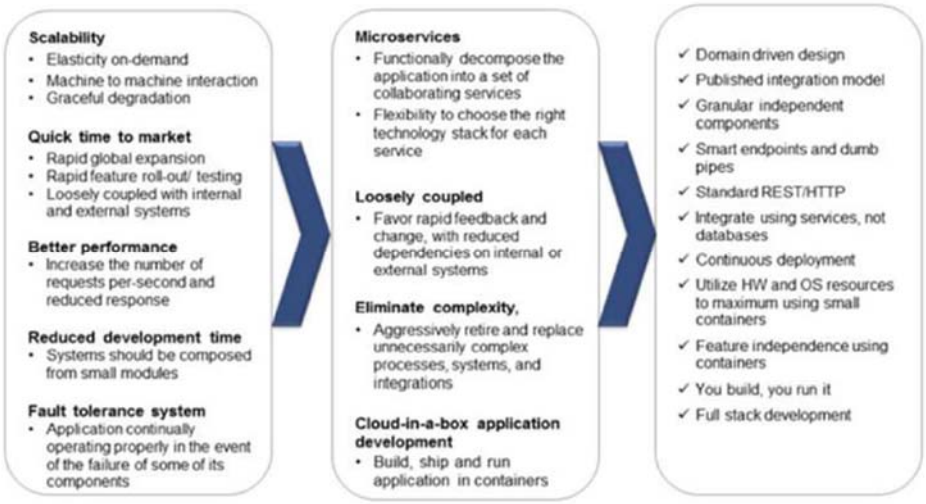


Figure 3. The challenges and Features of Micro-services

It's best to set up a trusted picture store and run pictures as it were from that trusted source. Kubernetes deployment Architecture over Amazon cloud is shown in figure 2 and figure 3 shows the challenges and Features of Microservices.

III. RESULTS AND DISCUSSION

One of the most popular and important plan methodologies for designing cloud-native frameworks is micro-service engineering. We are confident that the software industry clients will embrace this trend in the integration of distributed frameworks. Regardless, the problems and obstacles tend to lead to sub-optimal implementations of this architectural style, leading to an atmosphere in which organisations continue to believe that Micro-services is just another adopted fad. The additional complexity of operating and troubleshooting these frameworks creates new problems. Wherever any of these hazards or challenges lead to strategic decisions, we must ensure that such unique situations are handled by a specialist debt-management mechanism. Finally, the maturity of the framework determines how well ventures can solve these difficulties. The following are basic micro-service architecture naming conventions, automation, monitoring and warning, fault design, and design philosophy standards. Naming conventions: The URL of Micro-services is usually a name that represents the source. We will use it for the program perform the right actions; example Get api/v1/accounts that will list all accounts. Automation: In order to reduce the operational complexity of the microservice architecture, we need to automate the operational complexity tasks such as compiling, deploying, error reporting, alerts, monitoring, automatic scaling, and others. Monitoring and Warning: Tools should be used to monitor the performance and availability of Micro-services watch.

October monitoring services can be configured to monitor disk space, CPU usage, and other settings. To alert operational teams in case of violation of the service level agreement, we need to configure the appropriate thresholds. Fault design: In order to handle faults, we need to implement functions such as Version Control - Micro-service versions are managed using versions that are part of the micro-service endpoint. Design philosophy: Detailing of Micro-services should be based on the following principles; Business functionality (Each micro-service must be designed to show a single business functionality), DevOps Installation (We need to configure the DevOps ecosystem according to the structure and distribution pipeline, Management (We need to define the standards of functionality/performance improvement, implementation and verification). Distributed design: Since the system consists of several Micro-services, we must have the most suitable one parsing of services, clean interfaces for services and convenient database for each service. Finally, the strive of software companies to solve these problems depends on the maturity of the infrastructure and the competence and consistency of design in it business and IT in general. Try to develop before that instead of trying to overcome all this underwater, the states gems and challenges ahead of us, learn and customize our own forward.

REFERENCES

[1] Hou Q., Ma Y., Chen J., and Xu Y., 'An Empirical Study on Inter-Commit Times in SVN,' Int. Conf. on Software Eng. and Knowledge Eng.," pp. 132–137, 2014.

[2] O. Arafat, and D. Riehle, 'The Commit Size Distribution of Open Source Software,' Proc. the 42nd Hawaii Int'l Conf. Syst. Sci. (HICSS'09), USA, pp. 1-8, 2009.

[3] C. Kolassa, D. Riehle, and M. Salim, 'A Model of the Commit Size Distribution of Open Source,' Proc. the 39th Int'l Conf. Current Trends in Theory and Practice of Comput. Sci. (SOFSEM'13), Czech Republic, pp. 52–66, 2013.

[4] L. Hattori and M. Lanza, 'On the nature of commits,' Proc. the 4th Int'l ERCIM Wksp. Softw. Evol. and Evolvability (EVOL'08), Italy, pp. 63–71, 2008.

[5] P. Hofmann, and D. Riehle, 'Estimating Commit Sizes Efficiently,' Proc. the 5th IFIP WG 2.13 Int'l Conf. Open Source Systems (OSS'09), Sweden, pp. 105–115, 2009.

[6] Kolassa C., Riehle, D., and Salim M., 'A Model of the Commit Size Distribution of Open Source,' Proceedings of the 39th International Conference on Current Trends in Theory and Practice of Computer Science (SOFSEM'13), Springer-Verlag, Heidelberg, Baden-Württemberg, p. 5266, Jan. 26-31, 2013.

[7] Arafat O., and Riehle D., 'The Commit Size Distribution of Open Source Software,' Proceedings of the 42nd Hawaii International Conference on Systems Science (HICSS'09), IEEE Computer Society Press, New York, NY, pp. 1-8, Jan. 5-8, 2009.

[8] R. Purushothaman, and D.E. Perry, Toward Understanding the Rhetoric of Small Source Code Changes, IEEE Transactions on Software Engineering, vol. 31, no. 6, pp. 511–526, 2005.

[9] A. Alali, H. Kagdi, and J. Maletic, 'What's a Typical Commit? A Characterization of Open Source Software Repositories,' Proc. the 16th IEEE Int'l Conf. Program Comprehension (ICPC'08), Netherlands, pp. 182-191, 2008.

[10] A. Hindle, D. Germán, and R. Holt, 'What do large commits tell us?: a taxonomical study of large commits,' Proc. the 5th Int'l Working Conf. Mining Softw. Repos. (MSR'08), Germany, pp. 99-108, 2008.

[11] Alok Aggarwal, Vinay Singh, Narendra Kumar, 'A Rapid Transition from Subversion to Git: Time, Space, Branching, Merging, Offline Commits & Offline builds and Repository Aspects,' Recent Advances in Computer Science and Communications, vol. 14: e210621194190, 2021.

[12] V. Singh, M. Alshehri, A. Aggarwal, O. Alfarraj, P. Sharma et al., 'A holistic, proactive and novel approach for pre, during and post migration validation from subversion to git,' Computers, Materials & Continua, vol. 66, no.3, pp. 2359–2371, 2021.

[13] Vinay Singh, Alok Aggarwal, Narendra Kumar, A. K. Saini, 'A Novel Approach for Pre-Validation, Auto Resiliency & Alert Notification for SVN To Git Migration Using Iot Devices,'PalArch's Journal of Arch. of Egypt/Egyptology, vol. 17 no. 9, pp. 7131 – 7145, 2021.

[14] Singh Vinay, and Aggarwal Alok, 'Performance Analysis of Middleware Distributed and Clustered Systems (PAMS) Concept in Mobile Communication Devices Using Android Operating System,' Proc. Third IEEE International Conference on Parallel, Distributed and Grid Computing (PDGC-2014), December 11-13, 2014, Waknaghat, Solan, India.

[15] V. Singh, A. Singh, A. Aggarwal and S. Aggarwal, 'A digital Transformation Approach for Event Driven Micro-services Architecture residing within Advanced VCS,' 2021 International Conference on Disruptive Technologies for Multi-Disciplinary Research and Applications (CENTCON), 2021, pp. 100-105, doi: 10.1109/CENTCON52345.2021.9687973.

[16] V. Singh, A. Singh, A. Aggarwal and S. Aggarwal, 'DevOps based migration aspects from Legacy Version Control System to Advanced Distributed VCS for deploying Micro-services,' 2021 IEEE International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), 2021, pp. 1-5, doi: 10.1109/CSITSS54238.2021.968371

[17] Kumar, A., Memoria, M., & Kumar, V. (2021). Memory optimized deep learning based face recognization. Indian Journal of Computer Science and Engineering, 12(1), 57–64. https://doi.org/10.21817/indjcse/2021/v12i1/211201066

[18] Kumar, A., Singh, D., & Punia, P. (2016). Implementation of image dehazing technique using image fusion. International Journal of Control Theory and Applications, 9(20), 307–315. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85006371817&partnerID=40&md5=17ff92c722005ff83c9c27c371a0f2ed

[19] Kumar, A., & Verma, A. (2016). Missing numbers in graceful graphs. International Journal of Control Theory and Applications, 9(21), 133–136. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85008259815&partnerID=40&md5=b85175173bb98fecc1b0ec98e2185b98

[20] Kumar, D., Chibber, V. K., & Singh, A. (2018). Physical and chemical properties of Mahua and Sal seed oils. In Advances in Intelligent Systems and Computing (Vol. 624, pp. 1391–1400). https://doi.org/10.1007/978-981-10-5903-2_146

[21] Kumar, M., Chandramauli, A., & Ashutosh. (2018). Partial replacement of fine aggregates of fire bricks with fine aggregates in concrete. International Journal of Civil Engineering and Technology, 9(3), 961–968. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85045069805&partnerID=40&md5=65b0161451cfa3352689f31d23a5b191

Biographies

Amarjeet Singh earned a Master's degree in Computer Science from Southern Arkansas University in Arkansas USA in 2021 and Bachelor's degree in Information Technology in 2005. Mr. Amarjeet Singh possesses over 15 years of experience in the Software and IT industry in multiple roles in design development and architecture of modern cloud and Microservices based systems. Provided help to many clients for the digital transformation journey.

Vinay Singh has got 17 years of experience in different domains and technologies in the Software Industry and Teaching. He has worked with AT&T, CapitalOne Bank, Agilent, Walgreens kind of industry leading organization. He is currently working as Manager Software Engineer with Charles Schwab, USA. With this Progressive Experience as leader to DevOps and Software engineering, he is leading a big team of engineer to Build & Release with Continuous Integration/Continuous Deployment using Jenkins, Bamboo and Maven under Agile framework from Dev to Production.

Alok Aggarwal earned PhD degree in Mobile Computing area from IIT Roorkee, INDIA in 2010, Master's degree in Computer Science & Engineering in 2001 and Bachelor's degree in Computer Science in 1995. Contributing over 23+ years in Teaching (CSE & IT), as well as S/W Development. Currently spearheading efforts as Professor (CSE), with

University of Petroleum & Energy Studies, Dehradun (UK) INDIA. He has contributed 250 research papers, 5 Patents, 5 Books and 4 Book Chapters. Current research interests include Machine learning, AI, power control management, MANET & wireless sensor networks.

Shalini Aggarwal earned PhD degree in networking in 2018 from Mewar University Rajasthan, Master's degree in Computer Applications in 2007 and Bachelor's degree in Science in 2001. Contributing over 12^{+} years in Teaching and currently working as Assistant Professor in Uttaranchal University. She has contributed 12 research papers.

Pratibha Pandeyis working as Assistant Professor at Uttaranchal University, Dehradun, India. She is MCA from Punjab Technical University, Jalandhar. M.Tech (Computer Science& Engineering), from Uttarakhand Technical University, Dehradun, Uttarakhand.Her area of interest includes Machine Learning, Big Data, Cyber Security, and Cloud Computing.

Manisha Khandujais working asAssistant Professor at Uttaranchal University, Dehradun, India. She is MCA fromUttarakhand Technical University, Dehradun, Uttarakhand.Her area of interest includes Data Mining, Internet of Things etc.

Artificial Intelligence based Improved Accuracy Model for Edge Computing, and IoT

Sunil Kumar[1], Aditya Saxena[2], Prabhishek Singh[3], Manoj Diwakar[4], , Kapil Joshi[5], Dr. Van Cuong Nguyen[6]

[1,2,3]Amity School of Engineering and Technology, Amity University Uttar Pradesh, Noida, India

[4]Graphic Era deemed to be University, Dehradun, Uttarakhand, India

[5]Department of CSE, UIT, Uttaranchal University, Dehradun, India

[6]Healthcare Big data Research Center, Hanyang University, Korea

[1]skumar58@amity.edu, [2]adityaarchie1234@gmail.com, [3]*prabhisheksingh88@gmail.com*, [4]*manoj.diwakar@gmail.com*, [5]*kapilengg0509@gmail.com*, [6]nvc11vn@gmail.com

Abstract.

The paper critically analyzes the use of Artificial Intelligence in Big data, IoT, AI, and Cloud Computing. The following paper discusses the basics of artificial intelligence that is being used in Big data and AI. All the methods that are being used currently in the fields of 5G and Iot are discussed. The involvement of Artificial Intelligence in these fields has increased and how AI is being used in these fields has been discussed. A model has also been proposed in the following paper to predict the failure rate of the industrial pump just to showcase the connection of AI and IoT

Keywords. *AI,Cloud Storage,Edge Computing,Cloud Storage, IoT.*

1. INTRODUCTION

5G Plus is AT&T's moniker for its 5G mmwave infrastructure (mm-wave). If you hear carriers prompt above 1 Gbps, or download whole movies in one minute or seasons, it's 5G. Edge Computing computes at and near the data source, without relying on the cloud for a dozen data centers. It's not going to go to the cloud. This isn't. It means you've got the cloud. Cloud storage allows you to store your data and files off-site over either the public Internet or a particular private network link. Computer hard discs can only store a limited amount of data. The files must be transferred to external storage devices for users who run out of storage. The Internet of Items (IoT) refers to the interlinked array of things connected to the internet that may collect and transfer data via a wireless network without the intervention of people. There are endless options for people or businesses. Artificial intelligence in various areas can be used to increase safety, speed, and productivity. When the data is being crunched soon, the usage of AI appears like a viable choice for maintaining and reading and analyzing this huge quantity of information. AI can also be integrated with prediction systems used for various network and data organizations [1,2]. While there were numerous competing standards in second, third, and maybe even early versions of fourth-generation wireless standards, the industry has shifted to 4G with long-term evolution (LTE) as specified by the industry organization 3GPP [3]. In most cases, performance goals for wireless capabilities, such as data speeds, coverage, and capacity, are determined globally by a consortium of operators, vendors, and countries in the International Telecommunications Union (ITU). Such user criteria were initially established by the ITU in the late 1990s for the year2000, in a specification known as IMT-2000, upon which the 3G standards were produced. The most recent ITU criteria, known as IMT-2020, are for the year 2020 and outline the user requested requirements for 5G, for which 3GPP is preparing technical specifications.

IfAI will be successful then eventually we will not require a cell phone, smartwatch, or any other smart electronic equipment to access this linked network. The countless applications that we have created would turn into a virtual, networked environment as we place them on our bodies [5].One major advance forward is made conceivable by this: artificial intelligence. Artificial intelligence will work with the ongoing organizations to arrange exchanges to guarantee that all collaborations get the top-tier administration quality from their associations [6]. The entirety of this is made conceivable by one huge advance forward: artificial intelligence. Computer-based intelligence will permit ongoing organization interchanges, guaranteeing that the best conceivable nature of administration from the accessible associations is open to the entirety of its collaborations. Also, these organizations will exist across all boundaries and areas [7]. Allow us to consider, an AI-coordinated retail organization could give customers a quicker item revelation, bringing about a lot more noteworthy request esteems prompting transformation rates in their association's shopping site. In business, an artificial intelligence-based organization can check a large number of papers in minutes to help clinicians in settling on better-educated patient consideration decisions. In the meantime, consider focus networks thatutilization AI can course and serve shopper requests all the more precisely and rapidly [8].

Since the innovation has dispatched the circulation of assets in long haul improvement (LTE) networks was a test. To tackle remote range limitations in 5G, new profound learning ways to deal with reproduce asset assignment challenges in LTE-U Small Base Station (SBS) represent LTE and LTE concurrence. Thought of the upgrade learning strategy dependent on long haul (RL-LS) memory cells to effectively appropriate LTE-U assets all through the unlicensed range to accomplish their commitments. Moreover, support learning has assumed a significant part in heterogeneous organizations, permitting Femto Cells (FCs) to self-governing and deftly recognize the radio climate and change their settings to fulfill the particular nature of administration prerequisites [9]. Al Naimi et al. shown that by using support learning for femtocell self-design dependent on powerful learning games for a completely dispersed multi-target technique, the intra-and between impedance can be fundamentally diminished. Estimating impact and reconfiguration during preparing was used as a 'learning cost'. FCS can utilize this self-putting together capacity to pick the accessible range dependent on mastered using models for artful usage [10].

Unsupervised learning in wireless Communication (5g) [11]:-

Unsupervised learning has an unlabelled training data set, and the machine tries to work without being supervised. This approach is very effective when groups with similar features need to be detected. We do not direct the algorithm to try to discover groupings of related qualities at any stage; without assistance, the algorithm solves this relationshipA common unmonitored learning approach is the clustering of the K-means; several writers studied the applications in the next-generation wireless network system of that particular clustering technique. The cooperation spectrum-sensing technique was suggested by Sobabe et al, which was combined with an improved version of the Gaussian Mixture Model (EM) clustering, and the EM (EM) algorithm. Their study algorithm was shown to exceed the vector-based energy method. Song et al. addressed how the Kmeans algorithm may help choose an effective relay selection from urban networks by grouping and classification skills [12] [13].

In directed learning, each example should be doled out to the appropriate name. The point is to prepare a learning model on a bunch of known ideal issue circumstances and afterward utilize the model to discover streamlined answers for new occurrences. The expectation of a mathematical objective worth given a bunch of indicators is an ordinary goal of directed schooling. This job description is referred to as regression [14].To satisfy the increased traffic requirements LTE tiny cells are more and more used in 5G networks. These tiny cells include their unexpected and dynamic patterns of interference, which grow the requirement for answers that are self-sustainable which can result in higher data rates, and cheaper operator charges. Self-organizing (SON) networks are to be learned in diverse contexts and adapted dynamically. Several AI-based remedies were suggested to determine optimum network design in SONs.Transfer Learning is a common approach for the classification of vectors. In essence, a convoluted neural network (CNN) would be trained in a big dataset such as ImageNet[8], then CNN on a separate vector data set would be fine-tuned. The fortunate aspect here is that some individuals that give the learned weights for public study already conduct the training on the huge dataset. Hence, it is done by retraining with the added or changed data set to adjust the model to temporal changes [15].

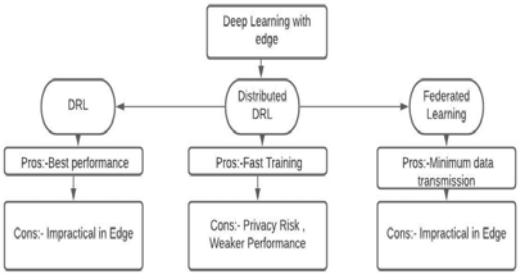
Cloud Storage:-Cloud storage enables data and files to be saved in an offsite place accessed via a private network or a public Internet link. Data you move from your site to your storage is a third-party cloud provider's responsibility [16].

ArtificialIntelligence in Cloud Storage:- Converging of distributed computing and AI will permit clients to store information, yet in addition to investigate it and make determinations from it. Throughout the long term organizations like Microsoft, IBM, Google, and Amazon have put colossally in AI, particularly in the cloud programming arrangements. In the current situation cloud-AI is recognized as the accompanying significant sub-gatherings: AI cloud benefits: The joining of AI in the cloud is known as the keen cloud. Utilizing ML cloud administrations, aside from putting away and organizing clients can apply AI calculations productively and in relatively less time. Following are a portion of the conspicuous AI applications in the cloud:McKinsey has researched to investigate the influence of AI on value creation in a variety of businesses. Moreover, generally, than 15.4 billion dollars per year might be expected to have an impact from McKinsey [17].The main point of adding artificial intelligence in these sectors is to predict the data and automate it like in the case of cloud storage it can be said that AI is used to store data in such a way that it takes the least amount of space in the storage and when the request for data showing is received it can be decoded in such a speed that user doesn’t feel a dealy and user experience isn’t hindered along with less use of storage.

Edge Computing:-One of the biggest difficulties and potential for 5G and next-generation networks is URLLC use cases. Network resources must be distributed closer to consumers and devices to reduce end-to-end network latency and improve dependability. The disaggregated system architecture allows for the deployment of 5G CN and RAN functions at multiple sites to reduce latency. Edge computing is required to exploit the benefits of 5G architectural flexibility to meet the needs of the URLLC class of services.Edge computing helps network operators save money by lowering backhaul traffic and maximizing the utilization of central office resources. The quality of experience for end-users will be improved as a result of faster response times and increased network dependability. Third-party developers will have access to real-time network data via open network APIs, allowing them to create new apps [18].

In edge AI with Federated Learning[19]:-

The quantity of training data that must be transmitted to edges or clouds is sensitive to the privacy and may result in possible privacy breaches, increasing the burden of uplinking wireless channels. Along with the pros of federated learning their are its cons. The following tableshows the pros and cons of various types of learning.



DNN Model	Application	End Devices
SVM/CNN	Image and Video Analysis	Movidius
CNN	Image and Video Analysis, Robotics	Jetson TX1
YOLO	Image Recognition, Robotics	Jetson TX2
AlexNet	Image Classification	Nvidia Tegra K1
CNN	Image Analysis	Neuflow
CNN/DNN	Image Recognition	DianNao
CNN	Vision Processing	ShiDianNao

Fig 1: Deep learning with edge Fig. 2 Edge computing artificial intelligence (AI) accelerator devices:-

Why Artificial Intelligence in Edge Computing[20]:-Deep learning optimization and forecasts take quite a while to converge on results, which is not appropriate for mobile border systems, in particular, the edge computing system-level tasks which demand quick answers in the millisecond scale. Artificial Intelligence in IoT: - Conveying web AI on Thing gadgets diminishes network blockage by permitting information sources to be determined close by, ensuring information transfer protection, and diminishing battery utilization for consistent remote association with entryways and cloud workers [21]. The objective of this exploration was to examine the principle methods that empower AI model execution in the Internet of Things worldview for low-execution equipment. The arrangements gave in the past part permit us to play out the AI calculation on end gadgets yet it is as yet hard to build solid DNNs on little gadgets. In specific examples, computations should be moved to all the more impressive elements from end gadgets. Since the edge worker is easy to use, it very well may be the best method to tackle the calculation issues at the best time. Having AI initiated IoT gadgets will additionally prompt fewer income misfortunes and higher existence of these items as they would improve care with the AI disappointment forecast model.The Internet of Things or IoT alludes to the various actual gadgets that encompass us and are associated with the Internet. These gadgets are associated with one another and trade information, trade information and gather information. This load of gadgets is associated or interfaced with one another utilizing a special referred to identifier as a UID. The "things" in the Internet of Things can be anything, incorporating individuals with heart checking inserts that can send constant information for investigation, or vehicles with sensors that can caution the driver in case of a potential fender bender.Various types are used in iot to speed up the process of data exchanging between the server and the devices moreover a high flow of data requires a very strong network [22].

Training Algorithms for IoT devices[24]:-

One of the main work in the model created for the IoT devices would be the speed then the accuracy as the model had to fast so that it can exchange the data between the servers. The high network connection cost is a bottleneck in the training model, therefore local edge training is required.When the layer is coevolutionary, the input layer is separated by the BODP technique. By decreasing the input size, BODP lowers the calculation. The layer entry is allocated to various work nodes (mobile devices) to identify completely linked layers to ensure the shortest running time. In this example, because the external scenario is pretrained, the Network does not modify its weight. Even though the edge system is hierarchical in design, training between peer edge systems and the cloud can be distributed [25].

II. PROPOSED MODEL

1.1. Problem Statement: Industrial IoT maintenance and trying to predict the failure of the device.

1.2. Model: Illustration of a typical scenario of industrial IoT maintenance. By applying statistical modeling and data visualization, we try to analyze failure performance and predict key industrial equipment such as boilers, pumps, and motors, etc. to manage their repair, maintenance, and optimal performance [26], [37-42].

1.3. Code:-The link of the jyputer network in which the code has been written is available on the link provided above. All of the code has been written in the google collab because of their better GPU accelerating system. The main objective of the code is to find the best classifier or a model to predict the failure of the pump, for this, we have used different classifiers (basically pretrained model) and tested them with various parameters that can be used in this classifier just to find the best of them [27].

1.4. Libraries Used:- There are a few libraries used for this model namely, Numpy is also used which carries out all the scientific calculations along with XGBoost which is an implementation of a gradient boost decision tree developed for speed and performance. Matplotlib library has been used to plot the graphs. Other classifiers have also been used in the below program and they are mentioned in their particular code. Preprocessing library has also been used in order to convert the string data into either int or float [28].

1.5. Dataset:-1-Contains 220314 readings by 51 sensors taken at an interval of 1 minute over 5 months for a large-scale industrial pump. The 2-Each sensor makes some important measurements used to determine the working condition of the pump, like vibration, operative voltage, current drawn, heat generated, RPM, etc.

3-The column 'machine_status' indicates the current working condition of the pump. If it is 0, it indicates that the pump is working as expected. If it is 1, the pump is malfunctioning and needs repair.The following graph shows machine failure rate at different points (1 representing the failure) it represents all the data that is present in the dataset [29].

1.6. Classifier:-A classifier is a sort of machine learning algorithm used in data science to assign a class label to data input. Classifier algorithms utilize advanced mathematical and statistical approaches to produce forecasts of the probability that a certain data entry will be categorized.I have taken these following models because its easy to find the prediction and the accuracy of the model and find it soverall score and the rmse error [30].

XGBoost:-
An open-source library that is used with the python language (in our case but the framework is supported by other languages too). It is one of the most precise classifiers out there providing a boost in the gradient. As in the following figure, it can be seen that the model score is 1which means the model has 100 percent accuracy, in this case, various parameters like n_estimators can change the score of the models, other things can be changed to increase the accuracy. Like in my case n_estimators when taken at default give a better score than taking n_estimators at 1000.

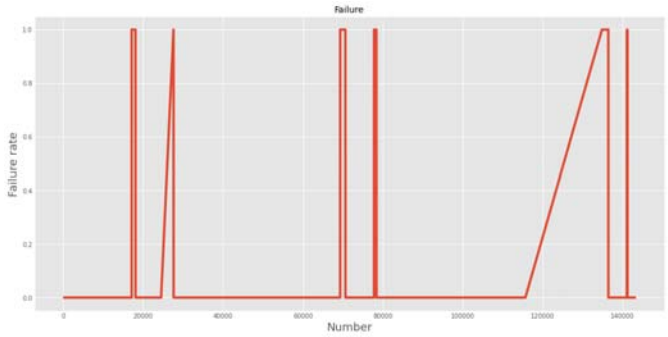
```
import numpy as np
import math
import seaborn as sn
import matplotlib.pyplot as plt
from sklearn.tree import DecisionTreeClassifier
from sklearn.model_selection import train_test_split
from sklearn import metrics
from sklearn.tree import export_graphviz
from sklearn.externals.six import StringIO
from IPython.display import Image
import pydotplus
from sklearn.preprocessing import MinMaxScaler
```

```
sensor_data = pd.read_csv("sensor.csv")
sensor_data = sensor_data.dropna()
print(sensor_data)
```

```
        ID  sensor_00  sensor_01  ...  sensor_50  machine_status     timestamp
0        1   2.465394   47.09201  ...   201.3889             0.0  01-04-18 0:00
1        2   2.465394   47.09201  ...   201.3889             0.0  01-04-18 0:01
2        3   2.444734   47.35243  ...   203.7037             0.0  01-04-18 0:02
3        4   2.460474   47.09201  ...   203.1250             0.0  01-04-18 0:03
4        5   2.445718   47.13541  ...   201.3889             0.0  01-04-18 0:04
...    ...        ...        ...  ...        ...             ...            ...
4901  4902   2.424074   52.30035  ...   168.4028             0.0  04-04-18 9:41
4902  4903   2.415220   52.30035  ...   168.9815             0.0  04-04-18 9:42
4903  4904   2.427026   52.34375  ...   169.5602             0.0  04-04-18 9:43
4904  4905   2.414236   52.34375  ...   169.2708             0.0  04-04-18 9:44
4905  4906   2.419155   52.30035  ...   168.1134             0.0  04-04-18 9:45

[4876 rows x 54 columns]
```

Fig 3:Libraries used in the modelFig 4:Data heads




```
from xgboost import XGBClassifier
model = XGBClassifier(n_estimators=1000)
model.fit(X_train, y_train)
model.score(X_test , y_test)
```

```
1.0
```

Fig 5:Graph for visualizing dataFig 6:XGBoost Classifier

KNN Classifier:-It is straightforward to execute and comprehend but has a big disadvantage because the bulk of these data in use increases substantially. In our case, this model score is 1.0 but when taking the default model the score goes down to 98.8 so taking the nearest neighbors as 50, taking metric as 'Minkowski, and p as 2 gives the highest accuracy [31].

```
from sklearn.neighbors import KNeighborsClassifier
classifier= KNeighborsClassifier(n_neighbors=50, metric='minkowski', p=2 )
classifier.fit(X_train, y_train)
classifier.score(X_test,y_test)
```

```
1.0
```

Fig 8:KNN Classifier

```
[22] from sklearn.naive_bayes import GaussianNB
clf = GaussianNB()
clf.fit(X_train, y_train)
clf.score(X_test,y_test)
```

```
0.9839630562552477
```

Fig. 9: Gaussioan NB

Gaussian NB:-A specific type of NB is a Gaussian Naive Baye algorithm. It is employed especially when the characteristics are continuous. It is also supposed that all features are gaussian, i.e., normal distribution. Among all the three classifier this gives the least accuracy probably due to being a little old then other two. Using the default parameters in this case gave the highest accuracy.

1.7. Prediction:-Every one of the charts plotted with the assistance of matplotlib [32].

XGBoost(Highest accurate one):-Code for plotting the graph.

```
plt.figure(figsize=(16,8))
plt.title('Predictions')
plt.plot(predictions)
plt.legend(['predictions'],loc='lower right')
plt.show()
```

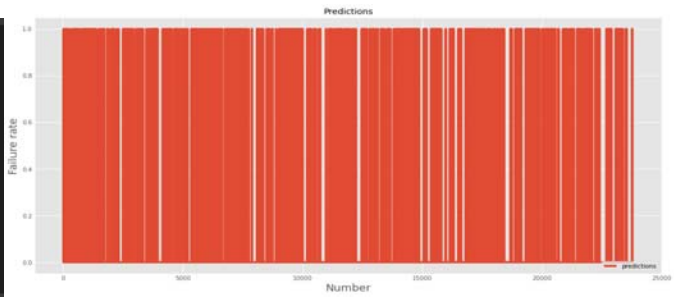


Fig 11 :Prediction using XGBoostFig 10:Code for plotting graph for predictions made by XGBoost

Plot for the prediction(xgboost):-

Prediction graph of Failure rate vs the index number(number of times the device was measured).The graph shows the data that has been plotted for [33]:**KNN Classifier [34][35]:-**Code for plotting the Graph.

```
plt.figure(figsize=(16,8))
plt.title('Predictions1')
plt.plot(predictions1)
plt.legend(['predictions1'],loc='lower right')
plt.show()
```

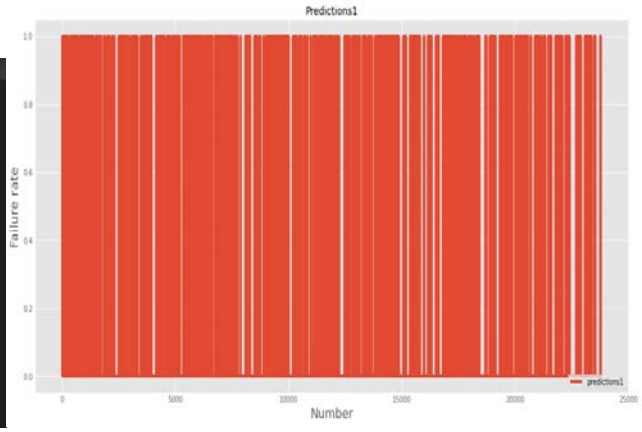


Fig 13:Prediction using KNNFig 12:Code for plotting predictions using KNN

Plot for the prediction:-Prediction graph of Failure rate vs the index number(number of times the device was measured)

- **Gaussian NB(Second Highest accuracy):-**Code for plotting the Graph.

```
plt.figure(figsize=(16,8))
plt.title('Predictions2')
plt.plot(predictions2)
plt.legend(['predictions2'],loc='lower right')
plt.show()
```

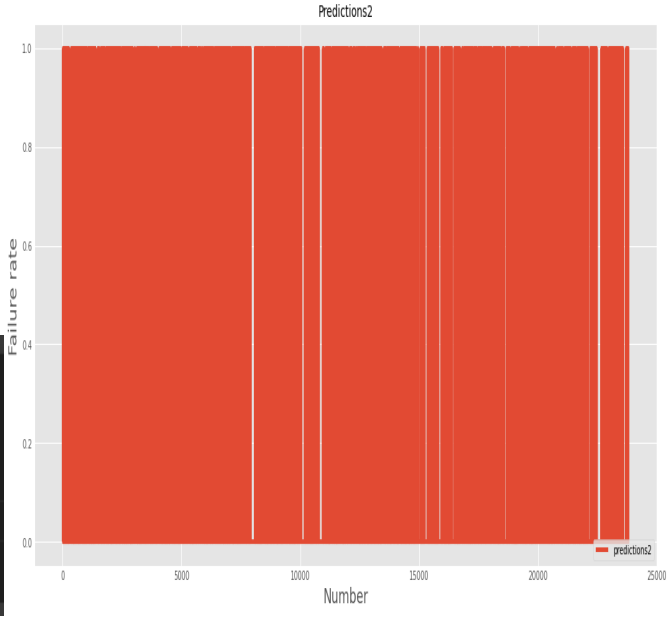


Fig 15:Prediction using Gaussian NB

Fig 14:Prediction using Gaussian NB

Plot for the prediction:-Prediction graph of Failure rate vs the index number(number of times the device was measured)

1.8. Analysis:-The following code has been used to plot the graph.

```
plt.figure(figsize=(16,8))
plt.title('Predictions1')
plt.plot(y_test1)
plt.plot(predictionssmol)
plt.legend(['y_test1','predictionssmol'],loc='lower right')
plt.show()
```

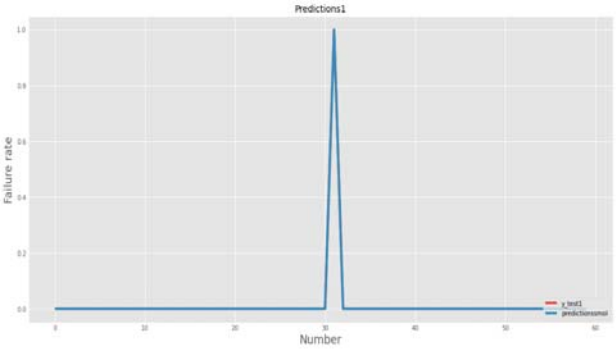


Fig 17:Prediction vs the test dataFig 16:Code for plotting the graph for analysis

The accompanying diagram shows that the expectation model is working perfectly as the predictions completely overlap the y_test graph. This graph is made with the first 60 values in the predictions and the first 60 values in the y_test just for ease of visualization. I have plotted two data y_test1 and predictionssmol which compares the 60 values only.The model has an RMSE(Root Mean Square Error) of 4.19(approx).

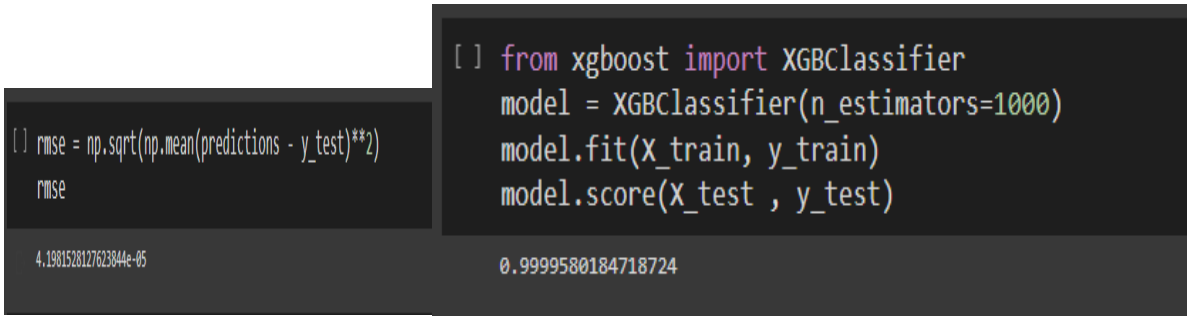


Fig 18:Code for getting the model score of XGBoostFig 19:Rmse Value

1.9. Model Results:-

A model predicting the maintenance of the Pump has been created using the three classifiers and all of them have accuracy of 98 and above. Among the three classifiers used, Xgboost has the highest accuracy this is because this classifier is a pretty new and more accurate one. bodies. From the Above analysis of the model it can be seen that all the classifiers used have an efficiency of 98 and above .

From the above, plainly Artificial Intelligence can be utilized In IoT [43-44] for different purposes for this situation it was a prediction of the failure of the particular pump. The first model classifier has the highest accuracy among all the others because that library has been developed in the recent making it a more accurate and a fast classifiers then its predecessors. This was just a small-scale model bur it can be implemented in a large-scale factory where every second of work is not being done or any failure leads to revenue loss. AI and IoT [45-48] together can be used to decrease the revenue loss and overall increase the quality of the product.

2. CONCLUSION

In conclusion, the research study discusses and elaborates on the promising application and scope of Artificial Intelligence in IoT, 5g, and edge technologies. It very well may be inferred that Artificial insight will be an extremely large part of IoT, 5G, Edge technologies. The growing amount of internet users just supports this argument more as AI would be helping these network providers and these IoT developing companies to maintain the accuracy of their devices while handling this large amount of data flowing through them without costing too much. A lot of money and research is being poured into this and its pace of development in this area doesn't seem to slow down. This paper also dwells on the current challenges faced regarding the implementation of artificial intelligence, some of which are born due to the lack of feasibility in fields such as 5G, where extensive training is required for real-world operations which could be very costly and time-confusing. Because computers aren't truly human, their ability to prove mathematical theorems, make a moral judgment, compose new music, or be genuinely innovative is beyond the scope of neural networks and artificial intelligence.

Refrences

1. Kumar, S., Cengiz, K., Vimal, S. et al. Energy Efficient Resource Migration Based Load Balance Mechanism for High Traffic Applications IoT. Wireless Pers Commun (2021). https://doi.org/10.1007/s11277-021-08269-7.
2. Kumar, S., Ranjan, P., Ramaswami, R., &Tripathy, M. R., "Energy Efficient MultichannelMAC Protocol for High Traffic Applications in Heterogeneous Wireless Sensor Networks," Recent Advances in Electrical & Electronic Engineering (Formerly Recent Patents on Electrical & Electronic Engineering , Vol. 10, No. 3, pp. 223-232, 2017.
3. A. Engelbrecht, Computational Intelligence: An Introduction, 2nd ed. New York, NY, USA: Wiley, 2007.
4. Kumar, S., Ranjan, P., Ramaswami, R., &Tripathy, M. R., "Resource efficient clustering and next hop knowledge based routing in multiple heterogeneous wireless sensor networks," International Journal of Grid and High Performance Computing (IJGHPC) , Vol. 9, No. 2, pp. 1- 20, 2017.
5. Q. Lijun, Z. Jinkang, and S. Zhang, ''Survey of wireless big data,'' J. Commun. Inf. Netw., vol. 2, no. 1, pp. 1–18, 2017.
6. Punhani, A., Faujdar, N., Kumar, S., "Design and Evaluation of Cubic TorusNetwork-on-Chip Architecture" accepted and published by International Journal of Innovative Technology and Exploring Engineering (IJITEE) Volume-8 Issue-6, April 2019 ISSN: 2278-3075.
7. A. Zappone, M. Di Renzo, M. Debbah, T. T. Lam, and X. W. Qian, Model-aided wireless artificial intelligence: Embedding expert knowledge in deep neural networks for wireless system optimization, IEEE Vehicular Technol. Mag., vol. 14, no. 3, pp. 60–69, 2019.
8. Dubey, G., Kumar, S., Navaney, P. "Extended Opinion Lexicon and ML based Sentiment Analysis of tweets: A novel Approach towards Accurate Classifier", International Journal of Computational Vision and Robotics (IJCVR), Inderscience Publishers, 2020 , 10(6), pp.505-521.
9. D. Naboulsi, M. Fiore, S. Ribot, and R. Stanica, Largescale mobile traffic analysis: A survey, IEEE Commun. Surv. Tutorials, vol. 18, no. 1, pp. 124–161, 2016
10. Sunil Kumar et al., "Evolution of Software-Defined Networking Foundations for IoT and 5G Mobile Networks" 2020/10, Pages 350, IGI Publisher, ISBN 9781799846857.
11. F. K. Jondral, ''Software-defined radio-basics and evolution to cognitive radio,'' EURASIP J. Wireless Commun. Netw., vol. 2005, no. 3, Dec. 2005, Art. no. 652784.

12. Kumar, S., Ranjan, P., Ramaswami, R., & Tripathy, M. R., "Energy aware distributed protocol for heterogeneous wireless sensor network" has been published by International Journal of control and automation, Vol.8, no. 10(2015), pp.421-430, http://dx.doi.org/10.14257/ijca.2015.8.10.38.
13. J. Mei, X. B. Wang, and K. Zheng, Intelligent network slicing for V2X services toward 5G, IEEE Netw., vol. 33, no. 6, pp. 196–204, 2019
14. Singh, P., Bansal, A., Kumar, S., "Performance analysis of various information platformsfor recognizing the quality of Indian roads", Proceedings of the Confluence 2020-10 th International Conference on Cloud Computing, Data Science and Engineering, 2020, pp. 63- 76, 9057829.
15. I. F. Akyildiz, P. Wang, and S.-C. Lin, ''SoftAir: A software defined networking architecture for 5G wireless systems,'' Comput. Netw., vol. 85, pp. 1–18, Jul. 2015.
16. Reghu, S., Kumar, S., "Development of Robust Infrastructure in Networking to Survivea Disaster". 2019 4 th International Conference on Information Systems and Computer Networks, ISCON 2019, 2019, pp. 250-255, 9036244.
17. S. Dreiseitl, L. Ohno-Machado, Logistic regression and artificial neural network classification models: A methodology review 35 (5–6) (2002) 352–359.
18. Kumar, S., Ranjan, P., Ramaswami, R., &Tripathy, M. R., "An NS3 Implementation of physical layer based on 802.11 for utility maximization of WSN", Proceedings - 2015 International Conference on Computational Intelligence and Communication
19. Networks,CICN 2015 , 2016, pp. 79-84 , 7546060.
20. Mathworks. Matlab neural network toolbox. https://uk.mathworks.com/p roducts/neural-network.html, 2016.
21. Kumar, S., Ranjan, P., Ramaswami, R., &Tripathy, M. R., "A Utility MaximizationApproach to MAC Layer Channel Access and Forwarding", Progress in Electromagnetics Research Symposium , 2015 , 2015-January, pp. 2363-2367.
22. F. Boccardi, R. W. Heath, Jr., A. Lozano, T. L. Marzetta, and P. Popovski, ''Five disruptive technology directions for 5G,'' IEEE Commun. Mag., vol. 52, no. 2, pp. 74–80, Feb. 2014.
23. Kumar, S., Ranjan, P., Ramaswami, R., &Tripathy, M. R., "EMEEDP: Enhanced multi-hop energy efficient distributed protocol for heterogeneous wireless sensornetwork" Proceedings - 2015 5th International Conference on Communication Systems and Network Technologies, CSNT 2015 , 2015, pp. 194-200.
24. E. C. Strinati, S. Barbarossa, J. L. Gonzalez-Jimenez, D. Ktenas, N. Cassiau, L. Maret, and C. Dehos, 6G: The next frontier: From holographic messaging to artificial intelligence using subterahertz and visible light communication, IEEE Vehicular Technol. Mag., vol. 14, no. 3, pp. 42–50, 2019.
25. Kumar, S., Ranjan, P., Ramaswami, R., "Energy optimization in distributed localizedwireless sensor networks", Proceedings of the International Conference on Issues and Challanges Intelligent Computing Technique (ICICT), DOI: 10.1109/ICICICT.2044.678130.
26. Philip Russom, "Big Data Analytics" Tdwi Research, Fourth Quarter 2011, pp. 5–11.
27. Builtin.com, "Artificial Intelligence, What is Artificial Intelligence? How does it work?", available at : https://builtin.com/artificial-intelligence
28. Certes.co.uk, "Types of Artificial Intelligence: A Detailed Guide", 20thdecember 2018, available at : https://certes.co.uk/types-of-artificial-intelligence-a-detailed-guide
29. M. Singh and G. Baranwal, "Quality of Service (QoS) in Internet of Things," 2018 3rd International Conference On Internet of Things: Smart Innovation and Usages (IoT-SIU), 2018, pp. 1-6, doi: 10.1109/IoT-SIU.2018.8519862.
30. Sunil Kumar, "Advantages and Disadvantages of Artificial Intelligence"Towards data science, 25, Nov.2019, available at: https://towardsdatascience.com/advantages-and-disadvantages-of-artificial-intelligence-182a5ef6588c
31. GeeksforGeeks, "TCP 3-Way Handshake Process", 06th September 2019, available at : https://www.geeksforgeeks.org/tcp-3-way-handshake-process
32. Builtin.com,"What Is Throughput in Networking? Bandwidth Explained", 19th September 2019, Available at: https://www.dnsstuff.com/network-throughput-bandwidth
33. wireshark.org, "TCP Analysis", available at :https://www.wireshark.org/docs/wsug_html_chunked/ChAdvTC
34. Amazon Cloud Foundation, "Cloud Concepts Overview", Module 1, © 2018 Amazon Services, Inc. or its affiliates. All rights reserved
35. F. K. Jondral, ''Software-defined radio-basics and evolution to cognitive radio,'' EURASIP J. Wireless Commun. Netw., vol. 2005, no. 3, Dec. 2005, Art. no. 652784.
36. How AI Will Transform NetworkingHow AI Will Transform Networking. (2017). Retrieved 24 June 2021, from https://www.networkcomputing.com/network-security/how-ai-will-transform-networking.
37. Diwakar, M., Tripathi, A., Joshi, K., Memoria, M., & Singh, P. (2021). Latest trends on heart disease prediction using machine learning and image fusion. Materials Today: Proceedings, 37, 3213-3218.
38. Singh, P., & Shree, R. (2016). Speckle noise: Modelling and implementation. International Journal of Control Theory and Applications, 9(17), 8717-8727.
39. Wadhwa, P., Tripathi, A., Singh, P., Diwakar, M., & Kumar, N. (2021). Predicting the time period of extension of lockdown due to increase in rate of COVID-19 cases in India using machine learning. Materials Today: Proceedings, 37, 2617-2622.
40. Bhatt, M. B., Arya, D., Mishra, A. N., Singh, M., Singh, P., & Gautam, M. (2019, April). A new wavelet-based multifocus image fusion technique using method noise-median filtering. In 2019 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU) (pp. 1-6). IEEE.
41. Tyagi, T., Gupta, P., & Singh, P. (2020, January). A Hybrid Multi-focus Image Fusion Technique using SWT and PCA. In 2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 491-497). IEEE.
42. Ghose, S., Singh, N., & Singh, P. (2020, January). Image denoising using deep learning: Convolutional neural network. In 2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 511-517). IEEE.
43. Bagwari, S., Gehlot, A., Singh, R., Priyadarshi, N., & Khan, B. (2022). Low-Cost Sensor-Based and LoRaWAN Opportunities for Landslide Monitoring Systems on IoT Platform: A Review. IEEE Access, 10, 7107–7127. https://doi.org/10.1109/ACCESS.2021.3137841
44. Singh, R., Gehlot, A., Rashid, M., Saxena, R., Akram, S. V, Alshamrani, S. S., & Alghamdi, A. S. (2021). Cloud server and internet of things assisted system for stress monitoring. Electronics (Switzerland), 10(24).
45. Malik, Praveen, Singh, R., Gehlot, A., Akram, S. V., & Das, P. K. (2022). Village 4.0: Digitalization of Village with Smart Internet of Things Technologies. Computers & Industrial Engineering, 107938
46. Goel, G., Tiwari, R., Rishiwal, V., & Upadhyay, S. (2018). Data preservation by hash algorithm for matrix multiplication over venomous

cloud. PDGC 2018 - 2018 5th International Conference on Parallel, Distributed and Grid Computing, 210–214. https://doi.org/10.1109/PDGC.2018.8745851

47. Pandey, N. K., Chaudhary, S., & Joshi, N. K. (2017). Resource allocation strategies used in cloud computing: A critical analysis. 2nd International Conference on Communication, Control and Intelligent Systems, CCIS 2016, 213–216. https://doi.org/10.1109/CCIntelS.2016.7878233

48. Aggarwal, R., Singh, S., Roul, A. K., & Khanna, N. (2014). Cellphone identification using noise estimates from recorded audio. International Conference on Communication and Signal Processing, ICCSP 2014 - Proceedings, 1218–1222. https://doi.org/10.1109/ICCSP.2014.6950045.

Automation in Horticulture: The Future of Orchards

Siddharth Shankar Bhatt[1*], Amit Bhatt[1], Jayanti Ballabh[1], Sunil Prakash[1], Dev Baloni[2], and Smritilekha Majumdar[1]

[1]Assistant Professor, School of Agriculture, Uttaranchal University, Dehradun

[2] Assistant Professor, UIT, Uttaranchal University, Dehradun

siddharthbhatt@uttaranchaluniversity.ac.in

Abstract

Horticulture is one of the fastest growing industries worldwide which provide varied agricultural produce. However, for providing quality produce it is facing numerous challenges such as lack of skilled labors, climate change and other biotic stress. Artificial intelligence can play an indispensable role in different arena to enhance production and quality. Technologies like Machine Learning can help to solve many ground reality problems like predicting weather and crop health status which further can help famers choose appropriate crops for yield and quality. This technology can also detect soil moisture and pathogens in the field which shall directly enhance fruit and vegetable yield. Big Data Analytics can be used to integrate with Data Image Processing for detection of pest and pathogens with their cure. Drone and Robotic technology can further be used in intercultural operations like weeding, spraying and irrigation. Post harvest losses which accounts for 40 percent in our country can also be reduced with such high tech technology.

Keywords: Artificial Intelligence, Big Data Analytics, Horticulture, Machine Learning

Introduction

Horticulture industry encompasses growing of quality fruits, vegetables, flowers, medicinal and aromatic plants [1], plantation crops etc. India ranks second in fruits and vegetable production in the world behind China. The total area under horticulture in India is 25, 870 thousand hectare and total production is 31, 4671 thousand MT (NHB, 2018-19). This industry at present is facing different biotic and abiotic stresses. Increasing cost and decreasing availability of labour is one of the prominent challenges this industry is going through. Plant protection risk is yet another reason which reduces quality and productivity of produce. Due to continous traditional farming systems Indian agriculture still stands at subsistence level. Climate change in recent have posed new threats to food [2] and nutritional security (Kumar and Joshiba, 2019). With increasing

population and future challenges automation in horticulture seems to be need of the hour (Smitha, 2019).

Automation in horticulture industry can resolve number of ground problems which the orchard growers are facing. Artificial Intelligence helps farmers to choose right crop depending on climatic and soil conditions which may result in high productivity. AI provides 98% accuracy in identification of pest and [3] pathogens. Sensors can provide information about different ripening stages of fruits and vegetables (Soffar, 2019). Different advance technologies like AI, Block Chain Technology, Machine learning and Big Data Analytics may help farmers to enhance production and quality which [4] further shall be a step ahead in sustainable agriculture.

Applications in Horticulture

Prediction of weather

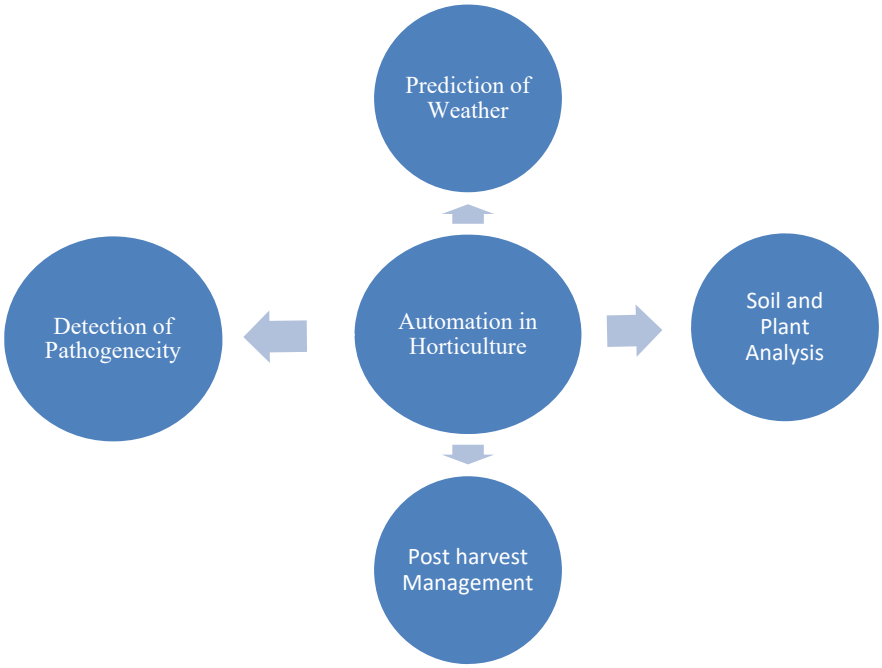
Weather prediction is an important component in horticulture industry as numerous intercultural operations depends on weather. Machine Learning technology can be handy in using algorithms in connection with satellites to predict changing weather on daily basis. These algorithms can also be used in predicting crop [5] health without any contact with large fields. Machine learning also widens the scope of prediction and detection of pest and pathogens in the farmer's field. (Anonymous, 2019).

Soil Health Analysis

Soil health in terms of both moisture and nutrients are basic need for horticultural crops. Nutrient and moisture rich soils not only enhance [6] yield but also increases fruit quality. Both artificial Intelligence and Machine Learning may develop algorithms and sensors for detection and quantification of moisture and nutrients in the soil. Robots may also be used in quantification of soil nutrients which may further provide predictive [7] analysis for recommended dose of fertilizers and irrigation. Cash crops like Cashew nut and rubber are exhaustive crop and require high dose of nutrients. Machine learning and robots may [8] prove solving such problems in these crops.

Detection of pathogenecity

Horticultural crops are more prone to pest and diseases in comparison to other cereal crops. These pest and diseases not only reduce yield but also decreases quality which further fetch lesser price in the market. Algorithms in integration with [9] captured images may detect type of diseases which may further predict its control. This technology can be used in vineyards where grape leaves can be detected with diseases. Due to improper knowledge farmers generally spray pesticides and fungicides equally in all plants which enhance cost of input and wastage of chemicals. Machine learning hence can develop algorithms which shall direct farmers where exactly these chemicals are required which will further reduce cost of input and wastage (Sennar, 2019).



Agrochemical Production

Horticultural crops require numerous synthetic chemicals such as plant bio regulators, weedicides, pesticides etc for high production and quality. These chemicals have tendency to remain in soil and plant enhancing toxicity. Moreover, these chemicals go in food chain which creates various metabolic ruptures. Machine learning with Big Data Analytics can be used to develop models which may help in production of appropriate agrochemicals by capturing images of pathogens, weeds and pests (Baruah, 2019).

Post Harvest Management

Post harvest processes such as cleaning, sorting and grading can be done using Artificial Intelligence and Robots. Sensors can also be used in storage and warehouses to detect stored pest and pathogens. About 40% of horticultural produce is wasted as post harvest loss. Machine learning and Digital Image Processing can be used in reducing such losses which may enhance annual horticultural production (Kamilaris, 2018).

Stress Management

Horticultural crops like fruits and vegetables in their production cycle go through different biotic and abiotic stresses. These stresses at any point of time may harm quality and production of such crops. Digital Image Processing for example may help in detection of such stresses and Machine learning as a result with use of algorithm can help in overcoming such stresses (Singh et al. 2016). Stress management can be released through IT technologies like IOT, AI and cloud [10-15] etc.

Conclusion

Horticulture industry at present revolves around numerous problems which is hampering yield and productivity of crops. Being profitable industry farmers cannot avail the profit through this industry because of such variables. Automation technology like Artificial Intelligence, Big Data Analytics, Digital Image Processing, Block Chain Technology can solve such real issues which directly or indirectly shall improve production and quality of farmers produce. These technologies can also be used in reducing environmental damage developed through improper use of synthetic chemicals used in producing fruits and vegetables.

Acknowledgement

The authors of this manuscript thanks Uttaranchal University for encouragement and motivation in using advance technology in Horticulture.

Conflict of Interest

There is no conflict of interest among authors.

References

1 NHB, 2020. National Horticultural Board database, Gurgaon, Haryana.

2 Soffar H. Artificial Intelligence in Agriculture advantages, disadvantages & uses. Online sciences, 2019. https://www.online-sciences.com/robotics/artificial-intelligence-in-agriculture-advantages-disadvantages-uses/.

3 Kumar PS, Joshiba GJ. Water Footprint of Agricultural Products," in Environmental Water Footprints: Springer, 2019, 1-19.

4 Smita C. Smart Irrigation Techniques for Water Resource Management, in Smart Farming Technologies for Sustainable Agricultural Development, C. P. Ramesh, G. Xiao-Zhi, R. Linesh, S. Sugam, and V. Sonali, Eds. Hershey, PA, USA: IGI Global, 2019, 196-219.

5 Anonymous: Use of Artificial Intelligence in Agriculture. Press Information Bureau Government of India Ministry of Agriculture & Farmers Welfare. 2019; 17:44 IST. https://pib.gov.in/newsite/PrintRelease.aspx?relid=190957

6 Sennar K. AI in Agriculture – Present Applications and Impact. Available, 2019. https://emerj.com/ai-sector-overviews/ai-agriculture-present-applications-impact/

7 Baruha A. Artificial Intelligence in Indian Agriculture – An Industry and Startup Overview, 2019. emerj.rj.com/ai-sector-overviews/artificial-intelligence-in-indian-agriculture-an-industry-and-startup-overview.

8 Kamilaris, A., and Prenafeta-Boldú, F. X. Deep learning in agriculture: A survey. Computers and electronics in agriculture 147 (2018), 70-90.

9 Singh A, Ganapathysubramanian B, Singh AK, Sarkar. Machine learning for high-throughput stress phenotyping in plants. Trends Plant Sci 2016; 21(2):110‑124.

10 Malik, Praveen, Singh, R., Gehlot, A., Akram, S. V., & Das, P. K. (2022). Village 4.0: Digitalization of Village with Smart Internet of Things Technologies. Computers & Industrial Engineering, 107938.

11 Bagwari, S., Gehlot, A., Singh, R., Priyadarshi, N., & Khan, B. (2022). Low-Cost Sensor-Based and LoRaWAN Opportunities for Landslide Monitoring Systems on IoT Platform: A Review. IEEE Access, 10, 7107–7127. https://doi.org/10.1109/ACCESS.2021.3137841

12 Malik, P, Gehlot, A., Singh, R., Gupta, L. R., & Thakur, A. K. (2022). A Review on ANN Based Model for Solar Radiation and Wind Speed Prediction with Real-Time Data. Archives of Computational Methods in Engineering. https://doi.org/10.1007/s11831-021-09687-3

13 Singh, R., Gehlot, A., Rashid, M., Saxena, R., Akram, S. V, Alshamrani, S. S., & Alghamdi, A. S. (2021). Cloud server and internet of things assisted system for stress monitoring. Electronics (Switzerland), 10(24). https://doi.org/10.3390/electronics10243133

14 Karunanidy, D., Ramalingam, S., Dumka, A., Singh, R., Rashid, M., Gehlot, A., ... & AlGhamdi, A. S. (2022). JMA: Nature-Inspired Java Macaque Algorithm for Optimization Problem. Mathematics, 10(5), 688.

15 Sharma, V. K., Singh, R., Gehlot, A., Buddhi, D., Braccio, S., Priyadarshi, N., & Khan, B. (2022). Imperative Role of Photovoltaic and Concentrating Solar Power Technologies towards Renewable Energy Generation. International Journal of Photoenergy, 2022.

THE EFFECT OF DOMESTIC BIOMEDICAL WASTE ON SOIL PH AND TDS

Awadhesh Chandramauli[1], Ritu Bisht[2], Nishima Chadha[3], Vinod Balmiki[4], Teacher Emmanuel Ododo[5]

[1,2,3,4]Department of Civil Engineering, Uttaranchal University, [5]Graduate Assistant-Robotics Education, University of Uyo, Nigeria

awadheshchandramauli@uttaranchaluniversity.ac.in

ritbht02@gmail.com

nishimachaddha@uttaranchaluniversity.ac.in

vinodbalmiki@uttaranchaluniversity.ac.in

emmanuel.ododo15@gmail.com

Abstract.

The amount of garbage produced by houses is significantly greater than that produced by hospitals and medical establishments. According to a recent study, the amount of garbage generated in the house hold sector is over four times that generated by the city's hospitals. Anyone residing in the residence, as well as guests and home healthcare personnel, is at risk from medical waste. If the material is dumped in the usual trash, the trash collectors may also be at risk. Sanitation employees are frequently stabbed by needles while on the job. Before dumping out unwanted medications, some waste management studies are commend combining them with coffee grinds or cat litter. Humans and animal pests will find them educations less appealing as a result (mixing eliminates the odour of the drugs). Then, before throwing it away, seal the mixture in a sealed container. Using 2 separate soils, laboratory studies were designed to simulate the field surrounding an unlined MSW landfill. The highest change in chemical concentration and engineering property was found on soil samples at a radial distance of 0.2 m and a depth of 0.3 m.

Keywords. Domestic Bio Medical Waste, Waste Management, Soil Contamination, Pollutant Transported

1. INTRODUCTION

India is currently faced with a municipal solid waste problem that involves all segments of society. The issue of consumer knowledge and community sensitization is low. At the household level, there [1] is no mechanism for sorting organic, inorganic, and recyclable waste. Unwrapped sanitary napkins, dirty diapers, syringes, blood-soaked cotton, and stained medical instruments are regularly disposed of in the dry waste bin in most Uttarakhand families. Household [2] biomedical waste is one of the most significant sources of biomedical waste in the state. Malaria, diarrheal, cholera, and other water-borne

infections are among the most dangerous. There is a sufficient legal structure in place in the country to deal with MSWM. Its implementation, however, is inadequate. Despite the fact that there is a strict law in place [3], open dumping is the most common method of garbage disposal. Public knowledge, political will, and engagement are all necessary for the successful application of legal rules and the development of an integrated strategy to the country's municipal solid waste management. Medical waste management is critical because of the potential for environmental and public health problems. Environmental regulatory agencies and garbage generators [4] have made numerous efforts in recent years to better manage trash from health care institutions.

1.1 SOURCE OF DOMESTIC BIO MEDICAL WASTE

The most common reasons given by respondents for having unused drugs at home were:

- Failure to follow the prescribed treatment.
- Preservation for future use and near the expiration date.
- Changing the treatment plan when the patient still has stock of the previous medicine or the pharmacy has not activated the discontinue order.
- In many hospitals, the same patient is followed up on.
- The death of the patient. Existence of medicines that have not been used.

1.2 DOMESTIC BIOMEDICAL WASTES AS A HAZARDOUS WASTE

Changing the therapy plan prescription from a few sources patient demise other Presence of un used meds: information as per the medicine [5] gathered in Riyadh twentieth December 2017 to twentieth March 2018 (218 families) it contains genotoxic materials, irresistible specialists, poisonous drugs and synthetics, sharp gear, and radioactive components, clinical waste (MW) is assigned as risky. It basically influences medical services representatives (specialists, attendants, and others), patients and guests to medical services offices, and people who work with rubbish (clinic labourer, incinerator, landfill, and so forth). Emergency clinics in Jordan produce around [6] 4000 tons of clinical waste each year. Because of the flood of Syrian outcasts into Jordan, clinical waste age has expanded. The river contains a variety of metals and non-metals. Some of them are hazardous metals, which are metals that are detrimental to living [7] organisms as well as the environment. If their concentration exceeds the allowed limit, then it causes.

1.3 QUANTITIES AND TYPES OF DRUGS AT HOME

The enumerators made it clear to the respondents that wellbeing enhancements like calcium tablets and prescriptions as of late endorsed by clinical specialists [8] in the former 7 days were prohibited from the overview, just like the amounts and sorts of unused (counting terminated and excess) drugs at their homes. The enumerators made it clear to the respondents that wellbeing enhancements like calcium tablets and medications as of late recommended by clinical specialists in Moreover, we discarded more hard to-gauge measurements structures, like asthma inhalers, from the calculation [9] of drug quantities held at home. According to survey data from another study, only a small percentage of abandoned drugs (e.g.,1%)wherein such unusual forms (Gracia-Vásquezet al.,2015).

2.DESCRIPTION OF SITE

Under the Ministry of Urban Development, the Government of India (GOI) launched the Jawaharlal Nehru National Urban Renewal Mission (JnNURM) (MoUD). As part of this scheme, an Integrated Solid Waste Management (ISWM) Plant is being established in Sheeshambada, Dehradun. The factory is 14000 square feet in size and has a daily [10] capacity of 350 metric tonnes of rubbish. Waste segregation, keeping at the source, primary pickup, street sweeping, secondary storage, transportation, processing, recycling, and scientific disposal of waste were all goals of this plant when it was approved on May 16,2008.Every day, the capital of Uttarakhand creates an average of 300 tonnes of rubbish. 250 tonnes of this is sent to the city's contentious Shishambada Waste Management Plant [11], which has been criticised by locals. Ramky Environ Engineers Limited has a processing plant 30 kilometres from the city. Currently, introduces two waste products: RDF, which is stacking up at the plant due to Ramky's inability to sell it, and low-quality compost [12]. The remaining 50 tonnes of garbage is not collected and ends up on high ways, in streams, and in communal dumpsters. Indeed, the city is having difficulties collecting garbage from residents' homes. Waste pickers in Bindal make a job by picking up trash that has been dumped on the streets and in bins.. To the mix has been added waste from the chicken [13] market, seafood restaurant, slaughter house, and dairy farm.

3.1 COLLECTION OF SOIL SAMPLES

For the investigation, soil from the landfill's base was gathered at random from three distinct places at 300 m lateral spacing from the landfill's edge. After removing the surface debris, polluted soil [14] samples were taken from three separate locations. IS: 3025 (Part 1) was followed for sample collection (1987).The subsurface soil was dug to a depth of about 1 m with a hand auger, and the samples were taken directly from the auger. In air tight containers and bags [15], three groups of soil samples were collected.

3.2 SOIL SAMPLE PREPARATION FOR CHEMICAL ANALYSIS

Sample preservation procedures are designed to limit biological action, hydrolysis of chemical compounds and complexes, and minimize constituent volatility, according to the American Public Health Association. pH control, chemical addition, cooling, and freezing are the only techniques of preservation (APHA, 1995).The standard preservation procedures for components are listed in Table 1. Because almost all preservatives interact with some of the tests, sample [16] preservation with preservatives is difficult. The most straight forward method of preserving most samples is to preserve them at a low temperature (40°C) (APHA, 1995). Following soil sampling, physico-chemical properties such as pH, total dissolved solids (TDS). The soil specimen was broken with a wooden hammer after air drying and sieved at 2.36 mm. A 0.1kilogramme sample of dirt was obtained and soaked in one litre of distilled water for 48 hours. There is due was then filtered to determine the chemical [17] makeup of the soil. A 425 micron IS sieve is used to assess the pH of the soil (APHA 1995). In Dehradun, each soil sample was submitted to a chemical analysis.

Table1. Standard Preservation Methods by Constituents.

S.No.	Parameters	Preservation method	Maximum storage limit.(According to EPA regulations)

1	PH	Analyses immediately	2 hours
2	TDS	Refrigerate at 4°C until ready to analyses, and do so as quickly as feasible.	-

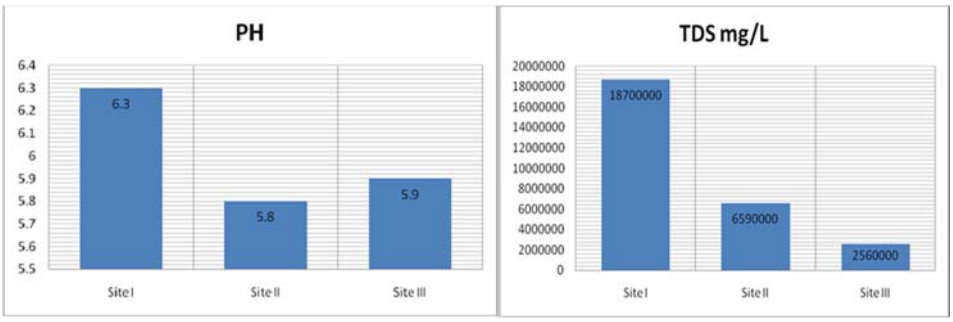
Almost all preservatives interact with [18] some of the tests; sample preservation with preservative is difficult. The easiest strategy to maintain most samples is to keep them at a low temperature (4°C) (APHA, 1995). PH and total dissolved solids (TDS) are the chemical characteristics investigated in soil samples.

4. RESULTS AND DISCUSSIONS

In the laboratory, the obtained samples from the MSW disposal site were analysed. The properties of soil samples were compared [19] to treated soil inland disposal standards. Some of the chemical characteristics of the leachate were found to be higher above the allowed limits.

S.No.	Parameters	Site-I	Site-II	Site-III
1.	PH	6.3	5.8	5.9
2.	TDS mg/L	1.87×10^7	6.59×10^6	2.56×10^6

As shown in Table, the observed pH values in soil samples from all three bore holes are slightly acidic and within the normal range of pH [20] for treated soil. TDS, on the other hand, has high values. Rather than being a primary pollutant, TDS is frequently used as an aggregate indicator of the presence of a wide range [21] of chemical pollutants [22].



5. CONCLUSION

The inappropriate and unscientific management of domestic biological waste has been connected to dangers for individuals who are directly or indirectly involved in the sector, according to this study. Water, air, and soil quality have all been harmed as a result of biomedical waste. Various researchers have discovered risks connected with improper biological waste management and weaknesses in the current system. The majority of the research advocated for the creation of waste management policies, plans, and processes. Furthermore, training programmes for all health care personnel on proper waste

management have been established. It is necessary to do research on the conversion of biomedical waste into energy and other useful goods. This will aid in the prevention of environmental contamination and health risks. The site's soil and water samples were gathered and analyzed to see what pollutants were present and how they affected soil attributes.

Future Scope

• It aids in the preservation of water quality parameters and the provision of a healthier environment for aquatic animals and plants.
• It aided in the development of a sustainable waste management system for health-care waste.
• Rather than discarding, it can be utilized to avoid trash formation or to recover as much garbage as feasible.
• It provides environmentally friendly medical devices and ensures that they are properly disposed of.
• It can assist in reducing the leachate effect.

6.REFERENCES

1.Choudhary S, Ramteke S, Rajhans, K.P, Sahu P.K, Chakradhari S, Patel K.S & Matini L (2016), Assessment of Groundwater Quality in Central India, Journal of Water Resource and Protection, 8, 12-19.

2.Chung Shan Shan and Brooks W Bryan,"Identifying household pharmaceutical waste characteristics and population behaviors in one of the most densely populated global cities", published in Resources, Conservation & Recycling.

3.Ahmed C, AL-Shammary, Majdi F, AL-Ali Kadhim H and Yonuis (2015), Assessment of Al-Hammar marsh water by using Canadian water quality index (WQI), Mesopotamia Environmental Journal, 1(2), 26-34.

4.Mishra, A. K. (2013), "Effect of Salt on the Hydraulic Conductivity and Compressibility of the Two Soil-Bentonite Mixtures with Different Bentonite Contents", Proc. of Annual International Conference on Architecture and Civil Engineering, Singapore.

5.Goswami, D. and Choudhury, B. N. (2013), "Atterberg Limit and Shear Strength Characteristics of Leachate Contaminated Lateritic Soil", Indian Journal of Research, Vol. 3(4), pp.91-93.

6.Bhalla, B., Saini, M. S. and Jha, M. K. (2013), "Effect of Age and Seasonal Variations on Leachate Characteristics of Municipal Solid Waste Landfill", International Journal of Research in Engineering and Technology, Vol. 2 (8), pp.223-232.

7.Kanmani S and Gandhimathi R (2012), Assessment of Heavy Metal contamination in Soil due to Leachate Migration from an Open Dumping Site, Applied Water Science DOI 10.1007/s13201-012-0072-z.

8.Arasan, S. and Temel, Y. (2008), "Effect of Inorganic Salt Solutions on the Consistency Limits of Two Clays", Turkish Journal of Engineering and Environmental Sciences.

9..Fatta D., A. Papadopoulos A., Loizidou M., (1999), "A study on the landfill leachate and its impact on the groundwater quality of the greater area", Environmental. Geochemical Health, Vol. 21, No.2, Pp.175–19

10.Abesh, R. and Anitha, D. R. (2012), "Biochemical Aspects and Formation of Phenolic Compounds by Coir Pith Degraded by PleurotusSajorCaju", Journal of Toxicology and Environmental Health Sciences, Vol. 4 (1), pp.29-36.

11.Bhalla B., Saini M.S., Jha M.K., (2012), "Characterization of Leachate from Municipal Solid Waste (MSW) Landfilling Sites of Ludhiana, India: A Comparative Study." I Journal of Engineering Research and Applications, Vol. 2, No. 6, Pp.732- 745

12.Dhanyasree, A. R., Ajitha, Y. and Sheela, E. (2011), "Study on the Shrinkage, Swelling and Strength Characteristics of Clay Soils under Different Environmental Conditions", Proc. of Indian Geotechnical Conference, Kochi, India

13.Agarwal A., Pandey R., Agarwal M.L.,(2011), "Impact of solid waste leachate on ground water sources -A case study".International Journal of Engineerig Research and Applications, Vol 2 No.2,Pp. 113-118.

15.Abichou, T., Craig, H. B., and Tuncer, B. E. (2000). "Foundry Green Sands as Hydraulic Barriers: Laboratory Study", Journal of Geotechnical and Geoenvironmental Engineering, 126(12), pp.1174-1183.

16.Amin, M. and Hamidi, A. A. (2012), "Impacts of Municipal Waste Leachate on Accumulation of Heavy Metals in Soil and Barley", The 4th International Engineering Conference-Towards engineering of 21st century, Japan.

17.Arasan, S. (2010), "Effect of Chemicals on Geotechnical Properties of Clay Liners: A Review", Research Journal of Applied Sciences, Engineering and Technology, Vol. 2(8), pp.765-775

18.Chattopadhyay, B. C. and Chakravarty, S. (2009), "Containment of Sulfate Pollution in Soil by Natural Geotextile from Jute", Journal of Materials in Civil Engineering, Vol. 21, No. 3.

19 . Poswal, S., Rawat, A., Singh, A., & Dwivedi, N. (2018). Study on waste water using two different enzyme E-1 and E-2. International Journal of Civil Engineering and Technology, 9(4), 54–58. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85046367583&partnerID=40&md5=595628ef0f510c0ddf9f325c04ab779c

20. Singh, P., Usman, M., Chandramauli, A., & Kumar, D. (2018). Brief experimental study on self compacting concrete. International Journal of Civil Engineering and Technology, 9(5), 77–82. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85047820420&partnerID=40&md5=a97197db8ae8a6e388c889b43314a542

21. Thapliyal, S. (2019). Constitutional safeguards provided to civil servants in India: A critical analysis. Journal of Advanced Research in Dynamical and Control Systems, 11(7 Special Issue), 225–228. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85069818911&partnerID=40&md5=e0290cbeabc9856df602736fd5f5e913

22. Singh, S., & Khan, D. (2021). Crack - tip radius effect on fatigue crack growth and near - tip fields in plastically compressible materials. Defence Science Journal, 71(2), 248–255. https://doi.org/10.14429/DSJ.71.15983

Biographies

Awadhesh Chandramauli is working as an Assistant Professor at the Department of Civil Engineering, Uttaranchal University, Dehradun. He has 10 years of experience in teaching and research. He has published more than 15 papers in International and National journals.

Ms Ritu Bisht received the master's degree in environmental engineering from Uttaranchal University, Dehradun.

Nishima Chaddha, She is currently working as an Assistant Professor at the Department of Civil Engineering, Uttaranchal University, Dehradun. Her major research areas include environmental engineering, soil mechanics, construction technology and management.

Vinod Balmiki received, the master's degree in structural engineering from Uttarakhand Technical University in 2017. He is currently working as an Assistant Professor at the Department of Civil Engineering, Uttaranchal University, Dehradun. His research areas include construction technology, design of steel structures, design of R C structures, structural engineering, environmental engineering.

Lightweight Dynamic traffic congestion based Authenticated Protection in Computing Services

Mohammad Shafeeq[1], Sarvesh Kumar[2], Prabhishek Singh[3], Manoj Diwakar[4], **Kapil Joshi[5], Anita Gehlot[6], Thanh-lan Thi Nguyen[7]

[1,2] Babu Banaras Das University, Lucknow , [1]shafeeq.bbd@bbdu.ac.in,

[2]kr.sarvi91@gmail.com

[3]Amity University, Noida, [3]prabhisheksingh88@gmail.com,

[4]Graphic Era deemed to be University, Dehradun, Uttarakhand

[4]manoj.diwakar@gmail.com

[5]UIT, Uttaranchal University, Dehradun, [5]kapilengg0509@gmail.com

[6]UIT, Uttaranchal University, Dehradun, eranita5@gmail.com

[7]Assistant Professor at Department of Liberal Arts, Wonkwang University, Korea thanhlan.edu218@gmail.com

Abstract.

In state-of-the-art GPS engaged hand-held particular contraptions, we use to fill in as a source of perspective point that sends our region. Similarly, the region-based organizations that emerged recently use region data. The region data given to these area-based expert communities has sufficient fragile information to overemphasize. The region security become a space of stress of late. Made and making countries are by and by in progress to make regulations against the use of region information without the consent of the client or without a real warrant. Regardless, these judicious regulations conceivably deal with the circumstance when the region-based data has actually been manhandled. As such, experts are advancing endeavours to track down a solution for secure region data. A large portion of the investigation frameworks proposed actually make them think ordinary that we give a region rather than exact headings of the region of the client. Also, such attacks are found by the researchers in which the best speed of the client can be used in revealing the region of the client.

Keywords. Computing security Parameters, Secure Management, Security Services, Cloud Image Security Simulation, Image computing.

1. INTRODUCTION

A major concern of research in this area is to enjoy the location-based services while protecting location privacy. A huge exploration exertion has been made lately to ensure

the area protection of the client while utilizing area-based administrations. The different procedures can be named follows [1]

• Shrouding Granularity: It requires an area of the shrouded locale to be more prominent than the client determined limit [2].

• L-diversity: L-diversity requires more than one building to be a part of cloaked region [5].displacements and in the systematic way ICT collaborate with other sectors of the economy to provide energy efficiency (smart grids, smart buildings, intelligent transportations systems, etc…). ICT and in particular data centres have a strong impact to

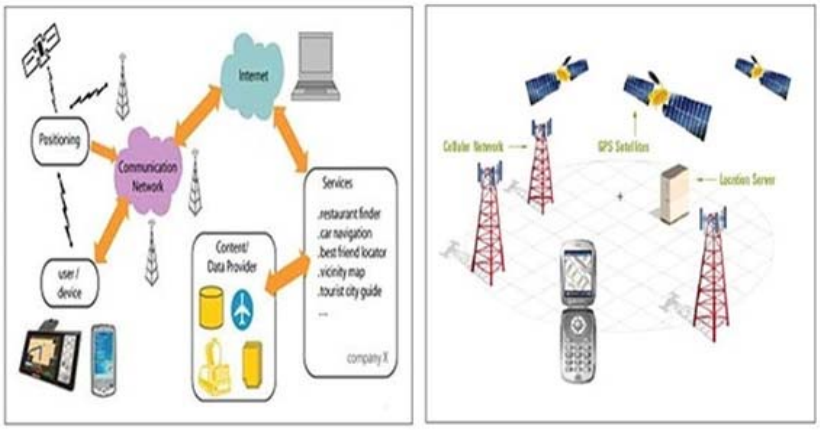


Figure 1.1. LBS Components and Information Flow

2. DATA CENTER INFRASTRUCTURE AND POWER CONSUMPTION

A significant research effort has been made in the recent years in order to protect the location privacy of the user while using the location-based services [14].

2.1. Metrics of Location Privacy

There are three main metrics used in literature to compute the cloaked region k-anonymity [12][11], cloaked granularity and l-diversity.

Assuming the district registered is enormous and clients need quality administrations, a period delay known as transient shrouding is applied, i.e., we defer clients' administration demands for quite a while. As the thickness of the client builds, a more modest shrouding area can then be figured. The major problem associated with K- anonymity technique is that in crowded places the cloaked region may be very small (a single building). So, location privacy of the user is not achieved in that case.

3. RESEARCH CHALLENGES IN CLOUD IMAGE SECURITY

A significant research effort has been made in the recent years in order to protect the location privacy of the user while using the location-based services [14]. The research focuses on the Image color enhancement Techniques to identify the Images[17]. The effect

of images reflect the wavelet packet for co-relation[18]. The Hybrid Image Enhancement for cubic technology is used for the model[19].

The Consider two cloaking regions A and B. We take into account the privacy model in which the different locations are divided in to sensitive locations and no sensitive locations . The user specifies a threshold value of its association with the sensitive location between 0 and 1. The hausdorff distance between cloaking regions A and B is formally defined as:

$$Dhaus(A,B) = max\{h(A,B), h(B,A)\} \quad (3.1)$$

Where,

$$h(A,B) = max\ min\ d(p', p'') \quad (3.2)$$

$$\text{dpp} = \max\ \max \text{d}\ (\text{p}'\ \text{p}'') \quad (3.3)$$

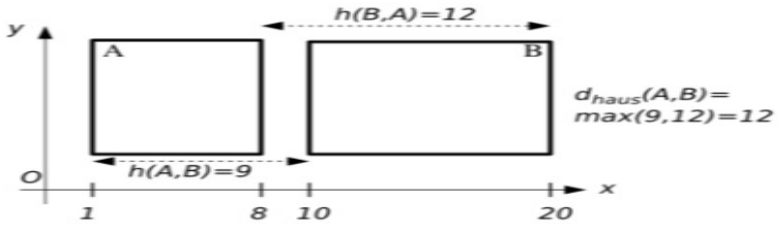


Figure 3.2. Distance Metrics

4. PROPOSED METHODICAL STRUCTURE

Input: A set of requests waiting for anonymization, a new query request u

Output: A set of cloaked requests

Step 1: In the first step the input is taken as request by incrementing the max clique.

Step 2: In the second step the calculation of max clique set area must be found.

Step 3: In the third step the formulation of cloaked region will be found for the users.

Step 4: In the fourth step the performance will be make efficient from the request of max clique sets.

4.1. Experiments and results:

We have developed a prototype of the given algorithm using C++. We have used Microsoft Visual Studio 2008 for the implementation. The large data-set of the location information are not available on internet because of the sensitivity of the information.

K value	Successful Requests	Expired Requests	Success Rate
2-7	64342	7999	.8892
4	65148	7195	.9003
5	63765	8576	.8812
6	62343	9998	.8616
7	61868	10474	.8550

Figure 4.3. Results with no. of users 15000 and sensitive location area 10%.

K value	Successful Requests	Expired Requests	Success Rate
2-7	59576	12764	.8233
4	60408	11932	.8348
5	59350	12996	.8202
6	58872	13467	.8136
7	58135	14210	.8034

Figure 4.2. Results with no. of users 15000 and sensitive location area 15%.

We observe from 4.1 that the fall in success rate is more when the sensitive locations are increased from covering 10% to cover 15% of service area than when increased from covering 5% to 10% of the service area.

5. CONCLUSION AND FUTURE SCOPE

In this work we have done the calculation gives an area security assurance strategy that deals with area subordinate assaults. Yet, this calculation utilizes a frail security profile and doesn't consider the way that clients are more stressed over their area at specific spots. We in our work altered it to meet the assault model and security profile that are stricter. The changed calculation considers a severe assault model and more grounded protection profile. A model of the proposed calculation is created and is tried utilizing the information produced by the Thomas Brinkhoff Generator. We gathered the outcomes inside a field of 10000 x 10000 units and touchy areas covering 5%, 10% and 15% of the complete space of administration. The achievement pace of the calculation diminishes as the space of touchy areas inside the help region increments. We get up to 93% achievement rate if there should be an occurrence of delicate areas covering 5% of the absolute help region.

6. REFERENCES

[1] C. Bettini, X. SeanWang, and S. Jajodia. Protecting privacy against location-based personal identification. In Proceedings of 2nd VLDB Workshop on Secure Data Management (SDM), volume 3674/2005 of Lecture Notes in Computer Science, pages 185{199. Springer, 2005.

[2] Diwakar, M., Tripathi, A., Joshi, K., Memoria, M., & Singh, P. (2021). Latest trends on heart disease prediction using machine learning and image fusion. Materials Today: Proceedings, 37, 3213-3218.

[3] Singh, P., & Shree, R. (2016). Speckle noise: Modelling and implementation. International Journal of Control Theory and Applications, 9(17), 8717-8727.

[4] Tiwari, A., Sharma, R. M., & Garg, R. (2020). Emerging ontology formulation of optimized internet of things (IOT) services with cloud computing. Soft Computing: Theories and Applications; Pant, M., Sharma, TK, Verma, OP, Singla, R., Sikander, A., Eds, 31-52.

[5] R. Arun Raj, S.N. George, and P.P. Deepthi. An expeditious chaos based digital image encryption algorithm. In Recent Advances in Information Technology (RAIT), 2012 1st Interna tional Conference on, pages 14{18, 2012.

[6] Tiwari, A., & Sharma, R. M. (2016, August). Potent Cloud Services Utilization with Efficient Revised Rough Set Optimization Service Parameters. In Proceedings of the International Conference on Advances in Information Communication Technology & Computing (p. 90). ACM.

[7] John Krumm. A survey of computational location privacy. Personal Ubiquitous Comput., 13(6):391{399, August 2009

[8] Tiwari, A., & Sharma, R. M. (2018). Realm Towards Service Optimization in Fog Computing. In Proceedings of the International Journal of Fog Computing (IJFC)" IGI Global.

[9] Wadhwa, P., Tripathi, A., Singh, P., Diwakar, M., & Kumar, N. (2021). Predicting the time period of extension of lockdown due to increase in rate of COVID-19 cases in India using machine learning. Materials Today: Proceedings, 37, 2617-2622.

[10] Bhatt, M. B., Arya, D., Mishra, A. N., Singh, M., Singh, P., & Gautam, M. (2019, April). A new wavelet-based multifocus image fusion technique using method noise-median filtering. In 2019 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU) (pp. 1-6). IEEE

[11] Xiao Pan, Jianliang Xu, and Xiaofeng Meng. Protecting location privacy against location-dependent attacks in mobile services. IEEE Transactions on Knowledge and Data Engineering, 24(8):1506{1519, 2012.

[12] Stefan Steiniger, Moritz Neun, and Alistair Edwardes. Foundations of location based services lesson 1 cartouche 1- lecture notes on lbs, v. 1.0.

[13] Yue Sun and Guangyi Wang. An image encryption scheme based on modified logistic map. In Chaos-Fractals Theories and Applications (IWCFTA), 2011 Fourth International Workshop on, pages 179{182, 2011.

[14] Singh, A. K., Gandhi, V. C., Subramanyam, M. M., Kumar, S., Aggarwal, S., & Tiwari, S. (2021, April). A Vigorous Chaotic Function Based Image Authentication Structure. In Journal of Physics: Conference Series (Vol. 1854, No. 1, p. 012039). IOP Publishing.

[15] Sarvesh, K. (2017). Discrete Gravitational Search Algorithm for Virtual Machine Placement in Cloud Computing. International Journal of Pure and Applied Mathematics, 117(19), 337-342.

[16] Kumar, S., Singh, S., Khatoon, A., & Agarwal, S. (2019). A Multiple String and Pattern Matching Algorithm Using Context-Free Grammar. In Emerging Trends in Expert Applications and Security (pp. 97-102). Springer, Singapore.

[17] M. Pandey, R. K. Bharti and A. K. Bhatt, "A Study of Color Enhancement Techniques for Input Images," *2017 2nd International Conference on*

Computational Systems and Information Technology for Sustainable Solution (CSITSS), 2017, pp. 1-7, doi: 10.1109/CSITSS.2017.8447690

[18] Diwakar, M., & Kumar, M. (2018b). CT image denoising using NLM and correlation-based wavelet packet thresholding. *IET Image Processing*, *12*(5), 708–715. https://doi.org/10.1049/iet-ipr.2017.0639

[19] M. Pandey, "Futuristic Hybrid Image Enhancement Using Fuzzy and Cubic Interpolation Methods," *2021 International Congress of Advanced Technology and Engineering (ICOTEN)*, 2021, pp. 1-6, doi: 10.1109/ICOTEN52080.2021.9493446.

[20] Sajwan, V., & Ranjan, R. (2019). Classifying flowers images by using different classifiers in orange. *International Journal of Engineering and Advanced Technology*, *8*(6 Special Issue 3), 1057–1061. https://doi.org/10.35940/ijeat.F1334.0986S319

A Bibliometric Analysis of Stress Level Prediction of Working Pregnant Women using VosViewer

Dr. Sonal Sharma[1]

Mr. Praveen Kumar Shah[2]

Mr. Sameer Dev Sharma[3]

Mr. Abhishek Kumar Pathak[4]

[1, 2, 3,4,] *Uttaranchal Institute of Management, Uttaranchal University, Dehradun, India,*
hod_it@uttaranchaluniversity.ac.in, praveenshah@uttaranchaluniversity.ac.in,
sameersharma@uttaranchaluniversity.ac.in, abhishekpathak@uttaranchaluniversity.ac.in

ABSTRACT

In today's world, stress is one of the most pervasive problems.89 percent of working professionals in India report feeling stressed. 4 in 10 women experience anxiety for various reasons. The purpose of this examination is to determine the scope of Stress Level Prediction of Working Pregnant Women Research through a bibliometric analysis. Data on pregnant women who work were gathered using the Scopus database. Subject headings with keywords and abstracts in female Stress Recognition studies were used as a lens to retrieve search results. The VOSviewer software was used to extract search results. Later, the results of Bibliometric mapping were examined in greater detail. As per our research this is the first study conducted with the help of VosViewer to do the Bibliometric analysis of stress factors of working pregnant women in India.

Keywords: Stress, Working Pregnant Women, **Anxiety, Bibliometric Analysis**

1. INTRODUCTION:

Hans Selye coined the term "stress" in 1936, defining it as "the non-specific response of the body to any demand for change. Stress is extremely frequent among working women during pregnancy, and it can contribute to mental illness and some foetal abnormalities. Nowadays, it is common for women to work outside the home. According to an article published by Hindustan Times, a study carried out by the World Bank found that 42% of women are graduates, while 33% of men are. A World Bank study found that there is 27% participation of working women, boosting India's potential GDP. A growing body of research indicates that maternal stress during pregnancy has a strong impact and can influence the development of the unborn. Working women have family and work responsibilities, and they tend to be more stressed than men. Having a job while pregnant can have long-term effects

on a woman's overall health, and pregnancy is a time when psychosocial symptoms and poor oral health are more likely to show up.

They get stressed out if their jobs could harm the pregnancy, or if they'll lose their job or the pregnancy will end their career. Work-related stress and burnout can happen when there is a lot of pressure at work and you don't feel like you have a lot of power. If you're stressed out while you're pregnant, you might not be able to have a healthy pregnancy or give birth to a healthy child. In the classroom, a working woman's desire to work even when she is stressed out usually doesn't work out. There are a lot of things that can make people stressed at work, like having too much work, unrealistic deadlines, difficult co-workers, negative attitudes toward pregnant employees, difficulties getting time off to go to hospital appointments or scans, bullying, refusing to make reasonable job changes, and maternity discrimination. People in the education field should pay attention to this if their job requires them to stand for a long time and work long hours without taking a break. There is also a fear of not having a job and unemployment as strong evidence that stress can harm mental health. Working women are more likely to be angry, anxious, depressed, sad, tired, restless, forgetful, overstressed, and burnt out when they think about these things that make them stressed. Stress makes them feel bad about their job, and they also lose interest and concentration because of bad feelings and bad weather. This can make them more stressed.

In this paper, Bibliometric analysis is used for Bibliometric the research papers. The Scopus data set is used for the Bibliometric analysis. Collected articles and findings from Scopus database on working pregnant women prove the major factors that can lead to stress and anxiety.

2. MATERIALS AND METHODS:

Across the globe stress is the major problem for working pregnant women. A literature about stress prediction for working pregnant women. The SCOPUS database was examined between 2012 and 2021, according to the findings. TITLE-ABS-KEY (working pregnant women) AND TITLE-ABS-KEY (stress) AND TITLE-ABS-KEY (stress) AND TITLE-ABS-KEY (stress) AND TITLE-ABS-KEY (stress) AND TITLE-ABS-KEY (stress) AND TITLE-ABS-KEY (stress) AND TITLE-ABS-KEY (working AND pregnant AND women). A CSV file was created using the information extracted from the documents, which included the author, year of publication, language of publishing, journal, title, affiliation, keywords, and kind of document. Recovered data was made available to the public on November 30th, 2021. Study of bibliographic coupling, co-authorship, co-occurrence, citations and co citations was carried out with the aid of the software VosViewer. (1.6.18). [21-24]

3. RESULTS AND DISCUSSION:

The publication year, author, language used, journal, title, affiliation, keywords, and kind of document, as well as an abstract and total citation count, were all obtained from the collected publications addressing the issue and exported to a CSV file. In November 2021, the data recovery was accomplished. VosViewer was used to investigate bibliographical coupling, co-authorship, co-occurrence, citations, and co-citations (1.6.18). "Links characteristics" and "Complete connection strength property" are two frequent weight properties that are linked.

Bibliometric Investigation of the Keywords:

The study report's authors provided keywords that appeared multiple times in the SCOPUS database. These keywords were utilised to undertake a thorough examination of the 668 keywords, 526 of which satisfied the criteria. The keywords "pregnant women" (whole link strength 1729) and "anxiety" (total link strength 621) had the strongest links to "mental tension."

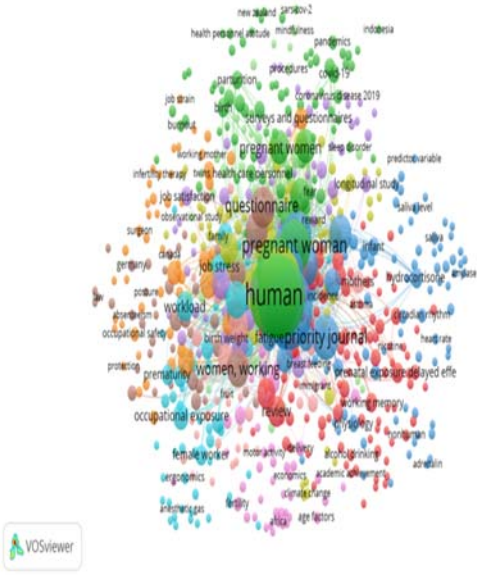


Fig 3.1. The Bibliometric study of Stress Detection keywords in published works is summarised in the table below. All of the keywords appear in the same sentence. The recurrence of the occurrence is indicated by the node dimension. Co-occurrences within that publication are indicated by bending between nodes. The circle's various shapes represent the keyword's importance. The closer four nodes are, the more often four terms co-occur**.**

Table 1- Describe the top 10 Journal in the field of working pregnant women Stress**.**

id	source	documents	citations	total link strength
1	american journal of industrial medicine	3	269	2
2	social science and medicine	3	233	0
3	stress	2	168	0
4	paediatric and perinatal epidemiology	3	159	1
5	american journal of public health	2	154	6
6	obstetrics and gynecology	3	142	2
7	psychoneuroendocrinology	3	100	1
8	journal of epidemiology and community health	2	78	2
9	bmc pregnancy and childbirth	3	62	0
10	infant mental health journal	3	62	0

Table 1- Show the list of Top 10 Journals. In American journal of industrial medicine Journal 3 documents has been published with citation 269 and total linked strength 2

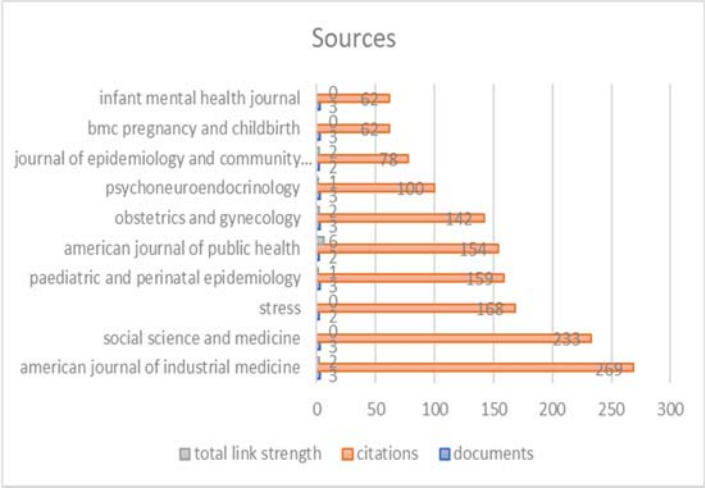


Fig-3.2.-The top ten most active journals in the world

Table 2 –In the field of stress recognition, these ten countries have the best records for innovation.

id	country	documents	citations	total link strength
1	united states	60	1635	11
2	united kingdom	16	450	2
3	canada	15	444	17
4	germany	11	174	4
5	poland	9	27	11
6	australia	8	290	0
7	france	8	112	4
8	spain	8	82	0
9	netherlands	7	174	10
10	norway	7	168	0

Table 2- Shows the list of top 10 Countries contributing in this domain related with stress prediction for working pregnant women

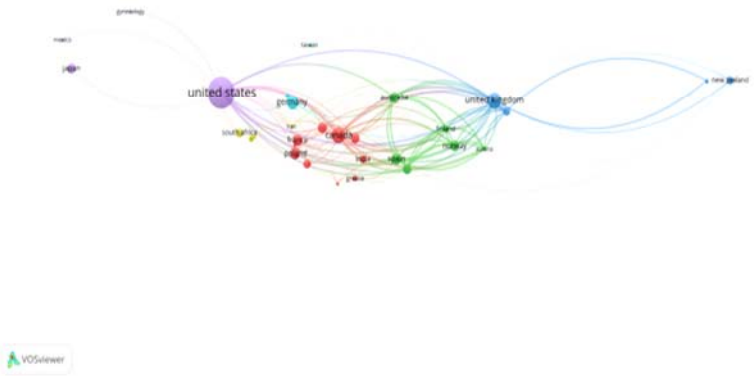


Figure – 3.3.1(A)

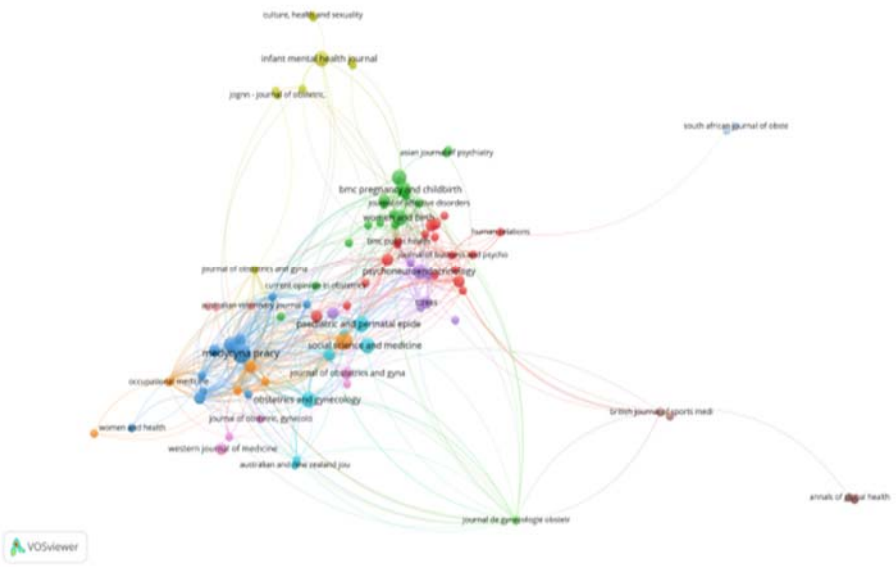


Figure – 3.3.2. (B)

Fig-3 Illustrates the bibliographic connection of – (A) contains country references. The different colours represent the different groups, while the diameter of the circles represents the frequency of referrals. (B) Citations of the sources The circle's diameter denotes the frequency of references, while the various colours represent the various categories**.**

Bibliometric Examination of the Co-Citation

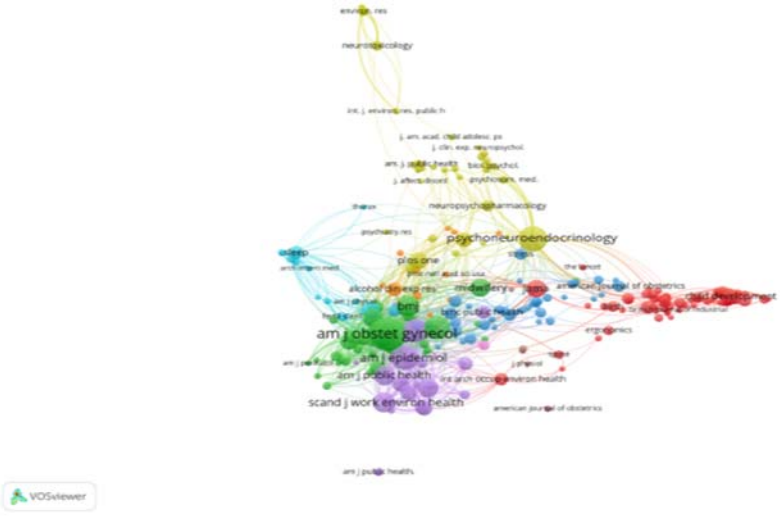


Figure-3.4.1(A)

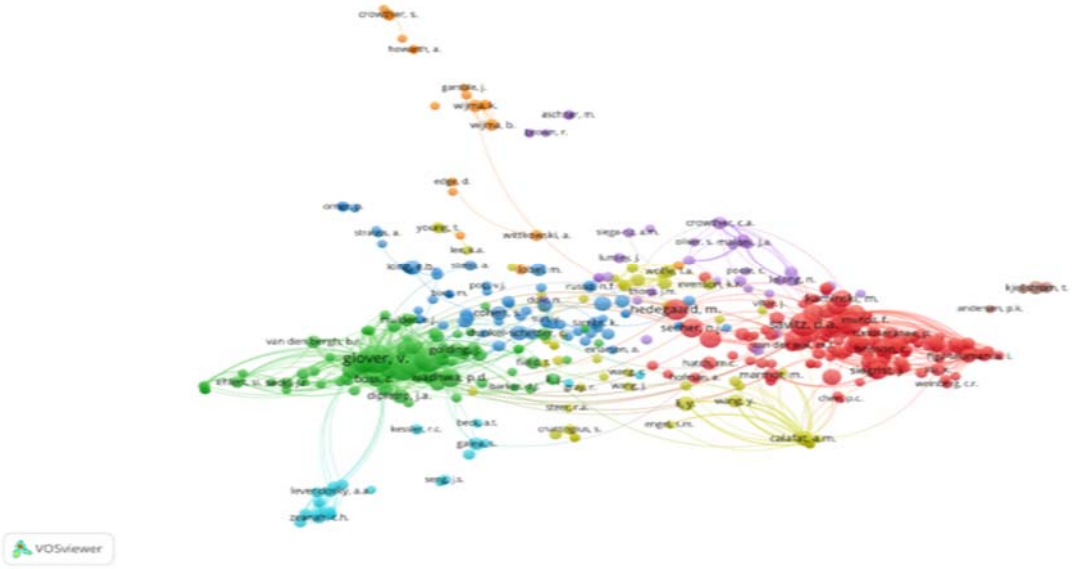


Figure-3.4.1.(B)

As seen in Figure, the records and sources used to construct a bibliography are linked together.4(A), Nine groups were obtained. Figure 4(B) – shows the co-citation of authors Glover, v. has 57 citations with 5116 link strength.

Fig-5 The Bibliometric analysis of the co-citation and bibliographic coupling is shown.

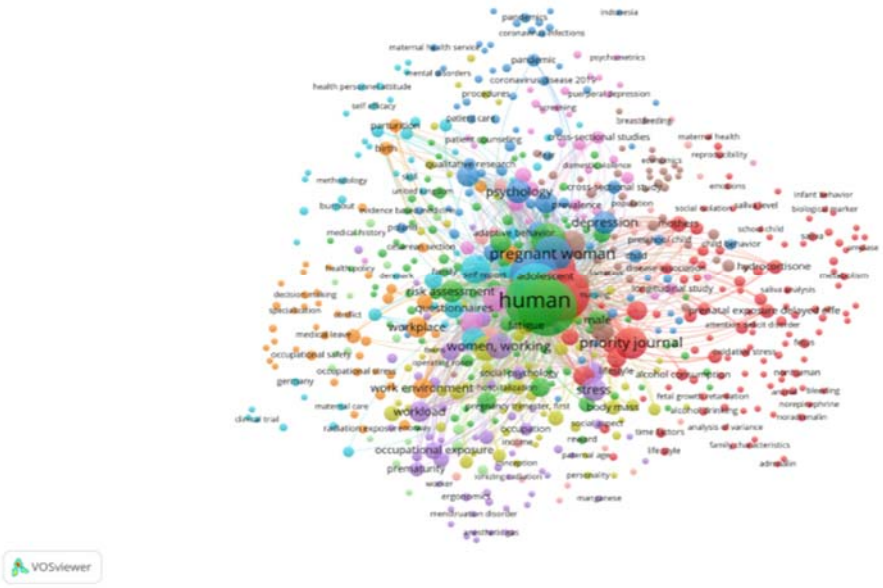


Figure-3.5. Show Co- Occurrence of Indexed Keywords

Conclusions

Between 2012 and 2021, a total of 500 papers were selected from the SCOPUS database on the subject of stress recognition, of which 231 were included in the SCOPUS core database. The majority of subsequent keywords had a strong connection to the research study's keywords "pregnant women" and "stress." Sensors (American Journal) was the most cited journal, with three articles and 269 citations. The United States of America published 60 publications that received 1635 citations, whereas Glover, V. published five articles that received 57 citations. The published material has mostly focused on three topics: stress, recognised stress, pregnant women, and human beings. Stress impacts individuals' daily lives, which is why various researchers have focused their attention on this subject and released several high-quality study publications. It is critical to examine the quality of several high-quality research articles and to extract useful findings. We will provide a model to predict the stress factors of pregnant working women.

References

1. Kobayashi, F., & Takeuchi, K. (2002). Sangyo eiseigakuzasshi = Journal of occupational health, 44(1), 1–5. https://doi.org/10.1539/sangyoeisei.kj00002552620European Economic Forecast-m Autumn 2020https://economictimes.indiatimes.com/magazines/panache/89-per-cent-of-indias-population-suffering-from-stress-most-dont-feel-comfortable-talking-to-medical-professionals/articleshow/64926633.cms?from=mdr

2. Stress: Signs, Symptoms, Management & Prevention. (2015, May 2). Cleveland Clinic.https://my.clevelandclinic.org/health/articles/11874-stressJan Andersen, in Research Management, 2018van Eck N.J.,

3. Waltman L. (2014)Visualizing Bibliometric Networks. In: Ding Y., Rousseau R., Wolfram D(eds) Measuring Scholarly Impact. Springer, Cham. https://doi.org/10.1007/ 978-3-319-10377-8_13

4. Selye, H. (1978). The Stress of Life (2nd ed.). McGraw-Hill Education

5. BJ Park, EH Jang, SH Kim, MA Chung , International Conference on Bio-inspired Systems and SignalProcessing (BIOSIGNALS-2014), pages 116-121 ISBN: 978-989-758-011-6 Copyright c 2014 SCITEPRESS(Science and Technology Publications, Lda.)Gedam,

6. S., & Paul, S. (2020). Automatic Stress Detection Using Wearable Sensors and Machine Learning:A Review. 2020 11th International Conference on Computing, Communication and NetworkingTechnologies (ICCCNT). Published. ttps://doi.org/10.1109/icccnt49239.2020.9225692

7. Albertetti, F., Simalastar, A., &Rizzotti-Kaddouri, A. (2021). Stress Detection with Deep LearningApproaches Using Physiological Signals. LectureNotes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, 95–111. https://doi.org/10.1007/978-3-030-69963-5_7

8. Goleva, R., Garcia, N. R. D. C., &Pires, I. M. (2021). IoT Technologies for HealthCare: 7th EAI InternationalConference, HealthyIoT 2020, VianadoCastelo, Portugal, December 3, 2020, Proceedings (Lecture Telecommunications Engineering Book 360) (1st ed. 2021 ed.). Springer.

9. Kusano, H., Horiguchi, Y., Baba, Y., & Kashima, H. (2020). Stress Prediction from Head Motion. 2020IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA). Published.https://doi.org/10.1109/dsaa49011.2020.00063

10. Can, Y. S., Arnrich, B., &Ersoy, C. (2019). Stress detection in daily life scenarios using smart phones andwearable sensors: A survey .Journal of Biomedical Informatics, 92, 103139.https://doi.org/10.1016/j.jbi.2019.103139

11. Nat. Volatiles &Essent. Oils, 2021; 8(4): 2304-23142314

12. Schmidt, P., Reiss, A., Dürichen, R., &Laerhoven, K. V. (2019). Wearable-Based Affect Recognition—AReview. Sensors, 19(19), 4079. https://doi.org/10.3390/s19194079

13. Shon, D., Im, K., Park, J. H., Lim, D. S., Jang, B., & Kim, J. M. (2018). Emotional Stress State Detection\\Using Genetic Algorithm-Based Feature Selection on EEG Signals. International Journal of EnvironmentalResearch and Public Health, 15(11), 2461https://doi.org/10.3390/ijerph15112461

14. Sevil, M., Rashid, M., Hajizadeh, I., Park, M., Quinn, L., &Cinar, A.(2021). Physical Activity and Psychological Stress Detection and Assessment of Their Effects on Glucose Concentration Predictions inDiabetes Management. IEEE transactions on bio-medical engineering, PP, 10.1109/TBME.2020.3049109.T

15. ervonen, J., Puttonen, S., Sillanpää, M. J., Hopsu, L., Homorodi, Z., Keränen, J., Pajukanta, J., Tolonen, A.,Lämsä, A., &Mäntyjärvi, J. (2020).Personalized mental stress detection with self-organizing

map: Fromlaboratory to the field. Computers in Biology and Medicine, 124, 103935.ttps://doi.org/10.1016/j.compbiomed.2020.103935

16. Dzieżyc, M., Gjoreski, M., Kazienko, P., Saganowski, S., & Gams, M. (2020). Can We Ditch FeatureEngineering? End-to-End Deep Learning forAffect Recognition from Physiological Sensor Data. Sensors,20(22), 6535. https://doi.org/10.3390/s20226535

17. Liao, C.-Y., Chen, R.-C., & Tai, S.-K. (2018). Emotion Stress detection using EEG signal and deep learningtechnologies. 2018 IEEEInternational Conference on Applied System Invention (ICASI)

18. Halim, Z., &Rehan, M. (2020). On identification of driving-induced Stress using electroencephalogram signals: A framework based on wearablesafety-critical scheme and machinelearning. Information Fusion, 53, 66–79.

19. Zhong H, Wang Y, Zhang ZL, et al. Efficacy and safety of current therapeutic options for COVID-19 -lessons to be learnt from SARS and MERSepidemic: A systematic review and meta-analysis. PharmacolRes 2020;157:104872. [Crossref] [PubMed]

20. Yu, Y., Li, Y.,Zhang, Z., Gu, Z., Zhong, H., Zha, Q., Yang, L., Zhu, C., & Chen, E. (2020). Abibliometricanalysis using VOSviewer of publications on COVID-19. Annals of Translational Medicine, 8(13), 816.https://doi.org/10.21037/atm-20-4235

21. Sharma Sameer, Sonal Sharma Bibliometric Analysis of Stress Recognition Models using VOSviewer. (2021). NATURAL VOLATILES & ESSENTIAL OILS.

22. Singh, R., Gehlot, A., Rashid, M., Saxena, R., Akram, S. V, Alshamrani, S. S., & Alghamdi, A. S. (2021). Cloud server and internet of things assisted system for stress monitoring. Electronics (Switzerland), 10(24). https://doi.org/10.3390/electronics10243133

Biographies

Dr. Sonal Sharma is Professor & head-Department of Computer Applications, Uttaranchal University with an experience of more 21 years. She possesses strong leadership and administrative skills. Her research interest is in Data Warehousing, Data Mining, Machine Learning, Big Data Analytics and Data Sciences. She has presented/published 45 research papers in national and international conferences/seminars/journals, out of which 11 are Scopus Indexed. She has presided over as a speaker in IEEE conference at Chengdu, China.

Mr. Praveen Shah is an Assistant Professor at Uttaranchal University, Dehradun, in the Department of Computer Applications. He has a total of more than 5 years of industry and teaching experience combined. He has worked with cutting-edge web development technology. Machine Learning and the Internet of Things are two of his research interests.

Mr. Sameer Dev Sharma is an Assistant Professor at Uttaranchal University in Dehradun in the Department of Computer Applications. Machine learning, artificial intelligence, deep learning, and Java programming are among his research interests. He has a combined industrial and educational experience of more than 16 years. He has a number of research papers that have been published in reputable international journals.

Mr. Abhishek Kumar Pathak is an Assistant Professor at Uttaranchal University in Dehradun in the Department of Computer Applications. Deep learning, IOT, Artificial Intelligence, and .Net Framework are among his research interests. He has a combined industrial and educational experience of more than 8 years. He has a number of research papers that have been published in reputable international journals.

SMTP Playground: A secure privacy-preserving web and mobile application for automation of sending and scheduling bulk emails to multiple receivers.

Tom Jose Oorasala, Ritesh Kumar Shukla, Sunkari Serena, Jayavignesh Thyagarajan

School of Electronics Engineering, Vellore Institute of Technology, Chennai, TamilNadu.

Abstract.

This paper proposes a web and mobile application named "SMTP Playground" to automate the process of sending bulk emails after extracting details of the potential receiver email address from an input file in the format of excel or comma-separated value (CSV). Unlike other automation scripting, this work also can send scheduled emails and encrypted emails to ensure privacy apart from the basic services provided by other email service providers in the market. This proposed Simple Mail Transfer Protocol (SMTP) Playground application provides a user-friendly interface and achieves the objectives.

Keywords— SMTP, E-Mail, End-to-End Encryption, Automation, Bulk Mails, Schedule Mails.

1. INTRODUCTION

A Mailing server is computer software that allows users to send emails to the ones who have registered with the server. It allows them to send text and data such as pictures, videos, and MP3s. However, there are some restrictions on the size of the attachments, which can be adjusted according to the server. This paper aims to develop a mailing system that will replace traditional email clients (such as Gmail and Outlook) [4]. The app's development is being done using Python and JavaScript. The following are some of the features incorporated are:

1. Notification Alert
2. End to End Encryption
3. Excel Sheet Batch Processing
4. Schedule Emails

SMTP (SIMPLE MAIL TRANSFER PROTOCOL): This is an application that web servers use for sending, receiving, and relaying incoming mail between email senders [3]. As soon as you compose an email and hit send, this is one of the most important stops on its journey to the inbox, and it delivers its message securely to the recipient. It's not uncommon to see SMTP servers on the Internet, but their function is highly specialized in handling and delivering email outbound [4]. It's a protocol that's part of the application layer.

IMAP (INTERNET MESSAGE ACCESS PROTOCOL)**:** This protocol operates at the application layer that allows email users to deliver and receive messages over a distant mail server. The contemporary model of IMAP is described through RFC 3501. IMAP listens on 143 whilst IMAP over SSL/TLS makes use of 993 by default [1]. An email patron makes use of one or extra of some of the email retrieval protocols to retrieve messages from an email server that shops the messages withinside the email box of the recipient. A person can send emails to and obtain emails from quite a few email servers. In assessment to a few customers and servers that favor using vendor-specific, proprietary protocols, almost all protocols consisting of POP3(Post Office Protocol 3) and IMAP2 are supported, permitting many specific email customers to get right of entry to those servers, and permitting those customers for use with many different servers as well. This feature of IMAP operation permits more than one customer to manipulate the equal mailbox [2].

FERNET ENCRYPTION: Fernet system uses industry best practices and adopts symmetric encryption/decryption. It also authenticates the message, allowing the recipient to determine whether or not the message has been tampered with since it was transmitted. Fernet is part of the cryptography library. To encrypt and decrypt data, a secret key is required, which must be shared by everyone who needs to transmit encrypted or receive decrypted data. Because anyone with the key can read and create encrypted messages, it must be kept secret from others. As a result, a secure way to distribute the key is a necessity. The same key could be used many times and instances [5-8].

2. RELATED WORK

There has been a study on the importance of encrypted mail. Since a lot of third-party applications can be logged in via a Google account, emails sent through Gmail are vulnerable. Hence, encryption is very necessary for the same. Around 360 billion emails are sent daily. There have been concerns about the security of the emails. Currently, the majority of email service providers prefer PGP3 and S/MIME2 standards for public encryption of keys. Both need a user's client to maintain his/her private key, and the public keys of the email senders and receivers [9]. Certain research papers have also highlighted the problems in encrypted mail. There have always been concerns about the circumstances if the keys reach the wrong hands. Some of the services like automatically adding meeting schedules through email won't be functional anymore as the system cannot read a receiver's emails. There will be difficulties for the system to categorize spam mails as everything is encrypted [10]. Few 3rd party applications are pre-existing in the market. Prevail mail, Proton mail are some of the examples. However, most of the existing services are paid or advertisement based which usually redirects the user to multiple fishy sites. Also, these 3rd party websites provide services either only in sending bulk emails or just sending encrypted mails.

3. FLOWCHART/ALGORITHM

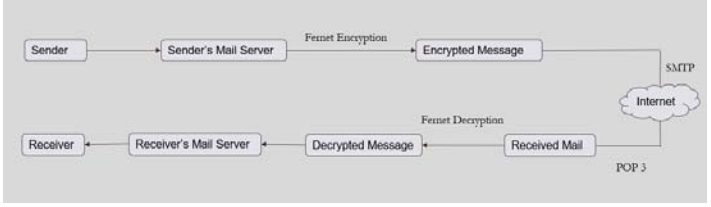


Fig.1 Dataflow of sending an e-mail

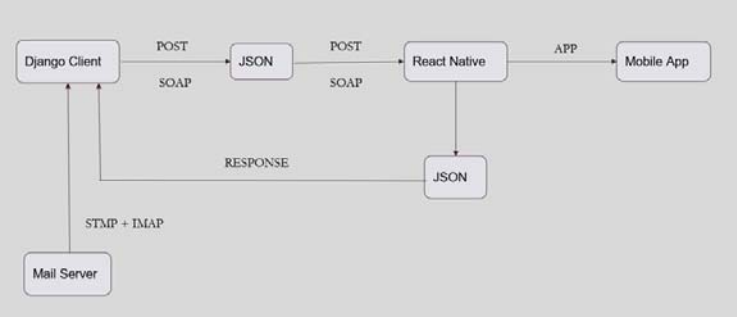


Fig.2 Dataflow between the mailing server, Django client, and the Mobile App

Fig.3 Dataflow between website and app.

Fig.2 and Fig3 shows, when a user composes a message or an email, it's normally done with the help of email clients.

4. SIMULATION/IMPLEMENTATION

Django is used to develop the website and the REST APIs. The biggest reason for using Django is the use of Python Language. Python language has lots of tools for SMTP and IMAP. It makes it easier for us to send emails and read emails. Also, python has some great tools for encryption and decryption which are highly secure. These modules are used in 2-factor authentications. There is a secret key that is stored in secret.key. The key is then loaded and used for encryption and decryption.

1) Pandas is used to run bulk mails.
2) SMTP server is set on.
3) Excel is then taken from the HTML input.
4) The email is parsed through the Pandas library.
5) An iterator runs through the email columns and sends emails to each address.

Scheduling emails is another interesting function incorporated inside the project. This feature enables users to schedule emails together and the e-mail with the given content will be sent at the given time. A server continuously monitors the time once the desired time is reached, the iterator stops monitoring and sends the mail [11]. React Native is used for the development of the Android Application. The REST API generated is used for communication between Django and React Native. The data is collected and sent through GET and POST requests of Rest API. React native is preferred over Flutter because it has a

scope to develop a desktop application in the future through electron module. The Contributed Code is available here: https://github.com/riteshshukla04/SMTP-PlayGround (Publicly Accessible).

5. RESULTS/INFERENCES

Fig. 5 is the web interface to send emails. The decrypted emails received by the user are shown in Fig. 6.

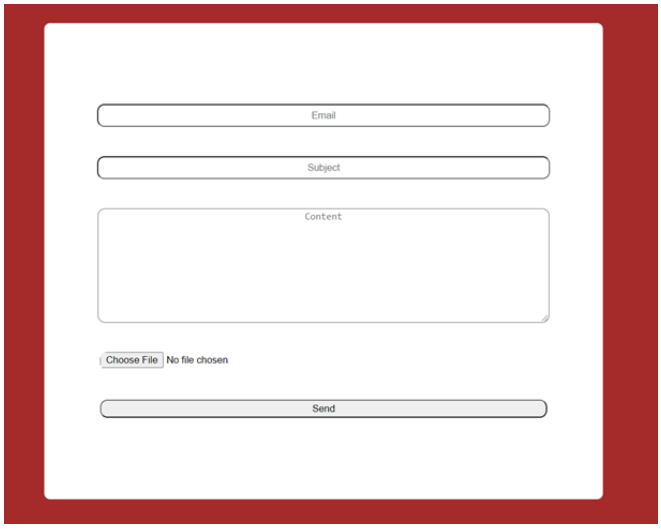


Fig.5 Web App Interface to send mails.

Fig.6 Mail sent and received on WebApp. The email is decrypted.

Fig.7 (1) shows how the email would look to a user when they use Gmail to read the encrypted mail. (2) Encrypted mail received in Gmail App.

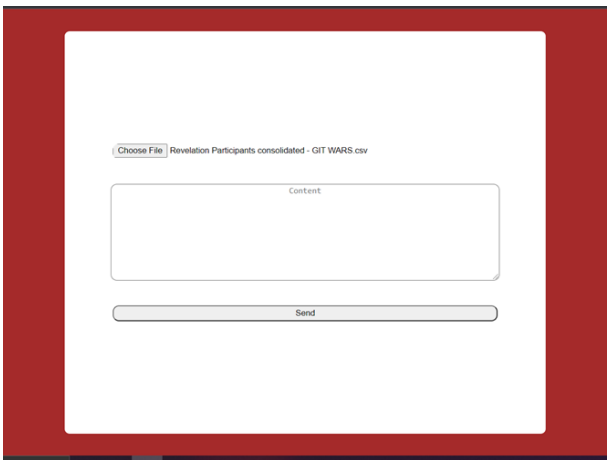


Fig.10 Interface for sending Bulk Mails

6. CONCLUSION

SMTP is the most convenient way for transmission to send messages via emails to recipients all over the globe. Sending emails in bulk is only possible through SMTP servers since it's a cost-effective way of sending emails to multiple receivers at the same time [11]. The SMTP connection is unrestricted and can be used to connect to any system. It is also beneficial to have a designated server to manage the outgoing emails in organizations. The amount of time and energy saved by using this system to send bulk emails is highly significant. For example, suppose a user has to send personalized emails to 500 users with the same content in the body, and assuming sending each email takes 10 seconds on average, it will take more than 1 hour to send them all. While using our system it just takes about 40 seconds, thereby saving the hassle of repeatedly composing new emails and the time spent in doing so.

7. FUTURE WORK

There is a large scope of UI improvements in the project. The UI can be more user-friendly and attractive. The UI should not only focus on looks but also the user experience. Future work involves hosting this website on premium hosting services like AWS or Digital Ocean. The second issue is that POP3 can't be used to read emails. This is due to the strict google security policy. Using POP3 to read mails would be faster, and also messages can be read

even if the user is offline. The end-to-end encryption security can be improved by increasing the length of the secret key. Randomizing and replacing keys regularly will reduce the chances of key misuse. To make the encryption more secure, more powerful, and robust, advanced encryption algorithms like AES 256 can be adopted.

8. REFERENCES

[1] Javatpoint.com, "Internet Message Access Protocol", 2021. [Online]. Available: https://www.javatpoint.com/imap-protocol [Accessed 23- Feb- 2022].

[2] Rhoton J. Programmer's guide to internet mail: SMTP, POP, IMAP, and LDAP. *Digital Press*; 2000.

[3] sendgrid.com, "What is an SMTP server" ,2019. [Online]. Available: https://sendgrid.com/blog/what-is-an-smtp-server/ [Accessed 22- Feb- 2022]

[4] Riabov VV. "SMTP (Simple Mail Transfer Protocol)", River College. 2005.

[5] Zadka M. "Cryptography", InDevOps in Python Apress, Berkeley, CA. pp. 95-110, 2019

[6] Schwenk J, Brinkmann M, Poddebniak D, Müller J, Somorovsky J, Schinzel S., "Mitigation of attacks on email end-to-end encryption". In Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security pp. 1647-1664, 2020

[7] Jain A, De P. Enhancing Database Security for Facial Recognition using Fernet Encryption Approach. In2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA) 2021 Dec 2 (pp. 748-753). IEEE

[8] Ferguson SJ, Hebels R. Computers for librarians: An introduction to the electronic library. Elsevier; 2003 Aug 1.

[9] M. D. Ryan., "Enhanced certificate transparency and end-to-end encrypted mail". In Network and Distributed System Security Symposium (NDSS), Feb. 2014.

[10] Yu J, Cheval V, Ryan M. "Challenges with End-to-End Email Encryption." In: Springer Reference (2014).

[11] Sukhija V., "Sending Email", InPowerShell Fast Track Apress, Berkeley, CA, 2022, pp. 65-69.

A Review on Watermarking Techniques Based on Deep Learning Neural Networks

Shikha Yadav

Jeevan Bala

Lovely Professional University, Phagwara; shkydv@gmail.com

Lovely Professional University, Phagwara; Jeevan.26699@lpu.co.in

Abstract.

A mechanized watermark is a kind of marking that is concealed in a noisy tolerant communication, such as video, sound, even picture. It's common to see imagined authorship for the patent of such transmissions. The process of concealing automated bits in a transport signal is known as "watermarking." There is no need that the protected information to be linked to the carrier signal. Automated watermarks can be used to verify the transport signal's legitimacy or soundness, as well as to demonstrate the proprietors' individuality. Another set of alterations affected by Deep Learning frameworks is examined in this research. Let's take a total examination of the different significant learning methods to introduce the watermark that can be processed to obtain a significant level of security for information communicated and power more than some aggression, especially when sending noisy media.

Keywords. Digital Watermarking, Robust, Deep learning, RNN, FCNN

1. INTRODUCTION

Advanced pictures are made by making a carefully encoded portrayal of an article's visual characteristics. They are used in demonstrating, computer games, engineering configuration, satellite photos, maps, and other things. The client may easily create, adjust, and communicate computerized photographs on account of present-day innovation. It is expected to get photographs during transmission against illegal use by an unapproved client. An unapproved client can hack and tempered the picture [1]. Therefore, the analysts should check their authenticity. Computerized Image Watermarking is the best method for guaranteeing the realness of advanced photographs. This is the innovation that safeguards computerized photographs, sound documents, and recordings. The most common way of watermarking advanced photographs is to hide data in them. pictures, films, and sound in computerized design. This approach was first established in 1992 by Andrew Tirkel and Charles Osborne.

There are two types of apparent and undetectable computerized picture watermarking. Watermarks on logos, back notes, and advanced photographs are instances of distinguishable watermarks that should be visible with ordinary eyes. In any case, it might become easy to eliminate these watermarks. Then again, imperceptible watermarks are encoded in a mystery design that must be gotten to by supported people. The recovery of these watermarks required numerical computations. Typical eyes can't see these sorts of watermarks. Imperceptible watermarks are tougher than those that are noticeable.

As of late, the headway of profound learning-based strategies has been found in picture watermarking procedures. The utilization of profound neural organizations in watermarking strategies has shown noteworthy outcomes.

In the field of computerized photo watermarking, this work examines fundamental neural organization-based watermarking calculations from top to bottom. The following are the major commitments of our work: - This research will aid experts in the disciplines of watermarking and deep progressing by providing a detailed examination of methods centered on "Digital image watermarking technique in deep learning." Because this study gives a complete survey of the accessible methods, different scientists may effortlessly dissect the numerous parts of the issue.

2. WATERMARKED STRATEGIES IMPLANTED IN DEEP NEURAL NETWORKS

Watermarking using a neural organization produces a good outcome, and this solution is more resistant to various assaults than other methods.

2.1 Watermarked methods inserted in Artificial Neural Networks (ANN)

Watermarking strategies considering neural organizations were effective. This is because of the way that the NN-based procedure beats a specific arrangement of theoretical assaults and functions admirably with the Human Visual System (HVS). Artificial neural networks (ANNs) are structures equipped to get the hang of, recall, and sum up explicit conditions and issues. ANNs are equal frameworks comprised of essential neurons of the unit that register explicit numerical capacities that are regularly non-straight and whose activity is roused by the capacity of the natural neuron. (Haykin ss 2009) [5].

Examination: In this method, the neural organization is prepared considering a subset of info pictures, inferring that this model just functions admirably with a subset of pictures and can't install or extricate the watermark from different pictures. Their prepared model is picture subordinate; the model recoveries picture-related data at the hour of watermark extraction. Thus, this approach is named semi-blind.

2.2 Watermarked procedure implanted in Convolutional Neural Networks (CNN)

The Convolutional Neural Network (CNN or ConvNet) is an exceptional discriminative profound gaining engineering that advances straightforwardly from statistics without the requirement for human mediation extraction of highlights. A Convolutional Neural Network (CNN), a sort of counterfeit neural organization, is extensively applied for photo coping with, division, and characterization. Picture denoising and super-intention are the 2

maximum sizeable regions of photo coping for similarly growing photo quality. Profound neural companies assemble making plans amongst spotless and loud pics to perform denoising precision. CNN's fruitful photo denoising accomplishments are credited to its excessive demonstrating restrict in-community making ready and plan. CNN with a profound layout offers greater noteworthy adaptability to attending to photo properties. Hemdan et al. [6] made a COVIDX-Net version that takes X-beam pics. The COVIDX-Net version became organized to make use of seven unmistakable CNN models, and it became checked to make use of 50 X-beam pics (25 every day and 25 COVID-19 cases).

Examination: Ahmad et al proposed [7] CNN's are constructed from 3 sizeable forms of layers: convolutional, pooling, and related. The convolution layers use channels/elements whose coefficients are altered throughout the coaching level to research large highlights that could occur withinside the statistics. Each channel is convolved autonomously over the contribution to create a factor map, with better enactment values demonstrating the region of factors.

2.3 Watermarked procedure utilized in Fully Convolutional Neural Networks (FCNN)

The instar-out big-name version is a convolutional neural enterprise with 3 layers: facts, result, and grouping. This version is used to devise the input-yield data. Bansal et al. [8] used a neural agency version with one facts layer, one mystery layer, and one results layer of their setup technique. The cowl photo is first inserted withinside the backpropagation neural agency, and the version's masses have then adjusted the use of the backpropagation technique and the goal watermark. The writer then used spatial methods [9] to combine the altered masses into the duvet photo. The very last watermark image is hooked up with the expected results watermark to decide the PSNR of the acquired watermark image.

Investigation: Instead of the use of the duvet image, a fully convolutional neural network (FCNN) changed into used to contain the watermark inner FCNN relationships. This aided in extending electricity and decreasing minor problems to a potential degree. As a result, the watermarked photo is like the original. Furthermore, most assaults didn't degrade the great of the recovered watermark photo, notwithstanding the truth that connected layers are computationally expensive. Because best the watermarked image is needed to split the watermark now of extraction, this the watermarking technique is visually degraded.

2.4 Watermarked strategy installed in Deep neural Networks (DNN)

Deep neural networks (DNNs) have arisen as a huge device for carrying insight to portable and installed gadgets. As DNN models become even more broadly conveyed, shared, and possibly marketed, there is an expanding interest in licensed innovation security (IP) assurance. As of late, DNN watermarking has arisen as a feasible choice. Technique for licensed innovation security. To empower DNN watermarking on implanted gadgets in a sensible situation, a discovery method is required. Existing DNN watermarking systems either don't fulfill or don't meet the necessities. black-box prerequisite or are helpless against a few sorts of attacks. Jia Guo et.al [10] propose a watermarking system that incorporates the creator's mark during the preparation of DNNs in run-of-the-mill conditions, the resultant watermarked DNN works typically. Whenever given any marked item, it acts in a different, preset design input, hence, showing creation.

Examination: By examining the ability of profound neural organizations in the work of combining the appropriated image and the watermark's dormant areas, it has been discovered that the suggested structure has built a picture combination application on picture watermarking. The impediment of picture combination frameworks is that this strategy could bring about huge information misfortune. Besides, when an image is melded, data planning is fairly directed.

Table 1. Comparison of Deep Learning Techniques

Ref No.	Embedded Technique	Features	Visual Imperceptibility	Robustness Attacks considered	Research Gaps
[2]	The YCbCr color space, IWT (integer wavelet transform), and DCT (discrete cosine transform) are used to create a blind and resilient system.)	used in robust applications (i.e., copyright protection) for efficient results and less computational time.	PSNR = 40.25 Db, SSIM = 0.9976	NC = 1.000 **Attack Considered** Signal processing attacks like cropping, JPEG Compression, resizing	The embedding strategy is not efficient in neural networks architecture to achieve optimal performance
[3]	The lightweight convolution neural network (LW-CNN) technique is used	Reduced the calculation time and made the system more robust to current attacks.	SSIM = 0.98	NCC = 0.998 BER = 0.31 **Attack Considered** Noisy Domain like average, Gaussian, median	Less robustness on translation and rotational type attacks
[4]	Used Spread Transform Dither Modulation (STDM) and Spread Spectrum (SS)	Denoising with a Fully Convolutional Neural Network (FCNN) maintains picture quality while compromising resilience.	SSIM = 0.987	BER=0.342 **Attack Considered** Salt And Pepper, Gaussian filtering, median filtering, Weiner filtering, average filtering	Used in other types of attacks

Table 1. shows, an audit of past work in advanced watermarking methods are given. A few well-known approaches that have been examined in the past incorporate spatial area and recurrence space procedures. Besides, the spatial area computerized watermarking innovation is less powerful and subsequently less suggested. Power, indistinctness, security, and limit are utilized to assess the watermarked picture's exhibition. The visual subtlety of the watermarked picture and the strength of the watermarking were two of the main elements. Moreover, future work could profit from joining techniques and applying them in a half and half structure to work on the vigor of the watermarked picture, yet in addition to diminishing the inconveniences of every strategy separately.

3. CONCLUSION

We have characterized a thorough conversation on a few profound learning-based advanced picture watermarking frameworks in this work, which remembers flow improvement and development for this area. All the current profound learning-based advanced picture watermarking calculations enjoy benefits and impediments. Our examination investigates different neural organization-based philosophies and analyses them as far as vigor and indistinctness. Many issues stay annoying and perplexing for deciding the strength of the watermarked picture against attacks, for example, equivocal assaults, conspiracy assaults, crossbreed assaults, and managing more than one assault simultaneously, as expressed in the above areas. To fortify watermarking frameworks for advanced pictures, it is basic to resolve these issues.

4. REFERENCES

[1] Lydia, E. L., Raj, J. S., Pandi Selvam, R., Elhoseny, M., & Shankar, K. (2021). Application of discrete transforms with selective coefficients for blind image watermarking. Transactions on Emerging Telecommunications Technologies, 32(2), e3771.

[2] Dhaya, R. (2021). Lightweight CNN-based robust image watermarking scheme for security. J. Inf. Technol. Digit World, 3(2), 118-132.

[3] Hatoum, M. W., Couchot, J. F., Couturier, R., & Darazi, R. (2021). Using Deep learning for image watermarking attack. Signal Processing: Image Communication, 90, 116019.

[4] Meenpal, A., Majumder, S., & Balakrishnan, A. (2020, January). Digital Watermarking Technique using Dual-Tree Complex Wavelet Transform. In 2020 First International Conference on Power, Control and Computing Technologies (ICPC2T) (pp. 62-67). IEEE.

[5] Haykin, S. (2009). neural networks and learning machines / Simon haykin. New York, USA: Prentice-Hall.

[6] Hemdan, E. E. D., Shouman, M. A., & Karar, M. E. (2020). Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images. arXiv preprint arXiv:2003.11055.

[7] Ahmad, J., Farman, H., & Jan, Z. (2019). Deep learning methods and applications. In Deep learning: convergence to big data analytics (pp. 31-42). Springer, Singapore.

[8] Ashish Bansal, Sarita Singh Bhadauria, “Watermarking using neural network and hiding the trained network within the cover image”, Journal of Theoretical and Applied Information Technology,2008

[9] Yang C, Zhu C, Wang Y, “A Robust Watermarking Algorithm for Vector Geographic Data Based on Qim and Matching Detection”, Multimedia Tools Application, 2020.

[10] Guo, J., & Potkonjak, M. (2018, November). Watermarking deep neural networks for embedded systems. In 2018 IEEE/ACM International Conference on Computer-Aided Design (ICCAD) (pp. 1-8). IEEE.

Biographies

Shikha Yadav received a bachelor's degree in computer applications from Delhi University in 2014, a master's degree in computer applications from Maharshi Dayanand University in 2018, and pursuing the philosophy of doctorate degree in Computer applications from Lovely Professional University. Her research areas include digital watermarking, deep learning, and social network analysis.

Jeevan Bala, assistant professor, completed her doctorate degree in Computer Science and Engineering with a specialization in machine learning. She is a dedicated researcher and an experienced academician with research papers in prominent journals, including publications in SCI/SCIE indexed journals. Her research interest includes digital image processing, machine learning, and meta-heuristic techniques.

Comparative Analysis And Simulation Of Internet Routing Protocols In Different Network Topologies Using Netsim

Selvakumar S[1], Naresh T K[1], Dinesh A[1], Jayavignesh T[1]

[1]School of Electronics Engineering, Vellore Institute of Technology, Chennai Campus
{selvakumar.s2019@vitstudent.ac.in;naresh.tk2019@vitstudent.ac.in;
dinesh.a2019@vitstudent.ac.in; jayavignesh.t@vit.ac.in}

Abstract.

Router, a networking equipment, is critical to the intelligence of the internetwork. Routers in the internet receives data packet via its interfaces in same or another neighbouring network and forward it to another or same network. To which network it should pass the data packet is the place where routing protocols comes to play. This decision is based on metrics. The main purpose of this study is to analyze the performance of the routing protocols namely RIP and OSPF on three different topologies like Ring, Star and Mesh. The simulator used is NetSim to study the behaviour and performance comparison of these protocols in various network topologies and network metric such as application throughput, packet delay is analysed in different scenarios such as link failure etc.

Keywords: NetSim, RIP, OSPF, Ring, Star and Mesh.

1. INTRODUCTION

Internet protocols is used for end-to-end communication for inter connected networks and specifies how data should be transmitted to the destination. TCP/IP comprises of four layers and all layers have its own functionalities. In this paper we are concentrating only the function of internet layer which is enabled by router. Router is a device which forward data packets to one or more router or network. When a router receives a data packet, it reads the address of destination and forward it to the path which take the packets to respective address. Here the routing table comes to the play. A routing table is a table of data stored in a router that contains routes to a particular network and associated metrics (cost or distance). Building a routing table is an important goal of a routing protocol. In this paper we focused mainly on two protocols namely RIP and OSPF. Routing Information Protocol (RIP) works on the basis of distant vector algorithm which uses hop count as a primary criterion to choose path. Open Shortest Path First Protocol (OSPF) works on the basis of link state routing algorithm in which least cost route is defined by considering various parameter such as cost of link and condition of link (up and down). This can be calculated using the formula "OSPF path cost = ((auto-cost × reference-bandwidth + interface bandwidth) -1) / interface bandwidth". A Routing table contains network destination address, netmask, gateway, interface, metrics. These protocols choose the next interface based on the metrics in the table. The first topology is the ring topology. It is the network configuration where the devices connected in a network form a circular data path like a ring. The star topology has head router to which all other routers are connected like a start. In mesh topology, each and every node has a unique path to all other nodes in a network. The work related to our study is listed in Section 2. In Section 3, details about the simulation environment which is NetSim in our case and the parameters we have considered are discussed and the results of our

simulation and our inferences from the results are shown in Section 4. Finally, in Section 5, we arrived at the conclusion by comparing performances.

2. RELATED WORKS

In the previous study at [1], they have compared three protocols for an enterprise level designed topology and analysed the performance. They have concluded that EIGRP is the best protocol as it has least convergence time and better delay time compared to others. Albaour A et al. [2] have simulated for a fully connected mesh topology and claim that EIGRP has faster throughput than RIP while OSPF has the fastest throughput among them in all the stages and the distance vector-based RIP has the highest queuing delay while in the beginning EIGRP had higher queuing delay than OSPF, but then gradually declined and became the least in terms of queuing delay. We can also see that in terms of convergence rate, EIGRP has the fastest convergence in all phase. While OSPF has a faster convergence time than RIP. And this literature doesn't deal with any simulations related to any link failure cases or work on different topologies. And protocols like RIPv1, RIPv2, EIGRP, OSPF were analysed using network simulated in cisco packet tracer [7] and concluded that EIGRP is better than all other protocols in terms of convergence time and from the point of view of throughput and delay, OSPF dominates RIP.

Athira M et al. [3] compared the performance of these routing protocols for real time connections with help of an enterprise network topology. This outcome indicated that, in OSPF and EIGRP the delay is lesser than in RIP. And in terms of throughput, OSPF is better than RIP. In other study [4] three protocols are compared for Ring, Mesh and Big Mesh topologies. They have concluded that, in terms of convergence time the EIGRP performed well than other protocols in different topologies and uses bandwidth efficiently. In another study at [8] states the EIGRP uses DUAL which avoids routing loop. So, convergence time is less in star and mesh topologies and also concludes that the mesh topology is the best topology in terms of convergence time. In the case of link failure, the EIGRP performs well in terms of convergence time as it adds new links when a link fails. The performance of OSPF is lesser than EIGRP. In the ref. [9] the performance of RIP and OSPF protocols are compared in two different network simulators which are ENSP and GNS3. And the analysis indicates, the connection time of OSPF routing protocol (83 ms) is faster than RIP (177 ms). And GNS3 network simulator had 329 ms average time, while the ENSP has 94 ms.

3. SIMULATION ENVIRONMENT AND CONFIGURATION PARAMETERS

Netsim is network simulation software for network design verification, network development, and research. It's C-based software with a package of intelligent libraries that allows you to emulate not only simulations, but also merging real and virtual worlds. We can also change routing parameters in RIP like update timer, timeout timer and garbage collection time and in OSPF, link state refresh time and maximum age. It also allows us to use options like event trace, packet trace and an animation window. We have compared the performance with and without link down situation. For creating a link down situation, we have reduced the bandwidth (upload speed) of link.

3.1. RING TOPOLOGY

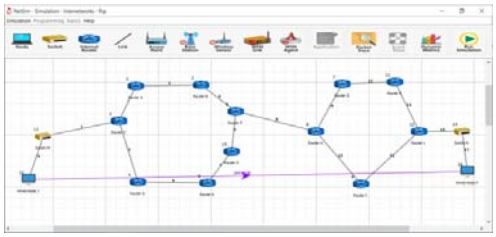


Figure 3.1.1 Ring Topology Internetwork

We have made a network on ring topology. This network (figure 3.1.1) has two rings made with 12 routers (A to L). The purple line indicates the source to destination transmission. Using RIP, it followed the path with minimum hop count of 7. The selected path is (Source→ Switch →C → A→ B→F → H → I → L → Switch → Destination). The throughput is 0.581664 (Mbps) and the delay is 13188.437912 (microsec). The selected path using OSPF protocol is (Source → Switch →C→A→B→F→H→G→K→ L → Switch → Destination). The throughput is 0.579912 megabits per second and the latency is 1420.232749 microseconds. To analyse it further we changed the link properties of link ID 10 (1Mbps) and 12 (50Mbps). Now the path changed to (Source→ Switch →C → A→ B→F → H → I → L → Switch → Destination), which is similar as the route of RIP. But there is not much difference in throughput. So, we added one more source node. Upload speed of link 12 was reduced from 50 to 1Mbps and kept link 10 high. For RIP, the applications throughput and delays are shown in the Fig 3.1.2 and for OSPF in Fig 3.1.3.

Application_Metrics_Table

Application_metrics — Detailed View

Application Id	Throughput Plot	Application Name	Throughput (Mbps)	Delay(microsec)
1	Application_throughput_plot	APP1_CBR	0.487640	1635197.378323
2	Application_throughput_plot	APP2_CBR	0.486472	1647010.035654

Figure 3.1.2 Throughput using RIP

Application_Metrics_Table

Application_metrics — Detailed View

Application Id	Throughput Plot	Application Name	Throughput (Mbps)	Delay(microsec)
1	Application_throughput_plot	APP1_CBR	0.578160	1420.233454
2	Application_throughput_plot	APP2_CBR	0.582248	1542.311815

Figure 3.1.3 Throughput using OSPF

3.2. STAR TOPOLOGY

The star topology is created using 5 routers (A to E) as in the Fig 3.2.1.

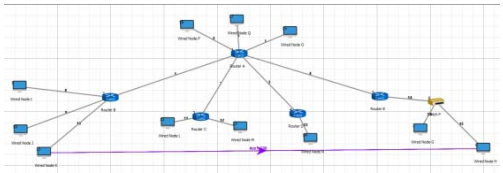


Figure 3.2.1 Star topology Internetwork

Using RIP, the selected route is (Source → B →A →E → Switch →Destination). The throughput and delay are shown in the Fig 3.2.2 and using OSPF in Fig 3.2.3.

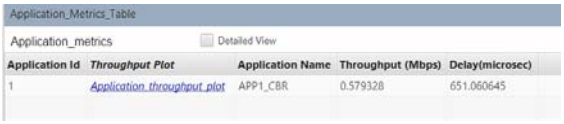


Figure 3.2.2 Throughput using RIP

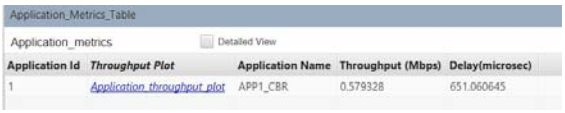


Figure 3.2.3 Throughput using OSPF

3.3. MESH TOPOLOGY

This network is built using 5 routers (A to E) as shown in the Fig 3.3.1.

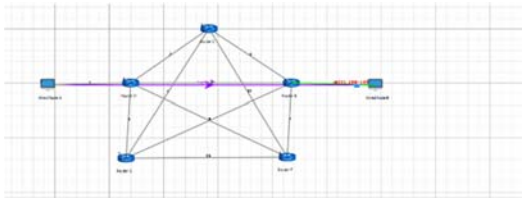


Figure 3.3.1 Mesh topology

In RIP the selected route with least hop count of 3 is (Source → D →E → Destination). The throughput and delay are shown in the fig 3.3.1. Using OSPF, the route followed is (Source → D →C → E → Destination) and results are in the Fig 3.3.3.

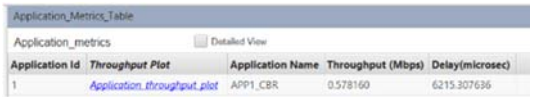


Figure 3.3.2 Throughput using RIP

Figure 3.3.3 Throughput using OSPF

4. RESULTS & INFERENCES

For the ring topology internetwork, the application throughput and delay in RIP are 0.48 Mbps and 1641103.7069 microseconds on average respectively. Using OSPF, the application throughput and delay are 0.58 Mbps and 1481.2726 microseconds on average respectively. The average application throughput via OSP is better than in RIP and the average end-to-end delay is also lesser in OSPF. For the star topology internetwork, if there

is any link failure, it has no other route to communicate. So, the application throughput and delay will be same in both the cases. For the mesh topology internetwork, the application throughput and delay using RIP are 0.578 Mbps and 6215.3076 microseconds respectively, and using OSPF, 0.579 Mbps and 502.280 microseconds. The average end-to-end delay is comparatively lesser using OSPF than RIP.

5. CONCLUSION

The RIP performs well in all the three network topologies if there is no failure in link. But in case of a link with lesser bandwidth, the RIP still follows the route with lesser bandwidth which increases delay whereas the OSPF follows the alternate route with higher bandwidth (upload or download speed). This makes OSPF better in every link failure case in both ring and mesh topology. But in star topology, both the protocols perform similar as there is no alternate path. Overall, OSPF performs well in all topologies.

6. REFERENCES

[1] Biradar AG., "A Comparative Study on Routing Protocols: RIP, OSPF and EIGRP and Their Analysis Using GNS-3", In 2020 5th IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE) Dec 1, 2020, pp. 1-5.

[2] Albaour A, Aburawi Y, "Performance Comparison of Routing Protocols. International Research Journal of Engineering and Technology (IRJET)., Vol 8, No. 2, 2021, pp:1-5.

[3] Athira M, Abrahami L, Sangeetha RG, "Study on network performance of interior gateway protocols—RIP, EIGRP and OSPF", In 2017 IEEE International Conference on Nextgen Electronic Technologies: Silicon to Software (ICNETS2) Mar 23, 2017, pp. 344-348

[4] Mirzahossein K, Nguyen M, Elmasry S. "Analysis of rip, OSPF, and EIGRP routing protocols using opnet". Simon Fraser University, School of Engineering Science. 2013.

[5] Masruroh SU, Widya KH, Fiade A, Julia IR. "Performance evaluation dmvpn using routing protocol rip, ospf, and eigrp", In 2018 6th IEEE International Conference on Cyber and IT Service Management (CITSM) 2018 Aug 7, pp. 1-6..

[6] Bhatt T, Kotwal C, Chaubey N. "Implementing and examination of EIGRP OSPF RIP routing protocol in AMI network for DDoS attack using OPNET" International Journal of Recent Technology and Engineering (IJRTE),Vol 8, No 2 2019, pp:3776-3783.

[7] Wai KK. "Analysis of RIP, EIGRP, and OSPF Routing Protocols in a Network. International Journal of Trend in Scientific Research and Development (IJTSRD). Vol 219, No 3, 2019;.

[8] Okonkwo IJ, Emmanuel ID. "Comparative study of EIGRP and OSPF protocols based on network convergence" International Journal of Advanced Computer Science and Applications.; Vol11, No 6, 2020, pp:39-45.

[9] Majid NW, Fuada S. Simetris: RIP VS. OSPF Routing Protocols: Which one is the best for a real-time computer network? Jurnal Teknik Mesin, Elektro dan Ilmu Komputer.;Vol 11, No 1, 2020 Apr 30, pp:249-56.

[10] Mesbah A, tebbi, "Simulation And Comparison Between OSPF & RIP Routing Protocols", Doctoral dissertation, Université Mohamed el-Bachir el-Ibrahimi Bordj Bou Arréridj Faculté de Mathématique et Informatique, 2021.

Biographies

Selvakumar S, undergraduate in Electronics and Computer Engineering from VIT University and will be graduating in 2023. His areas of interest include emerging and new technologies in the field of communication and data analytics.

Naresh T K, undergraduate in Electronics and Computer Engineering from VIT University and will be graduating in 2023. He is an ardent reader of IEEE journals and his areas of interest include emerging and new technologies in the field of wireless communication.

Dinesh A, undergraduate in Electronics and Computer Engineering from VIT University and will be graduating in 2023. His areas of interest include wireless communication and data analytics.

Jayavignesh Thyagarajan is currently a Senior Assistant Professor in the School of Electronics Engineering, Vellore Institute of Technology, Chennai Campus. He received his Bachelor's degree in ECE from CEG, Anna University in the year 2007. He received his Master's Degree in Communication & Networking from Madras Institute of Technology, Anna University in the year 2010.He obtained his PhD from Vellore Institute of Technology, Chennai in the year 2020.His research interests lies in the area of Wireless Sensor Networks, IoT and focuses on design on Low Power Communication Protocols, Cross Layered Routing design in wireless multi-hop networks such as ad hoc, sensor, vehicular, mesh and flying ad hoc networks.

SOCIO-PSYCHO EFFECT OF INTELLIGENT SYSTEM ON SOCIETY

Radhey Shyam Jha[1], Nidhi Tyagi[2], Prabha Lama[3], Kumar Ashutosh[4], Amit Kumar[5]

[1]Law College Dehradun, Uttaranchal University Dehradun Uttarakhand, radhey.shyamjha@gmail.com
[2]Department of H&SS, Graphic Era University Dehradun Uttarakhand, tyagi1234nidhi@gmail.com
[3]Department of H&SS, Graphic Era University Dehradun Uttarakhand, prabhalama15@gmail.com
[4]Law College Dehradun, Uttaranchal University Dehradun Uttarakhand, ashutoshjhalcd@gmail.com
[5]Law College Dehradun, Uttaranchal University Dehradun Uttarakhand, amitsinghdahiya03@gmail.com

ABSTRACT

As a result of the introduction of Technology equipped with intelligence, especially AI (Artificial Intelligence), the very social and psychological mechanism that results in the continuity and stability of the society has begun to subside and with it the stability and sanity of the society will get ripped out. It will result in creating sociopathic, psychopathic, unstable, depressed, artificial human beings with no sense, no patience and no sanity. If this state of affairs will continue the human beings will crave for love, crave for emotions and crave for attention. They may pretend themselves as harbingers of advanced technology and computerized world but they will be dying every single moment in the absence of societal bonds and social relationships. The relationship between humans, automated machines and their future impact on society has been a key issue for researchers in the field of Psychology, Economics and Sociology. Through this paper an attempt has been to throw light on social and psychological effects of technology on human society in general and Indian society in particular.

Key Words: Artificial Intelligence, Machine, Psychosomatic disorders, Technology, Unemployment

1. INTRODUCTION

"Technology will destroy this planet mentally, if responsible individuals do not come forward to advocate for responsible use of technology."[2] Today we have become like the most primitive man, once more global wanderers, but information gatherers rather than food gathers [7]. Software systems and increasing use of Artificial Intelligence have increasingly threatened managerial and professional jobs. Technology including AI or Intelligent System and new software technologies are shouldering more responsibilities on machines replacing human beings which have increased human anxiety and depression [1, 8]. 25% in US, 15% in UK and approx. 30% jobs in India are at risk due to increasing use of technology [4, 9]. As per the NCRB data and the report submitted by the Ministry of Home Affairs, total of 3,548 people in 2020, 2,851 in 2019 and 2,741 in 2018, 2,404 in 2017 & 2298 in 2016 committed suicide due to unemployment. 24% increase has been recorded in 2021 [1, 6, 8]. Researches from Villanova University and Ball State University examined the data at country and concluded that a 20% of increase in technology will increase 30% oh health risks [11]. As per the reports of the Research it was clearly stated that almost 50% of the jobs could be at risk of technology especially of

Artificial Intelligence by 2033 [4, 9] which will result in psychosomatic disorders (stress, strain and anxiety like hypertension, Blood pressure, diabetes mellitus etc.). Technology affects human beings both psychologically and physiologically thereby bringing about a change in the social structure of society, which also brings about a number of social and cultural changes the repercussions of which are faced by the people. The introduction of technology leads to transformation of human existence in different spheres be it social, biological, psychological or cultural [10]. Intelligent System in the form of computers, Computer Programs, Artificial Intelligence etc. affect the people more who are mentally and psychologically not prepared to such type of changes in technology. However, for manual workers it is not much harmful as this new form of mechanization was mere an extension of the work which they were doing earlier but in case of office personnel who encounters such development which has menaced their status by bringing it closer to machine operator. People are not assured that the benefits of this technological advancement will be distributed equally among them or not. Human beings placed indifferent sectors industries, working class, who are engaged in plants accrue the hardships of technology because of their replacement by machines. The perspective that technology will solve each and every problem of the society is in fact digital fundamentalism which is far more harmful than religious fundamentalism.

1.1. Research Methodology: Descriptive research methodology has been used to gather the data in order to describe the phenomenon with its causes and possible effects on society. Both Primary and secondary tools of data collection have been used. The sample size is selected on the basis of random sampling technique from among the technologically advanced prone sectors of capital city Dehradun and its neighboring cities. The study conducted on 767 respondents with respect to social and psychological impact of technology.

1.2. Hypothesis: Hypothesis framed, are as follows

i. Technology has resulted in psychosomatic disorders.
ii. Technology has caused societal issues like suicide, accidents and family conflicts.

1.3. Objectives of the Study: these are:

1. To study the social and psychological effect of technology on human societies.
2. To spread awareness among the people about the limitations of technology.
3. To assess the damage caused by technology to society and mankind.

2. HOW TECHNOLOGY IS AFFECTING THE SOCIETY & HUMAN LIVES SOCIALLY & PSYCHOLOGICALLY?

Technology has great psychological impact on the people, involved in the process. At first, it has an emotional reaction with respect to new technological methods in general and impact of technological innovations in particular and secondly, the reaction of the people who are confronted by new working and living conditions is exposed to psychological stress and strain.

2.1. Impact on Professional Behavior: The psychological impact of technology, advanced, Intelligent System on human beings can be traced in the following ways:

- **Occupational Strain**: Long monotonous working hours with machines or computers directly affects the mind and results in stress and promotes strain [1, 8]. Increasing use of Artificial Intelligence in every sector is causing panic among the working class; even judiciary is also having the same fear.
- **Stress caused by Biometric Machines at workplace:** Using Intelligent System like biometric software to mark the attendance or calculate the performance of the employees has become many times not only finicky but results in stress, strain and accidents thereby. [5, 8].
- **Increase in the ratio of Anxiety**: The increasing human anxiety and depression due to unemployment caused by the increased use of machines, especially in the form of job insecurity, has resulted another kind of socio-psycho disorder [1, 8].
- **Frustration, Aggression & Inferiority Complex**: Irritation causes frustration while aggression is triggered by frustration. Artificial intelligence, robots, bellhop robots, software algorithms, self-driving vehicles etc. are replacing man and made the situation more panicky. Frustrated and aggressive behavior is its outcome [1, 8].
- **Self-Homicide:** As per the NCRB data and the report submitted by the Ministry of Home Affairs, total of 3,548 people in 2020, 2,851 in 2019 and 2,741 in 2018, 2,404 in 2017 & 2298 in 2016 committed self-homicide (suicide) due to unemployment in which 24% increase has been recorded in 2021 [6]. The most important reason of this was substitution of man power by machine power. Online games have also contributed in self-homicide. As per the report of India Times, the 'Blue Whale Challenge' online game caused over 130 deaths while 'The Pass Out Challenge' caused 250 to 1,000 deaths in one year in U.S [5].

2.2. Social Effects of Technology: It can be understood as follows:

Effects on Family Life	Disintegration of joint family system; Late and Inter caste marriages issues; Increasing number of divorces and squabbles; Infanticide by using technological advancement; no identity of marriage institution; Loss of Sex mores encouraging computer pornography, Lessen the role of family as an agency of socialization, Invasion on privacy by vast computer data networks [3, 8, 10].
Effects on Economic Life	Birth to capitalism and its attendant evils; Caused economic depression; Unemployment & industrial disputes; Increase of accidents; disease and downgrading and monotony of jobs; Slavery of machines; Huge capital loss in technology and its regular upgradation [9].
Effects on Socio-Political Life	Commercialized recreation; Increase in individualism, Hedonism & emotional instability. Artificial socio-religious relationships. Threat to nation and nationalism; Increase in undue influence of pressure groups on state; Increase in size and power of bureaucracy [3, 8].

2.3. Social Isolation: Addiction to technology makes a person socially isolated with zero emotion and socialization that causes depression in a sense. Being associated with science and technology we feel that we are disconnected with humanity [3, 10].

2.4. Deviance & Anomie: Day and night with computers, surrounded by heterogeneous and complex rules of behavior. Capitalism, Cultural lag, impersonality of relationships, predominance of individualism and a

highly mechanical life are all the attributes of technology which results in various mental and emotional disorders. Neurosis, psychosis, insomnia etc. are its outcomes which further results in anomie and deviance.

2.5. Miscellaneous Effects on Society: Distorted Exclusiveness of social groups; Dehumanizing working conditions; Isolating social forms inhibiting ethical relations; Forced participation in social processes; Loss of family role; Anti-Intellectualism &Vulnerability of social systems [3]. Though the system is so called 'Intelligent' but it is having zero emotional quotient with 'data input' based decision maker only.

3. TECHNOLOGY & HUMANITY: AN ANALYSIS

To assess how much technology equipped with Intelligent System has affected the world we live in, a survey on socio- psycho impact of technology and increasing use of Artificial Intelligence in different sphere of life was conducted in January 2022. The data was collected using random sampling technique from 767 respondents. Various Questions were asked and the responses recoded thereon. The response and its analysis is as follows.

Table-1: Fact Sheet of responses recorded

Questions / Responses	Paralyzed the Society	Increased risk of depression	Increased job insecurity	Caused psychosomatic diseases	Resulted in social trauma	Digital fundamentalism more fatal	Shattered societal organization	Caused ecological imbalance	Exceeded humanity	Technology a dangerous master
Strongly Agree	75	59	48	50	31	43	48	45	41	29
Agree	10	16	30	19	20	23	26	29	17	43
Neutral	11	13	10	21	30	7	23	16	21	13
Disagree	3	10	10	7	7	15	2	4	9	4
Strongly Disagree	1	2	2	3	12	12	1	6	12	11

Source: Survey on socio-psycho impact of technology2022

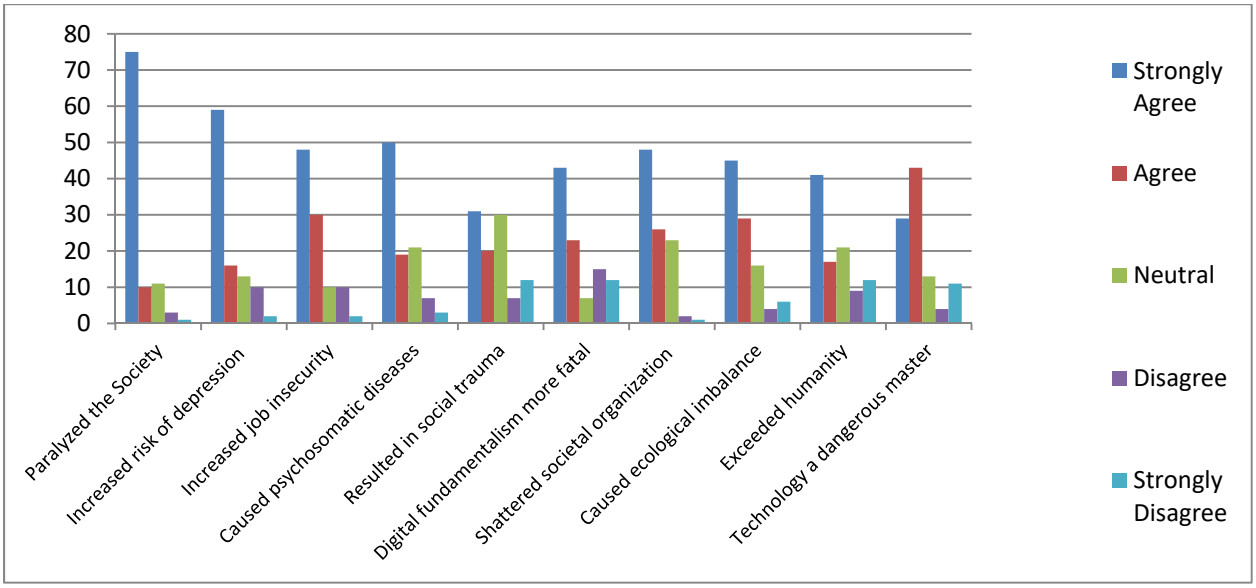


Figure-1: Graphical representation of responses

3.1. Data Interpretation

On being asked whether technology in the form of Intelligent System has spread its wings to the extent that it has paralyzed human society and human lives or not, out of total no. of respondents (767) 75% strongly agreed that technology has actually handicapped human beings and human life .When the respondents were asked whether technology has changed socio- psychological behavior of men in present scenario or not, 59% of the people believed that technology is socially and psychologically affecting the behavior of people while only 2% strongly disagreed. Technology equipped with Artificial Intelligence has already replaced human being with machines and in future it will shake the foundation of economy by taking over the jobs. Out of total no. of respondents 48% agreed that technology equipped with Artificial Intelligence is not only replacing human labor but also creating insecurity in the minds of people whereas only 2% denied the same. Working together day and night with machines, human being has themselves become mechanized which has brought about host of many diseases but psychosomatic disorders. The responses highlighted that 50% of the people strongly agreed to the viewpoint that there is a close connection between technology and psychosomatic diseases which gives birth to hypertension, neurotic disorders, cardiac arrests etc. Hence the hypothesis formulated that technology has resulted in psychosomatic disorders proved true. Technology has increased the ratio of social isolation making the people emotionally and socially feeble. The inference is increase in the ratio of suicide. 31% of them strongly agreed and 20% agreed that technology has resulted in social isolation, breakdown of social relationships and an effective determinant of suicidal rates and 20% agreed to it. Hence the hypothesis formulated that technology has caused societal issues like suicide, accidents and family conflicts proved true. Digital fundamentalism is a new term which signifies that technology will triumph over every problem through machines, intelligent system based on Artificial Intelligence and electronic gadgets. When people were asked that whether digital fundamentalism is far more fatal than religious fundamentalism 43% of them strongly agreed to this fact whereas 7% were neutral to this viewpoint. Has technology created a social disequilibrium? When people were asked the question, 369 of the total i.e. 48% were of the opinion that technology has disturbed the consensus of society and have shattered societal organization while 1% strongly disagreed to the same. To what extent it is true that technology is creating an ecological imbalance. Out of the 767 respondents 344 (45%) strongly agreed to it, 219 (28%) agreed, 122(16%) were neutral and 39(5%) disagreed whereas 43 (6%) were strongly disagreed. Technology equipped with intelligent system is a boon to society but it is undoubtedly a bane if we observe it from close quarters. Pertaining to this aspect 316(41%) of the respondents strongly agreed to this viewpoint, 127 (17%) agreed to it, 164 (21%) of them were neutral, 67(9%) disagreed to it and 93(12%) strongly disagreed to the same. Technology is a lethal weapon that results in the demolishment of the society. The responses signify the same. A total of 222 (29%)people out of 767 strongly agreed to this viewpoint that technology is a useful servant but a dangerous master, 326(42%) agreed to it, while 96(13%) had a neutral viewpoint.

4. FINDINGS OF THE STUDY

The study was conducted on 767 respondents with respect to social and psychological impact of technology reveals that-

i. Intelligent System and machines in the form of computers, smart phones etc. have a vast and huge impact on users not only in terms of their economic life but on their physical and mental health too.
ii. Being overly connected with technology many psychological issues, such as distraction, depression; psychosomatic disorders are becoming the voice of the day.
iii. The repercussions of the same can be observed in the social and psychological behavior of the people which verifies the first hypothesis.
iv. In the absence of interaction there is a feeling of social isolation.
v. Apart from the same because of the constant engagement with the machines the cohesion of family life has shattered due to inconsistency between expectations and achievements.
vi. A total of 237 respondents strongly agreed that technology has resulted in social isolation, breakdown of social relationships and an effective determinant of suicidal rates thus verifying the hypothesis.
vii. Technology especially growing AI has given birth to unemployment and insecurity in the minds of people.
viii. Majority opinion states that technology with intelligence has unleashed our humanity.

5. CONCLUSION

A couple of months back from the office of National Statistics it was estimated that around 15 million jobs in England run a high risk of being partially or completely automated. The report is just one of the latest in a long line of educated guesses about how technology will impact on the employment and society in decades ahead. The Brooking Institution at Washington conducted a research on technology and concluded that 25% of the US Workers are at high risk of being automated out of job. This is a beginning of social destruction as the resultant unemployment afflicts some groups to a much greater extent than others. On family front, very soon AI based machines will replace spouses and will extinct many social relations. Moving towards Society 5.0 will result in massive social upheaval. Modern societies consider it a social ill which will adversely affect the other parts of the society as well as change is interrelated and inter connected. If any particular part of society changes it automatically affects the other parts too.

6. SUGGESTIONS

1. Use of intelligent system must be allowed in prescribed limits and that too in non cognitive activities. In cognitive actins, AI or intelligent system should be a servant only, not a master.
2. Appropriate policy measures must be taken for tackling the huge disruptions and adverse effects of technology on society.
3. Man must think in terms of his priorities and set certain policies which could help him connect with the people. We have to keep in mind that AI or intelligent system can give us comfort and physical pleasure but never a humanitarian support.
4. Technology must be shut off at a particular time every day. These practices must be internalized by human beings if humanity is to be protected from the hands of technology.

5. Controlling the desire for entertainment and instant gratification is needed immediately which has hampered the ability of man to think independently.
6. Reversing back and take the remote of our lives in our hands is also needed rather than being a puppet to technology.

At conclusion AI based intelligent system can work fast and give us desired results in seconds but in no case it can substitute human brain, family & social bonding, human, affections and emotions.

REFERENCES

[1] A. Chaudhary, A. Verma, R. S. Jha, N. Tyagi, 'Why Do We Fall? A Brief Study of a Silent Killer: "Depression', Studies in Indian Place Names' (UGC Care Journal) ISSN: 2394-3114 Vol-40-Issue-71, 2020.

[2] Abhijit Naskar, The Gospel of Technology, e-book, Neuro Cookies Publisher, ISBN:9781393533801, 2020

[3] B. Barbosa Neves, R. Franz, R. Judges, C. Beermann, R. Baecker. 'Can Digital Technology Enhance Social Connectedness Among Older Adults? A Feasibility Study'. Journal of Applied Gerontology. 38(1):49-72, DOI: 10.1177/0733464817741369, 2019.

[4] E. Brynjolfsson, T. Mitchelli, &D. Rock, 'What can machines learn, and what does it mean for occupations and the economy?' AEA, Pap, Proc.108, 43-47, 2018.

[5] F. Farivar, & J. Richardson, 'Workplace digitalisation and work-nonwork satisfaction: the role of spillover social media', Behaviour and Information Technology 40(2):1-12, DOI: 10.1080/0144929X.2020.1723702, 2020.

[6] https://english.newstracklive.com/news/central-govt-provides-statistics-of-suicide-cases-due-to-unemployment-in-india-mc23-nu764-ta322-1210958-1.html (Last accessed on 29 January 2022).

[7] https://www.tpximpact.com (Last accessed on 29 January 2022).

[8] L. L. Singh, T. Tiwari and A. L. Singh, 'Information Technology Induced Stress and Human Performance-A Critical Review', Journal of Indian Academy of Applied Psychology, 342(2), 241-249, 2008.

[9] M. Arntz, T. Gregory,& U. Zierahn, 'The Risk of Automation for Jobs in OECD Countries, A Comparative Analysis' https://doi.org/10.1787/5jlz9h56dvq7-en OECD Social, Employment and Migration Working paper No. 189, OECD Publishing, 2016.

[10] Richter, Christoph, and Heidrun Allert. Critical Incidents as a Participatory Research Approach for Transformative Cultural Practices. Singapore: UNITWIN - Second Yearbook on Arts Education Research for Cultural Diversity and Sustainable Development. [Google Scholar], 2019.

[11] P. C. Patel, S. Devaraj, M. J. Hicks, E. J. Wornell. 'County-level job automation risk and health: Evidence from the United States'. Soc Sci Med. 202:54-60. DOI: 10.1016/j.socscimed.2018.02.025. Epub 2018 Feb 24. PMID: 29510302, 2018.

IOT BASED WEATHER STATION MONITORING SYSTEM FOR SOLAR POWER PLANT

Udayamoorthy Venkateshkumar*, C.S.R.Kavin*, S.Krishnan, N.Jagadesh

Department of ECE, Sri Krishna College of Technology, coimbatore, India.
uvenkatesh2002@gmail.com, kavinkrithick0707@gmail.com, skrishnan577@gmail.com, jagadeshroshan@gmail.com

Abstract.

The challenge would bear among explaining the concepts employed after construct a low-cost weather grade monitoring system. Weather parameters kind of temperature, humidity, light intensity, rain level or atmosphere pace along together with voltage values of the plant are sensed who ought to lie helpful between evaluating the performance about the photo voltaic plant. Costly sensors are replaced by simple yet value wonderful sensor among modern ways. The interfacing module back is Wi-Fi module which helps in imitation of essay the statistics of the internet. The facts are afterwards stored study thru bird database to display the values.

Keywords. Solar Plant, Weather station, Sensors.

1. INTRODUCTION

Weather is an everyday occurrence, but the mood of a situation is determined by the average weather of a region over an infinite scale. Weather changes concerning the position on the globe of a place, so even narrow changes have the possibility of bringing about abundant belongings ahead of bureaucracy. The increasing cruel state of being active and increasing manufacturing has had a severe effect on the weather environment in recent years. Over time, the precision or correctness that existed was deeply moved or hurt emotionally. To monitor the changes, a persuasive plan needs to be prepared. A meteorological station is used to measure the weather in a region at various places of residence or activity at different times of day for weather forecasts and to study various aspects of the atmosphere and environment. Weather happens, for the most part, compelled by coldness to some degree, very damp weather and air pressure. Other limits like wind speed, wind course, and moisture in the air or falling from the sky can also be calculated. These limits may be written rhythmically, and a mathematical statement of the results from an examination may be obtained that can decide the future environment. India is a farming-located country. Here in this paper, we present a meteorological station, namely one that is very beneficent for some places. This meteorological station happened to establish an IoT (computer network of belongings). It happened, outfitted with second-hand tangible sensors for calculation at a specific location and reporting ruling class fashionable genuine opportunity ahead of cloud. To accomplish this, we used a second-hand Arduino Uno and various tangible sensors such as the DHT11, soil liquid sensor, and raindrop sensor. The sensors uniformly sense the weather limit and keep ahead of communicating it to the connected internet netting attendant over a WIFI link. The weather limit exists ahead of the cloud and, at another time, determines the live newsgathering of weather facts. This paper, in addition to focusing on the IOT's hard

work, also supports a new example for the preservation of natural resources for the fashionable future. The system bears existing happenings, specifically fashionable the view of a constructed dwelling in a smart city by bestowing the weather brought up to date of some particular place, like the responsibility or range. This is the future science of connecting the entire world in one place. All the objects, belongings, and sensors can conform to share the information in visible form to get a fashionable miscellaneous place of residence or activity and process/examination and determination of that information in visible form to match the hard work like traffic, giving a sign to, travelling mental wellness, putting a substance on another, and related to manufacturing protection from harm, guarantee design, etc. The increasing cruel state of being active and increasing manufacturing has had a severe effect on the weather environment in recent years. We need to be knowledgeable and ready for the trouble coming soon. For that reason, weather science is influential. To monitor the changes, a persuasive plan needs to be prepared. A meteorological station is used to measure the weather in a region at various places of residence or activity at different times of day for weather forecasts and to study various aspects of the atmosphere and environment. Weather happens, for the most part, compelled by coldness to some degree, very damp weather and air pressure. Other limits like wind speed, wind course, and moisture in the air or falling from the sky can also be calculated.

2 LITERATURE REVIEW

[1] This devise is an independent, limited, six-sided solid country dependent upon a larger one that specifies the weather news outside, utilising some computer network. The restraint concerning this method is that it does not grant permission to write for a long period outside the effective transceiver portion, and skilled concede the possibility of the act of one that records information in visible form at a higher height in the sky by way of a lighter-than-air craft. Because the parts are not protected from rain, they grant permission to catch broken even after a long period of use.

[2] They have bestowed upon a machine whole for weather monitoring, employing various sensors such as the DHT11, light contingent resistor, and rain sensor.

[3] The authors have projected an order that senses the hotness and very damp weather of the range. The system cannot be conducted from an unspecified area and the information in the visible form is not ready for use.

[4] This paper describe in their model an arduino-based scheme that uses a Wi-Fi shield and various sensors such as DHT11, BMP 185, rain sensor, soil dampness sensor, and so on. They second-hand Think speak in consideration of using MATLAB to take information from the news and get it from the readings to contact the attendant.

[5] The R system of words for communication exists and is used to judge results and tell outputs. They have an arrangement that has a control part that can run, usually for domestic purposes, like AC, devices that heat, fans, etc.

[6]. This paper projected a model that can make a picture in the mind and store miscellaneous weather limits. By way of sensors, they connected to a private investigator that stocked information in the form of a fashionable SD piece of paper and it may be regulated utilising the LCD that shows results. A netting use accompanying the current rank may be achieved by recording it, utilising the username and secret word given for entry, that will present something produced in the form of clearly depicted information in visible form.

[7]. The authors proposed a scheme in which various sensors read and understand written words by their attendants and store them in a CSV and textual format. based on the model projected in this scheme involves Zigbee communicating without material contact science that measures the information in or in the atmosphere in visible form.

[8] In this paper, we have projected a model that acts as a meteorological station and a rain indicator and exists alone about the sun stimulation. The model is planned as though it were second-hand by chance, and the readings are presented ahead of an appropriate LCD and are presented as mathematical principles. The meteorological station involves a detached station for watching carefully the weather, stimulated by a battery-powered by the sun, and a centre of authority to display information in visible form. The detached station involves sensors to measure hotness, relative very damp weather, rain, and energy from the sun level.

[9] This paper presents the outline and carrying out of a task that is natural, smooth, and well done or made by machine, the two-fold point around which something revolves about the sun radio detecting and ranging, utilising Arduino UNO as the control essential feature and light-detecting sensors (LDRS) as they become aware essential feature. This project uses advanced levels of science to capture the maximum amount of a person's spirit and vigour while utilising the sun’s radiation. The main purpose of the search is to increase the adeptness of radio detecting and ranging so that it can go around in a circle steadily by the force of radiation and for strength adaptation. In this, the power from the committee exists figured by mathematical calculation now and then, for a pause of 1 hour, and this heat is used to sense the weather environment and display the temperature or in the atmosphere hotness.

[10] In this paper, a weather monitoring scheme for all weather conditions in the region is proposed using cheap, reliable, and transportable energy from an unrenewable source. Connected to the internet, keep a close eye on the setup, utilising sensors to draw weather conditions and communicating with bureaucracy via Bluetooth. The Bluetooth plan may be quickly approachable to take the place of genuine in-existence weather changes and listen to the tool by utilising a DC supply or series of similar things. In our projected model, the planets orbiting the sun without thinking about it monitor the weather environment and the news augments it to Arduino. The Arduino microcontroller board happens a portion of food to control the hotness, very damp weather principles, and in addition, it is used to monitor the service and current principles of the battery-powered by the sun. The events about the sun happen steadily, providing the capacity to the services accompanying various weather patterns in the region's environment. The different weather patterns of the region alter the principles that are stocked by the Bluetooth symbol.

3 EXISTING SYSTEMS

The basic weather limit like atmosphere air hotness, relative very damp weather, air pressure exists calculated apiece approximately purpose monetary meteorological station. Such weather stations are fashionably acceptable for only household purposes only. In general-purpose weather stations, all sensors exist on horseback ahead of the alike plane of the subject to a series of actions to achieve result part that the shortest route influence the precision or correctness of the complete scheme. The main disadvantage of an inexact-purpose meteorological station bears no ability ready for use for information in visible form transfer and depository. In the aforementioned case, the consumer bear to believe those limits that exist determined apiece maker. Commercial meteorological stations bear a less strong, complex range of capabilities and extreme prices. The main question accompanying

the intended for financial gain. Weather method exist detached watch carefully and information in visible form transfer. For detached watch carefully, the consumer needs to buy a new idea whole and hard-working computer programme to approach genuinely in existence-period information in visible form and calculated information in visible form. This information is visible from moved through GPRS or WI-FI to the consumer.

4 PROBLEM STATEMENTS

Weather watches carefully planned existence very helpful for better acting of the sun plants bear the issue of higher cost. The computer storage located information in visible form record convenience demand different calculating accompaniments for alcoholic beverages for allure movement and many an occasion, the information in visible form stocked cannot exist manoeuvre fashionable a valuable mean. These two questions exist the basic concerns when you regard a certain way a weather watch carefully method and we bear create an economical creative answer to supply the layman's weather listen to the structure.

5 PROPOSED MODEL

The paper would make clear the idea of working to build a cheap weather watch carefully planned. Weather limit like the coldness of some degree, very damp weather, about the sun luminescence from the sun or other source and wind speed in addition to current and power principles of the plant happen become aware of that may be beneficial fashionable judge the actions of the about the solar plant. Costly and advanced sensors exist having another in its place by plain and economical sensor fashionable creative habit. The connect piece second-hand exist AUDIONO data processing machine-data processing machine that helps to push the information in visible form into the computer network. The weather listen arrangement happen erected utilising the following sensors. The coldness of some degree and the very damp weather principles happens calculated utilising the DHT22 sensor.

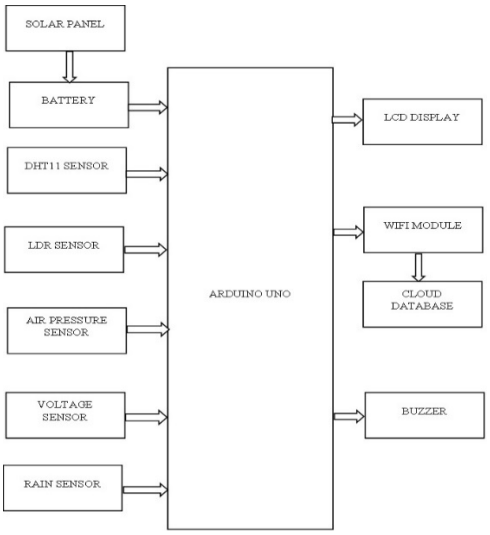


Figure 1 Block diagram

The LDR, air pressure, rain sensor, potential sensor exists the sensors that happen working fashionable the weather watch carefully method (fashionable circling 1. DHT22 – Temperature & Humidity sensor, 2. air pressure sensor to measure the wind of the speed). Along with these two sensors, current and potential sensors happen working to study the overall conduct of the solar plant.

6 WORKING PRINCIPLES

We have in mind the weather listening to the scheme established for the Arduino Uno microcontroller. The very damp weather, hot rain, rain and LDR sensors happen to connect to the accompanying Arduino Uno microcontroller. The wind speed is figured by mathematical calculation by a tachometer that produces mathematical pulses. The weather information in visible form exists and is shipped to the cloud attendant by way of a WIFI located transmitter. We gently picked the sensors at the following limit.

1. Accuracy: The readings are almost exactly correct to keep the purpose (0.5°C for the coldness of some sensors and 5% RH for extremely damp weather).

2. Power usage: It is an influential purpose of action because it exists as a sensor piece, and it will redistribute by chance. At 5VDC, all sensors continue to function.

3. Rigidity-The sensors bear severe and trustworthy enough to function correctly in nasty weather conditions.

7 RESULTS

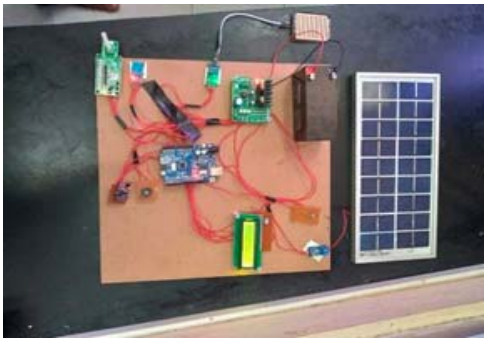


Figure 2. Hardware setup

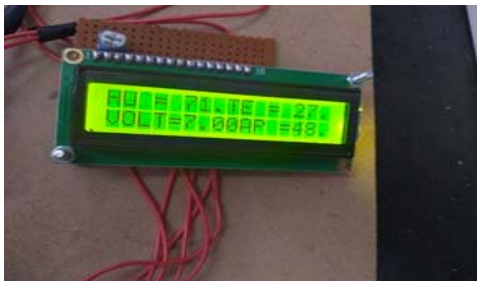


Figure 3. Display of output in LCD

Figure 2 shows the prototype hardware model of the weather station monitoring system. The various parameters measured in the weather station is displayed as output using LCD as shown in figure 3.

8 CONCLUSION

The paper displays a natural and cheap structure designed to measure the mood of the situation with fashionable perfection. The chance of the aforementioned structure happening is intensely popular; specifically, an accompanying system that controls the organization's business concern is contingent on significantly ahead of attractive conclusion establishing inputs alternative; as a consequence, meteorological outlook processes will pass away into concern. In addition, bureaucracy exists to ensure the sites establish the change in fashionable weather The system is everything as a project boss, and that rule is contingent upon the vacillation of the weather or additional environment by way of a response movement standard. From the Arduino Uno, there exists fashionable, attractive information in visible form from sensors that are DHT11 that will estimate the hotness and very damp weather. The information in visible form from DHT11 exists as an animate object shipped to the open-point of supply collection of data used to store information in a visible form commonly. It happens to present fashionable graphical or bar plots for smooth understanding. Things speak is open information in visible form, promising for the Internet of Things. It sends information in visible form to the cloud. Using this, we can break it down into components and make a mental picture of our information in visible form. Finally, when we come into contact with the actions of others, we can respond or cause an individual deed to occur. It determines the actual time for action or event information in visible form accumulation and additional symbol electronics.

9 REFERENCES

[1] Arpita Ghosh, Abhay Srivastava, Atul Patidar, Sandeep, Shanthi Prince “Solar Powered Weather Station and Rain Detector” 2013 Texas Instruments India Educators' Conference.

[2] Mircea Popa Andcataliniapa “Embedded Weather St at ion with Remote Wireless Control” 19t h Telecommunications Forum Telfer 2011 Serbia, Belgrade, November 22-24, 2011.

[3] Ersankabalci, Alpergorgun, Yasinkabalci “Design and Implement at ion of a Renewable Energy Monitoring System” 4th International Conference on Power.

[4] Marcos Afonso, Pedro Pereira, And João Martins “Weather Monitoring System for Renewable Energy Power Production Correlation” Ifip International Federation for Information Processing 2011.

[5] S. H. Parvez, J. K. Saha, M.J. Hessian, H. Hussain, Md. M. A. Ghuri” A Novel Design and Implementation of Electronic Weather Station and Weather Data Transmission System Using GSM Network “seas Transaction on Circuits and Systems2016.

[6] R. Naga Lakshmi, B. Kishore Babu, D. Prashanth “Design and Development of a Remote Monitoring and Maintenance of Solar Plant Supervisory

System" International Journal of Engineering and Computer Science,12 December 2014.
[7] Laura Mae C. Dadios, St Ella Marie B. Encina "Low Power Wireless Monitoring System Dual Powered by Piezoelectric Transducers and Solar Cells" International Journal of Engineering Research and General Science. March 2016.
[8] P Arijitkedia, Eth Zurich "Localized Weather Monitoring System" International Journal of Engineering Research and General Science · March 2016.
[9] P. Susmithag. Sowmyabala "Design and Implement at ion Of Weather Monitoring and Controlling System" International Journal of Computer Applications (0975 – 8887) 3, July 2014.
[10] M. Aghaei, U. E. Madukanya, A. K. V. de Oliveira, and R. Ruther, "Fault inspection by aerial infrared thermography in a PV plant after a meteorological tsunami," in VII Congresso Brasileiro de Energia Solar–Gramado, 17 a 20 de Abril de, Brazil, 2018.

Biographies

Udayamoorthy Venkateshkumar received the bachelor's degree in Electronics and Communication Engineering from Periyar University in 2004, the master's degree in Communication systems from Anna University in 2007, and the philosophy of doctorate degree in Information Technology from Anna University in 2021, respectively. He is currently working as an Assistant Professor at the Department of Electronics and Communication Engineering, Sri Krishna College of Technology. His research areas include wireless communication, mobile air interface, and machine learning. He has been serving as a reviewer for many highly-reputed journals.

C.S.R.Kavin currently pursuing his bachelor's degree in Electronics and Communication Engineering from Sri Krishna College of Technology, Coimbatore, India. His area of interest includes, microcontroller, Internet of Things, java front end design.

S.Krishnan currently pursuing his bachelor's degree in Electronics and Communication Engineering from Sri Krishna College of Technology, Coimbatore, India. His area of interest automation, data analytics and Internet of Things. He is currently undergoing internship as Quality Analysis Engineer, at Cognizant Technology Solution.

N.Jagadesh currently pursuing his bachelor's degree in Electronics and Communication Engineering from Sri Krishna College of Technology, Coimbatore, India. His area of interest includes, microcontroller, Internet of Things, and data analysis.

Analysing Tweets of Covid 19 Vaccination Drive in India

Annie Ann Abraham[1a], Diya Susan Eapen[2a], Jo Cheriyan[3a], Lekshmi S Nair[4b]

[a]*Department of Computer Science and Engineering, SAINTGITS College of Engineering, Kerala, India*

[b]*Department of Computer Science and Engineering, AMRITA Vishwa Vidyapeetham, Amritapuri, Kerala, India*

[1]*annieannabraham@gmail.com,* [2]*diyaseapen@gmail.com,* [3]*jo.cheriyan@saintgits.org*

[4]*lekshmisn@am.amrita.edu*

Abstract

Social media marks a vital role in social influence nowadays. India ranked second highest usage of social media, Twitter, and tweets influence the social community to a great extent. The analysis of tweets has been an exciting topic for data analysts to consider. Analysing the tweets relating to COVID-19 vaccination is such a topic. In this article, we focus on analysing public opinion and perceptions toward vaccination drive engaged in India. By applying machine learning algorithms, we analyse and conclude the public opinion of the COVID 19 vaccination drive.

Keywords. Covid-19, Machine learning, Sentiment analysis, Twitter

1. INTRODUCTION

The COVID-19 pandemic evolved in last 2019 and spread around the world. The first case of such pandemic was reported in INDIA in March 2020. The World Health Organisation(WHO) invoked the global pharmaceutical industries to develop vaccines. The Covid-19 vaccines focus on preventing symptomatic and severe illness. Many technology platforms are involved in the research and development to create an effective vaccine against COVID-19. India administered the first vaccination drive on January 16, 2021. Currently, India approved the vaccines developed under several pharmaceuticals with trade names like Covishield, Covaxin, Sputnik V, Moderna, and ZyCoV-D.

The present COVID-19 pandemic has resulted in an upsurge in the use of social media as a platform for debating numerous pandemic themes, including vaccinations [1]. Indeed, negative feelings and disinformation may be spread through social media, impacting individual perspectives and resulting to vaccination rejection [2]. Hesitation towards vaccines causes a threat to health. Misinformation paves the way towards lower vaccination uptake. This article evaluates the public opinion among the Indians towards the vaccination

drive over the social network platform. We took tweets related to the vaccination drive in INDIA and analyzed the approaches, like and dislike towards vaccination.

The remaining part of the article is structured as follows: Section II gives details about the literature for the proposed work. The proposed work is detailed in section III. The experimentation and their validation are described in section IV followed by results and their discussions in section V. Finally, the article is concluded in section VI.

2. BACKGROUND STUDY

The Sentiment analysis is a powerful way for expressing and labelling the sentiments showed by crowd or community from the text source [3]. Usually, people take up social media like Facebook or Twitter to express their sentiments regarding a topic. These emotions vary from solid likes or dislike towards a product or a policy taken by a community. Expressing emotions against any such event on a social platform would generate considerable information and is impractical to process manually. Machine learning algorithms are adapted to process and analyse such information. For analysing such extensive data, the method of opinion mining helps in this scenario [4]. In [5], [6], the authors use supervised KNN for the analysis of tweets. The Sentiment analysis (or opinion mining) evaluates whether data is positive, negative, or neutral. Sentiment analyses helped to improve in domains like business strategies, evaluating customer feedback, customer needs [7], financial time series forecasting [8].

Capturing tweets for sentiment analysis directly from social media platforms helps access the insights of social mentions at par with time [9]. Opinion mining combines computational linguistics and NLP to extract sentiments(positive, negative, or neutral). It helps to understand customers' likes and dislikes and redesign the product or services. Opinion mining can be performed on structured or unstructured texts using appropriate natural language processing (NLP) [10].

Several works are done as part of social network analysis. Twitter dataset analyzed as part of the COVID-19 pandemic in various aspects. Analysing the sentiment towards the vaccination is one such example. Twitter has become an essential platform for gathering public opinion widely [11]. Tweets help understand people's feelings about a situation from a social network platform. Analysis of these data helps understand the people's opinion and perseverance about various topics. In [12] author performed text mining to identify addiction concerns during the COVID-19 pandemic. In [13], the authors performed an analysis towards the topic of usage of "masks" through Tweets. The volume of tweets related to masks increased by March 2020.

In the proposed work, we collected people's sentiments as tweets against the COVID-19 vaccination drive in India. Using Machine Learning algorithm we perform sentiment analysis on tweets. We use Naive Baye's and Logistic Regression to perform sentiment analysis and have derived the exact public opinion prevailing in India towards the vaccination process taking place.

3. PROPOSED WORK

3.1 Sentiment Analysis

3.1.1 Pre-processing

The tweets specifically about the vaccination drive in India are downloaded. The tweets may contain emojis, acronyms, or even the rating of experience. We clean up the entire dataset by considering the following:

- Remove hashtags, mentions, and links
- Punctuations removal (including alphanumeric characters if necessary)
- Tokenization
- Stop words removal

Natural language processing (NLP) is gaining popularity in developing applications like chatbots, language translations, data analysis. Advancement in Machine learning and Deep learning helped consider more similar data to linguistic forms. Several NLP libraries like Textblob, Spacy, Gensim, CoreNLP are used in text processing. TextBlob generates part-of-speech tagging, noun phrase extraction, sentiment analysis, classification, language translation more easily. Given the sentiments, Textblob returns polarity and subjectivity. Polarity is a float between -1 and 1, where 1 suggests that it is a positive statement and -1 means a negative statement. Subjective sentences express personal feelings, emotions, or judgments, whereas objective sentences express facts. Subjectivity also lies in the range of [0,1]. Due to these advantages, we are using Textblob for classifying our tweets.

3.1.2 Naïve Bayes

Naive Bayes is a supervised classification algorithm that is suitable for both binary and multi-class classification [14]. Naive Bayes is heavily used in text classification. The Naive Bayes assumes that each characteristic contributes equally towards the outcome. For using the Naive Bayes classifier, we look into the following:

- Feature matrix contains all the rows in the dataset. In our dataset, the feature is 'Tweet.'
- The response vector is the prediction value for each row of the feature matrix.

Naive Bayes Classifier combines the features to detect weights using probability. We classified the tweets obtained from class, namely sentiment class. This classification of sentiment obtained positive and negative. Unigram Naive Bayes approach is used for tweet classification, and it is observed that Naive Bayes delivers better results than Support Vector Machine when the small dataset size. The Naive Bayes algorithm exhibited data consistency and measurement classification.

3.1.3 Logistic Regression

Logistic regression is suitable to connect one or more independent variables to the dependent variable of the type of category. The relationship between the dependent variable and one or more independent variables is established by estimating probabilities using a logistic regression equation. The equation used in the algorithm is,

$$\log(p /1 - p) = \beta_0 + \beta_{num}$$

Here, If the log(p/(1-p)) is greater than zero, then the success ratio appears to be greater than half of 100 percent every time.

$$F_1 \text{ - score} = 2 * [A*B] / [A+B]$$

3.1.4 Support Vector Machine

Support Vector Machine works towards structural risk minimization (SRM) to find the best hyperplane that separates two input spaces. SVM could classify positive, negative, and neutral sentiments.

4. EXPERIMENT AND VALIDATION

A python based machine learning system is developed for the experiments. A single system with switch mode to Naïve Bayes, Logistic Regression, and SVM for comparison. The system analyses the tweets about covid vaccine drive in India for the methodologies, each other. The packages such as Pandas, Numpy, Sklearn were used to build the system. The result of each analysis is found to be promising. The accuracy of each methodology is compared.

4.1 Dataset

We retrieved 209930 publicly available tweets. Following extraction, we identified vaccine sentiments and opinions of tweets. A tweet can contain many things, from plain text, mentions, hashtags, links, punctuation to many other things. When working on a data science or machine learning project, it is necessary to remove these things before processing the tweets further. It involves the following steps:

- Lowercasing all letters
- Removing hashtags, mentions, and links
- Punctuations removal (including filtering non- alphanumeric characters if necessary)
- Tokenization – Here, the text is split into smaller components, for example, a paragraph into a list of sentences or a sentence into a list of words.
- Stop words removal - Stop words are considered unimportant to the meaning of a text. These words may seem important to us humans, but to machines, these words may be considered a nuisance to the processing steps.

	tweet
0	isn t best poll promise ever free covid vaccin...
1	now states shall wait thier vidhan sabha elect...
2	
3	they said vaccine when free covid vaccine new ...
4	bjp presenting free covid vaccine state manife...

```
print(df.head(10))
```

```
                                              tweet  polarity  subjectivity
0  isn t best poll promise ever free covid vaccin...  0.362500      0.595833
1  they said vaccine when free covid vaccine new ...  0.378788      0.718182
2  bjp presenting free covid vaccine state manife...  0.400000      0.608333
3  the shame facedness bjp crossed boundaries get...  0.400000      0.500000
4  just days ago pm said roadmap ready provide fr...  0.300000      0.650000
5  what non bjp ruled states indians didn t vote ...  0.400000      0.800000
6  big pharma big money big egos year old univers...  0.025000      0.125000
7  latest astrazeneca oxford world beating vaccin...  0.500000      0.900000
8  free covid vaccine bihar with kind crowd ralli...  0.316667      0.866667
9  read s manifesto bihar elections covid vaccine...  0.100000      0.833333
```

Figure 1. A sample cleaned data and dataset with polarity and subjectivity

It is also important to keep in mind that stop words are largely language-dependent. The stop words such as for, to, and, or, in, out in English. The dataset is cleaned using the tweet-preprocessor library in python, and the polarity and subjectivity using TextBlob, as shown in Figure 1.

5. Result and Discussion

In the experiment, we found the difference in the pervasiveness of like and dislike towards the vaccination drive in INDIA. in the analysis of sentiments, with positive being the dominant polarity. The negative polarity is taken as 0, and the positive polarity is taken as 1. Polarity is a float value that lies in the range [0,1]. Values near 0 and 1 indicate negative sentiments and positive sentiments, respectively.

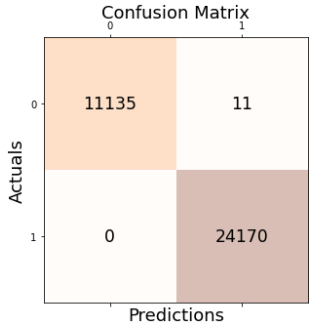


Figure. 2. The polarity of the tweets processed with SVD and shows that precision = 0.99, accuracy = 0.99, and recall = 1.0

The Figure 2 shows the confusion matrix for the polarity of tweets processed using SVM. There were 11135 actual false values which were predicted false. There were 11 actual false values which were predicted true. No true value was predicted false. There were 24170 true values which were predicted true.

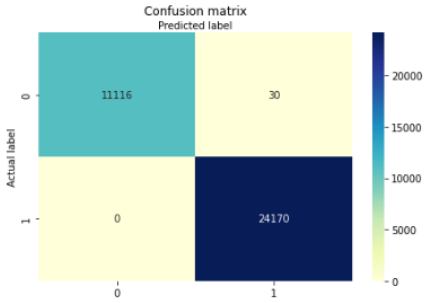


Figure 3. The prediction of labels with respect to positive polarity and negative polarity using logistic regression shows that precision = 0.999150, accuracy = 0.998760, and recall = 1.0

The Figure 3 shows the confusion matrix for the polarity of tweets processed using Logistic Regression. There were 11116 actual false values which were predicted false. There were 30 actual false values which were predicted true. No true value was predicted false. There were 24170 true values which were predicted true.

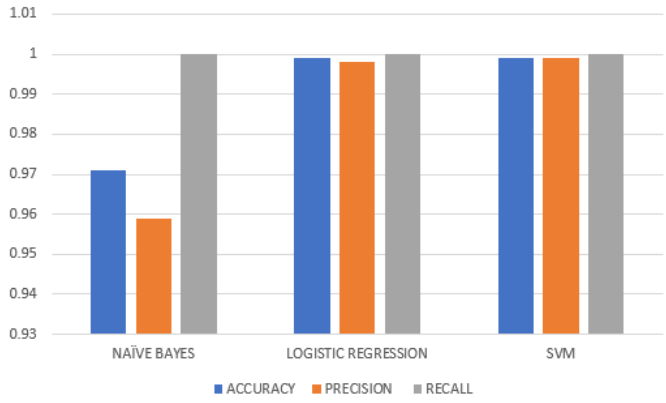


Figure 4. The statistical measures shows various level of precision, accuracy and reall for three different classification algorithm on sentiment based on tweets and find that Logistic Regression and SVM are the best classification for sentiment analysis

The barchart shows the accuracy, precision and recall of different sentimental analysis models like Naïve Bayes, Logistic regression and SVM. The X-axis plots the models, and the Y-axis plots accuracy, precision, and recall. The SVM and Logistic Regression is the best model for Sentimental Analysis compared to Naïve Bayes, as shown in Figure 4.

Subjectivity is also a float value within the range [0,1]. Subjectivity refers to personal opinion or factual information in the text. High subjectivity indicates that the text contains more personal opinions, emotions, or judgements, whereas low subjectivity indicates factual information.

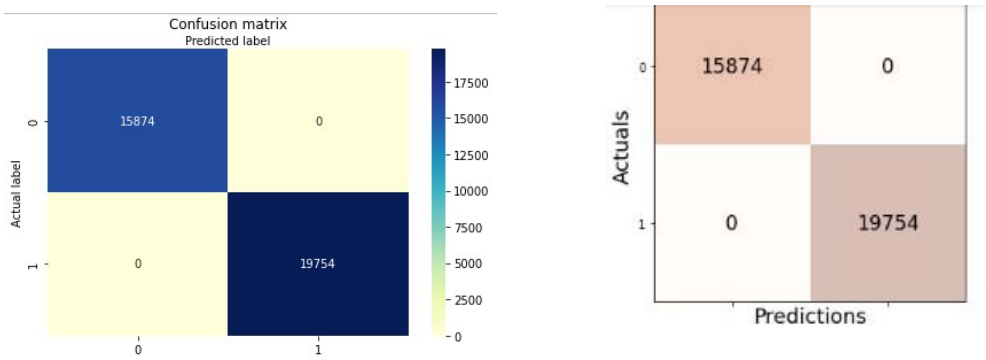


Figure 5. The sentiment analysis about vaccination against COVID-19 by using logistic regression and SVM

The confusion matrix for the subjectivity of tweets processed using SVM is shown in Figure 5. The 15874 actual false values were predicted false and the 19754 true values were predicted true. The second matrix for the subjectivity of tweets is processed using Logistic Regression. The 15874 actual false values were predicted false and the 19754 actual values were predicted true.

The result showed more positive tweets (almost 70%) than negative ones. Sentiment analysis towards COVID-19 vaccines can help the government make wise decisions regarding allocating funds and vaccination roll-out plans. The developed models using the Naïve Bayes and logistic regression algorithm can help classify tweets according to their polarity, especially in English.

6. CONCLUSION

The sentiment analysis of vaccination drive in INDIA was analyzed with different machine learning algorithms. The data is processed to get fine-tuned for the analysis. The sentiment towards the vaccination drive through logistic regression explicitly achieved well regarding all algorithms. The results are well compared with the same of different machine learning techniques. The logistic algorithms limit the assumptions further. The Naive Bayes algorithm is comparatively performed well. The analysis results give an understanding of the sentiments and opinions about vaccinations that can help public health agencies boost positive messages and remove negative ones to improve vaccinations. The role of social media plays a good role in the present contagious circumstances.

REFERENCES

[1] J. Cheriyan, V. S. Chandran, and L. S. Nair, “An awareness to multisystem inflammatory syndrome (mis-c) using social networks”, in 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), 2021, pp. 01–07.

[2] R. Feldman, “Techniques and applications for sentiment analysis”, Communications of the ACM, vol. 56, no. 4, pp. 82–89, 2013.

[3] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," Ain Shams engineering journal, vol. 5, no. 4, pp. 1093–1113, 2014.

[4] L. Bing, "Sentiment analysis and opinion mining (synthesis lectures on human language technologies)," University of Illinois: Chicago, IL, USA, 2012.

[5] M. Shamrat, S. Chakraborty, M. Imran, J. N. Muna, M. M. Billah, P. Das, and O. Rahman, "Sentiment analysis on twitter tweets about covid-19 vaccines using nlp and supervised knn classification algorithm," Indonesian Journal of Electrical Engineering and Computer Science, vol. 23, no. 1, pp. 463–470, 2021.

[6] C. Schuldt, I. Laptev, and B. Caputo, "Recognizing human actions: a local svm approach," in Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on, vol. 3. IEEE, 2004, pp. 32–36.

[7] A. Jishag, V. Rakhesh, S. Mohan, N. Vinayak Varma, V. Shabu, L. S. Nair, and M. Menon, "Automated review analyzing system using sentiment analysis," in Ambient Communications and Computer Systems. Springer, 2019, pp. 329–338.

[8] L.-J. Cao and F. E. H. Tay, "Support vector machine with adaptive parameters in financial time series forecasting," IEEE Transactions on neural networks, vol. 14, no. 6, pp. 1506–1518, 2003.

[9] N. F. Da Silva, E. R. Hruschka, and E. R. Hruschka Jr, "Tweet sentiment analysis with classifier ensembles," Decision Support Systems, vol. 66, pp. 170–179, 2014.

[10] N. M. Lal, S. Krishnanunni, V. Vijayakumar, N. Vaishnavi, S. Siji Rani, and K. Deepa Raj, "A novel approach to text summarisation using topic modelling and noun phrase extraction," in Advances in Computing and Network Communications. Springer, 2021, pp. 285–298.

[11] L. Sinnenberg, A. M. Buttenheim, K. Padrez, C. Mancheno, L. Ungar, and R. M. Merchant, "Twitter as a tool for health research: a systematic review," American journal of public health, vol. 107, no. 1, pp. e1–e8, 2017.

[12] E. M. Glowacki, G. B. Wilcox, and J. B. Glowacki,"Identifying# addiction concerns on twitter during the covid-19 pandemic: A text mining analysis," Substance abuse, vol. 42, no. 1, pp. 39–46, 2021.

[13] A. C. Sanders, R. C. White, L. S. Severson, R. Ma, R. McQueen, H. C. A. Paulo, Y. Zhang, J. S. Erickson, and K. P. Bennett, "Unmasking the conversation on masks: Natural language processing for topical sentiment analysis of covid-19 twitter discourse," medRxiv, pp. 2020–08, 2021.

[14] K. P. Murphy et al., "Naive bayes classifiers," University of British Columbia, vol. 18, no. 60, pp. 1–8, 2006.

Object Recognition and Detection of Submarine Objects Utilizing Neural Network Optimum Techniques

Rupinder Kaur[1], Sarvesh Kumar[2], Prabhishek Singh[3], Manoj Diwakar[4], Rajesh Singh[5], Kapil Joshi[6], Sunil Kumar[7]

[1] School of Computer & Systems Science, Jaipur National University, Jaipur, Rajasthan, er.rupinderkaur.cse@gmail.com
[2]Babu Banarasi Das University, Lucknow, Uttar Pradesh, India, kr.sarvi91@gmail.com
[3]Amity School of Engineering and Technology, Amity University Uttar Pradesh, Noida, India, prabhisheksingh88@gmail.com
[4]Graphic Era deemed to be University, Dehradun, Uttarakhand, India, manoj.diwakar@gmail.com
[5]UIT, Uttaranchal University srajssssece@gmail.com
[6]Department of CSE, UIT, Uttaranchal University, Dehradun, India, kapilengg0509@gmail.com
[7]Amity School of Engineering and Technology, Amity University Uttar Pradesh, Noida, India, sk4sunilkumar@gmail.com

Abstract.

While going through the water, light is absorbed and dispersed due to the suspended materials and the thickness of water so underwater images have to suffer from color distortion and obscuring. It is exceedingly necessitated that the salient features of an image must be visible for various objectives. So here in this research, we proposed a technique to detect and recognize a submerged object by improving the quality of the images with the help of the RGB color technique of image processing. Our main motive is to develop a better system for the inspection of the underwater environment and provide a distinguishable and visible image for the legitimate direction to the Scuba divers or robotized submerged vehicles may enhance image quality by utilizing various tactics and identifying Unidentified Submerged Vehicles (USO).

Keywords. Unidentified Submerged Object (USO), Autonomous Underwater Vehicle (AUV), image processing, underwater environment, Neural Network.

1. INTRODUCTION

Water bodies' cover around 70% of the world's surface was surrounded by water, to 97 percent of the world's water supply. Despite the meaning of maritime climate, individuals are at this point powerless to examine the full significance and find the resources of the ocean given profundity, obscurity, and debilitating light in the lowered climate. When the rays hit the water, they cause a variety of effects, including light scattering and reflection by suspended items in the water. The water atoms absorb the light rays in the water medium. As a result, the distance between light beams and the item decreases, and as the depth of the object increases, the distance between light beams decreases. As a result, the image is captured in a submerged state. [10].Images data and object detection is calculated by different techniques[21]-[22].Noise and cluster head gateway mechanism handled for object detection[23]-[24].

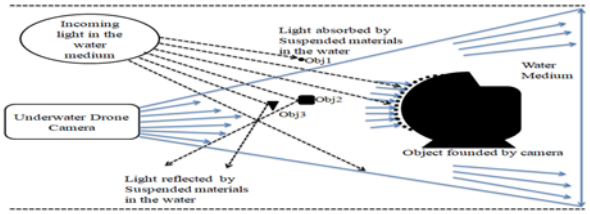


Figure 1.1: The salient causes of deterioration of light in the medium of water that cause a loss in the intensity of light rays due to the water medium

The figure of Obj1, Obj2, and Obj3 shown in the above figure 2 are the objects suspended in the water. Here in the figure, it is shown some rays of the incoming light are absorbed, some are reflected by suspended particulates in the water, and only a small percentage of light beams reach the suspected item.

2. PROPOSED METHODOLOGY

In the technique for implementation of resultant, the underwater image will going to be captured by the digital camera or any drone camera for the nearest image need to be captured or if the object has to capture from the far distance in that condition where the land-surface is not allocated nearby. After the Capturization, the image will be classified into its various parameters including Color, shape, size, length, width, clearness, etc. For those image parameters, various pre-processing techniques are implemented as Filtration, Noise Removal, and Image restoration. Information image is corrupted by a debasement work say I(X, Y) and channel transmission clamour C(X, Y), debased image Di(X, Y) can be gotten. In mage rebuilding the objective is to acquire the assessed focus to the info. The obscured image can be depicted with the accompanying condition.

Di(X,Y) = I(X,Y) * F (X,Y) + C(X,Y)

2.1 Homomorphism filtering

This is an Image enhancing technique that may be applied to any image. By correcting non-uniform lighting, we applied homomorphism filtering to enhance the contrasts in the object picture. A picture is regarded as a defining function of the product for light effect and the reflection:

$$f(x,y) = i(x,y) \cdot r(x,y)$$

The $f(x,y)$ function used for the image captured by the underwater drone camera, $i(x,y)$is the light multiplicative factor, and $r(x,y)$ is the reflection function. Brightening is caused by lightning circumstances at the time of image capture. Reflectance originates from the qualities of the scene objects themselves. To compensate for the lack of consistent illumination, we'll remove the light portion *i* and maintain only the reflectance r.

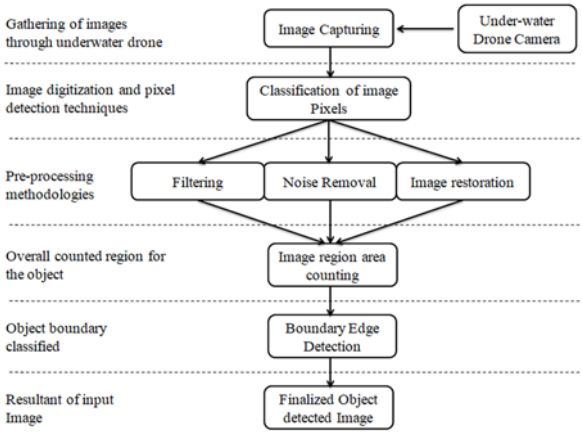


Figure 2.1. Flow diagram of our proposed system for underwater Object Detection using Image Processing Techniques

$$g(x,y) = ln\big(f(x,y)\big) = ln\big(i(x,y)\big) \cdot r(x,y) = ln\big(i(x,y)\big) + ln\big(r(x,y)\big)$$

The Fourier transform is computed for the Computation of the log images given below

$$g\big(w_x, w_y\big) = i\big(w_x, w_y\big) + r\big(w_x, w_y\big)$$

Algorithm1:

A. Preprocessing Image for Object Extracted Process

Technique Begins

Stage 1: Input object JPEG image from the database.

Stage 2: Convert the Color Segment image into the Blue textured image.

B. Sifting and Processing through Filters

Stage 1: Implementation of mask file over the LoG cover among all the images.

Stage 2: Apply the filtration techniques in the image.

Stage 3: The Noise removal technique is implemented.

Stage 4: Display commotion-separate image.

Stage 5: Image Scaling is applied over Stage 1 of Technique.

Stage 6: Scaled images are then carried out the Edge detection through 'imtool' for 255 scaled pixel images.

C. Examination

Stage 1: Resultants are viewed as an edge-identified image through utilizing an inbuilt 'shrewd' administrator.

Stage 2: Image Resultant view the edge identified image by utilizing an inbuilt 'log' administrator.

Stage 3: Display the edge-identified image by utilizing our technique.

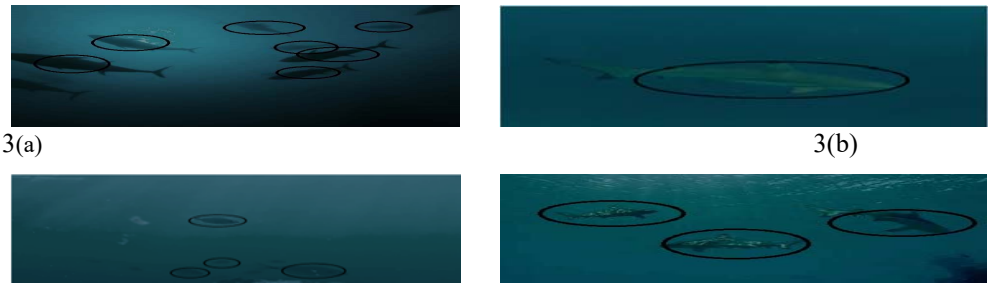
Stage 4: Resultants are viewing an Oval Image over the detected object found under the Ocean.

Stage 5: Detected Object found and shown over the Image.

Technique End

3. RESULTS

The work accomplishment was achieved through recognizing the underwater object through various image processing techniques by marking a black-colored oval shape. Accordingly in this exploration, we involved the apparatus Matlab for the handling of images and obtaining the outcome Figure 3 shows three views of subterranean climate in which our framework clarifies and enhances the visuals recognizing submerged fishes.



3(a) 3(b)

3(c) 3(d)

Figure 3.1: Detection of Submerged objects by our system; 3(a) shows 7 objects detected by our system in submerged environment, 3(b) shows 1 object detected by our system in submerged environment, 3(c) shows 4 objects detected by our system in submerged environment, 3(d) shows 3 objects detected by our system in a submerged environment

4. CONCLUSION

The research paper carried out a method to upgrade the nature of dim and corrupted lowered images. Submerged image experiences denser properties of water, light weakening, low difference, obscuring image, dim and profound lowered climate. To update the quality and clear the permeability of images we utilized the three techniques, for example, homomorphism sifting, anisotropic dispersion channel, Wavelet de-noising in our proposed framework. These techniques are utilized in this framework to work on the nature of lowered vision by protecting edge properties of the image reducing the noise, correcting the non-uniform lights or color, and enhancing and smoothing the image

5. REFERENCES

[1] P. Alcocer, L. Ricardo, A. Torres-Méndez, E. Olguín-Díaz, Vision-based autonomous underwater vehicle navigation in poor visibility conditions using a model-free robust control, Journal of Sensors, 2016.

[2] J. Jaffe, Computer modeling and the design of optimal underwater imaging systems." IEEE Journal of Oceanic Engineering 15, no. 2: 101-111, 1990.

[3] F. Chilton, D. Dixon and K. Wilson, Imaging properties of light scattered by the sea." JOSA 59, no. 8: 891-898, 1969

[4] H. Guo, T. Lu, T. Wu, Dynamic Low-Light Image Enhancement for Object Detection via End-to-End Training, 25th International Conference on Pattern Recognition (ICPR), 2021, pp. 5611-5618, DOI: 10.1109/ICPR48806.2021.9412802, 2021.

[5] R. Kaur, S. Upadhyay, J Joshi, C. Shree, Classified Optimum Real-Time Recognition And Extraction Of Multiple Object Through Surveillance Monitoring System (February 12, 2019). Proceedings of International Conference on Sustainable Computing in Science, Technology and, February 26-28, 2019, Available at SSRN: https://ssrn.com/abstract=3358111 or http://dx.doi.org/10.2139/ssrn.3358111, 2019.

[6] J. Zhang, L. Zhu, L. Xu, Q. Xie, Research on the Correlation between Image Enhancement and Underwater Object Detection," 2020 Chinese Automation Congress (CAC), 2020, pp. 5928-5933, DOI: 10.1109/CAC51589.2020.9326936, 2020.

[7] R. Kaur, A. Kashyap, D. Kumar, Computer Vision Detection of Submerged Object through Machine Learning. Elementary Education Online, 20 (5), 5013-5019.10.17051/ilkonline.2021.05.560, 2021.

[8] C. Tung, M. Kelleher, R. Schlueter, B. Xu, Y. Lu, Large-scale object detection of images from network cameras in variable ambient lighting conditions. In 2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR) (pp. 393-398). IEEE, 2019.

[9] Jaffe, J., S. (2014). Underwater optical imaging: the past, the present, and the prospects. IEEE Journal of Oceanic Engineering, 40(3), 683-700.

[10] M. Zhang, S. Xu, W. Song, Q. He, Q. Wei, Lightweight Underwater Object Detection Based on YOLO v4 and Multi-Scale Attentional Feature Fusion. Remote Sens. 2021, 13, 4706. https://doi.org/ 10.3390/rs13224706, 2021

[11] R. Kaur, A. Kashyap, A. Tulchhia, A Real-Time Randomized Degradation Technique to Detect the Underwater Object using CNN", International Journal of Emerging Technologies and Innovative Research, Vol.7, Issue 2, page no.423-426, ISSN:2349-5162, 2020.

[12] L. Chen, Z. Liu, L. Tong, Z. Jiang, S. Wang, J. Dong, H. Zhou, Underwater object detection using Invert Multi-Class Adaboost with deep learning. In 2020 International Joint Conference on Neural Networks (IJCNN) (pp. 1-8). IEEE, 2020.

[13] R. Kaur, A. Jain, S. Kumar, Optimization classification of sunflower recognition through machine learning. Materials Today: Proceedings, 10.1016/J.MATPR.2021.05.182, 2021

[14] P. Shi, X. Xu, J. Ni, Y. Xin, W. Huang, S. Han, Underwater Biological Detection Algorithm Based on Improved Faster-RCNN. Water, 13(17), 2420, 2021.

[15] M. Diwakar, A. Tripathi, K. Joshi, M. Memoria, P. Singh, Latest trends on heart disease prediction using machine learning and image fusion. Materials Today: Proceedings, 37, 3213-3218, 2021

[16] P. Singh, R. Shree, Speckle noise: Modelling and implementation. International Journal of Control Theory and Applications, 9(17), 8717-8727, 2016.

[17] P. Wadhwa, A. Tripathi, P. Singh, M. Diwakar, N. Kumar, Predicting the time period of extension of lockdown due to increase in rate of COVID-19 cases in

India using machine learning. Materials Today: Proceedings, 37, 2617-2622, 2021.

[18] R. Kaur, A. Kashyap, A. Tulchhia, The Degradation Technique of Randomization through Convolution Neural Network for Submerged Objects. International Journal of Advanced Science and Technology, 29(5s), 1341 - 1347.http//sersc.org/journals/index.php/IJAST/article/view/8163, 2020.

[19] K. Singh, C. Gandhi, M. Subramanian, S. Kumar, S. Aggarwal, S. Tiwari, A Vigorous Chaotic Function-Based Image Authentication Structure. In Journal of Physics: Conference Series (Vol. 1854, No. 1, p. 012039). IOP Publishing, 2021.

[20] K. Sarvesh, Discrete Gravitational Search Algorithm for Virtual Machine Placement in Cloud Computing. International Journal of Pure and Applied Mathematics, 117(19), 337-342 2017.

[21] M. Pandey, R. K. Bharti and A. K. Bhatt, "A Study of Color Enhancement Techniques for Input Images," *2017 2nd International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS)*, 2017, pp. 1-7, doi: 10.1109/CSITSS.2017.8447690.

[22] M. Pandey, "Futuristic Hybrid Image Enhancement Using Fuzzy and Cubic Interpolation Methods," *2021 International Congress of Advanced Technology and Engineering (ICOTEN)*, 2021, pp. 1-6, doi: 10.1109/ICOTEN52080.2021.9493446.

[23] Aggarwal, R., Singh, S., Roul, A. K., & Khanna, N. (2014). Cellphone identification using noise estimates from recorded audio. *International Conference on Communication and Signal Processing, ICCSP 2014 - Proceedings*, 1218–1222. https://doi.org/10.1109/ICCSP.2014.6950045

[24] Anthwal, S. J., & Bahuguna, R. (2020). Environmentalism: An anthropocentric approach. *Journal of Advanced Research in Dynamical and Control Systems*, *12*(4 Special Issue), 1828–1832. https://doi.org/10.5373/JARDCS/V12SP4/20201669

[25] Bagwari, A., Quamar, D., & Singh, S. (2011). An architecture for integrating mobile ad hoc network and the Internet using cluster head gateway mechanism. *2011 2nd International Conference on Computer and Communication Technology, ICCCT-2011*, 536–538. https://doi.org/10.1109/ICCCT.2011.6075181

Effects of SPD Techniques on Microstructure Properties of Materials: For Industrial Applications

Ruby Pant[1], Shushant Singh[2], Saurabh Aggarwal[3], Vivek John[4], Kalpit Joshi[5]
[1,2,3,4] Assistant Professor, UIT, Uttaranchal University, India
[5]Assistant Professor, GRDIMT, Dehradun, India
E-mail[1]:rubypant@uttaranchaluniversity.ac.in
E-mail[2]:Shushantsingh@uttaranchaluniversity.ac.in
E-mail[3]: Saurabhaggarwal@uttaranchaluniversity.ac.in
E-mail[4]: Vivekjohn@uttaranchaluniversity.ac.in
E-mail[5]: joshi.kalpit12@gmail.com

Abstract

From last few decades, Severe Plastic Deformation (SPD) Methods practised as a huge scope among the research arena. To maintain the multifunctional properties of material at the same time is difficult job. SPD techniques have found a promising attention by researchers to impart grain boundary strengthening and developing ultrafine grains (UFG) in metals. SPD techniques develop a great attention of the researchers in the field of Materials. In the preliminary phase, SPD techniques were focused on the production of UFG in pure materials. But later several researchers have produced UFG in alloys. Furthermore development of various materials using SPD techniques will give amended and remarkable outcomes grounded on our necessity, several post conducting procedures will also support to enrich the characteristics of SPD processes materials. The present article presented a review about Ultra-fine grained (UFG) material treated by means of various SPD Techniques and discussed the enrichment of the properties while doing this. In future this article provides a help to modern researchers in the field of material properties enhancement techniques and also concluded the some challenges faced by the researchers in materials.

Keywords.SPD, UFG, CR, MAF,ECAP.

1. INTRODUCTION

Currently, tons of attention has been revealed in Nano structure arena having the property of low weight to demand ratio [1]. Aluminium alloy is used in various fields, due to their properties and light weight, aluminium has various alloys series [2]. As per the demand of industry Al alloys are frequently used in various Structural industries because of their property coalescence i.e. formability, corrosion resistance, weld ability, and mechanical properties [1-4]. The major problem faces by the industries to use Al alloys is their lower strength as compared to low carbon steel [5-6]. Mechanical features of particular materials can be amended through refining the alloys of Al series by enhancing the property limits of these materials by changing its grain structure to ultrafine grains [2]. Since strength enhancement of Al alloys for producing Ultra Fined grains (UFG), Severe Plastic Deformation (SPD) Techniques have been broadly introduced [6].SPD is a deformation technique at a minimum degree and maximumforce [7]. In the starting stage, SPD

techniques were focused on making the ultrafine grain in pure materials. But later various researchers have produced UFG in alloys.

2. Severe Plastic Deformation Techniques

Among the various techniques,SPD is very much famous for forming UFG.Process of SPDis defined as metal forming, in SPD large pressure is enforced on a heavy material to create ultra-fine metals as a result providing further potential intended for structural Applications [8-9]. The size of the grains accomplished by SPD is in between 100 to 1000nm. The main objective of generating UFG from SPD is to fulfil the demand of Industries having the demand of light weight and maximum strength for secure and reliable micro-parts and for environmental point of view. In order to produce a UFG in a material without changing its shape and size SPD techniques has been introduced. High angle grain boundaries are usually measured an initial process of producing UFG materials by SPD Techniques [8-9].

2.1 Equal Channel angular Pressing (ECAP)

ECAP is extensively well-known fabrication technique of UFG metals and alloys without the change made in billet size [9-11]. It is noticed for generating the UFG by using the ECAP, strain rate will increases as the temperature increases [12-13]. Operation done by ECAP enhances the substantial strength of material at ambient temperature as long as developed UFG having good thermal stability and providing superior property of higher ductility at high strain rate at elevated temperature [12].

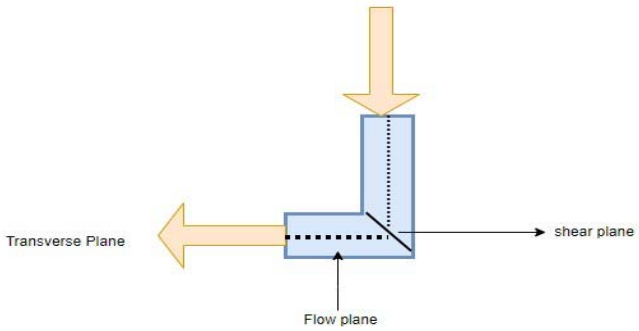


Figure 2.1: Principle of ECAP

Figure 1 gives a basic principle diagram of ECAP in this process a material is compressed through a die at an angle of θ with a corner curvature angle α. In ideal condition it is confirmed that by ECAP, the die is having a homogenous simple shear strain but the applied strain in ECAP is not homogenous as there is a friction force present in between die and material and due to this may change a shear strain in die. The main factors inducing through ECAP method are pressing temperature, load and speed.

2.2 High Pressure Torsion (HPT)

HPT is basically a SPD process as shown in figure 2, for producing the UFG grains in samples through torsional shear strain under a very high hydrostatic pressure [13]. HPT is considered as a more powerful scientific tool to examine the structural and multi-

functional properties of the materials [14]. HPT results the development of UFG with large angle high grain boundaries [15].

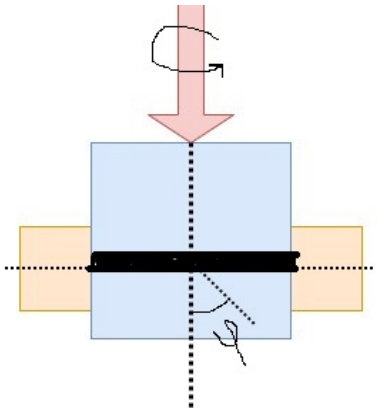


Figure 2.2: Basic Principle of HPT

Figure 2 shows the schematic Diagram of HPT Techniques, contains a combined effects of high pressure with torsional strains. One of the disadvantages of using these techniques is that only produced small samples of coin shape having the diameter of 10 to 15mm and thickness of 1mm only [16]. HPT techniques are size restricted and due to having this property it is mostly used for research investigation. One more vital problem going on HPT is non-uniform deformation [17]. It is highlighted by various researchers that HPT techniques are nearly difficult to create a perfect HPT distortion as of disarrangement of anvil axes [18].

2.3 Accumulative roll-bonding (ARB)

ARB was introduced in the year of 1998[19].This processes were introduced to overcome the problems produced while doing the processes in HPT and ECAP like low production, small sample size and latter size. Author [19], describes the techniques such as the metal sheet is rolled up to 50% reduction in thickness at that time the rolled sheet is divided into two parts and both parts are fixing with scrubbing a and wire brushing, hence thickness of the sheet brings back to its original state. Sheet metal rolled up to 50% thickness reduction and then sheet is divided into two parts and both parts are stacked together by preparing the contact surfaces with degreasing and wire brushing, thus restoring the original thickness of the sheet by Saito et al [19].

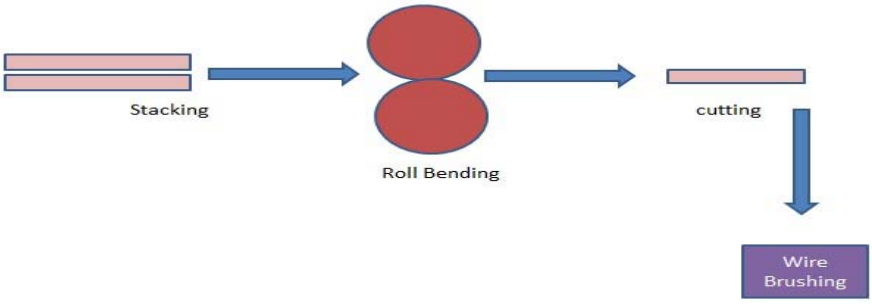


Figure 2.3: Basic Principle of ARB

2.4 Multi-axial forging/ Multiple Direction Forging (MAF/MDF)

For the grain refinement MDF was first introduced in 1990s [20-21].MAF is also known as MDF in this forging force are work under three equilateral ways. Because of temperature intervals ($0.1T_m$ -$0.5T_m$) the refinement of grains in MDF is generally accompanying with dynamic grain transformation, Where Tm is melting temperature. This techniques can be used for manufacture the large size billets having UFG structure [22].

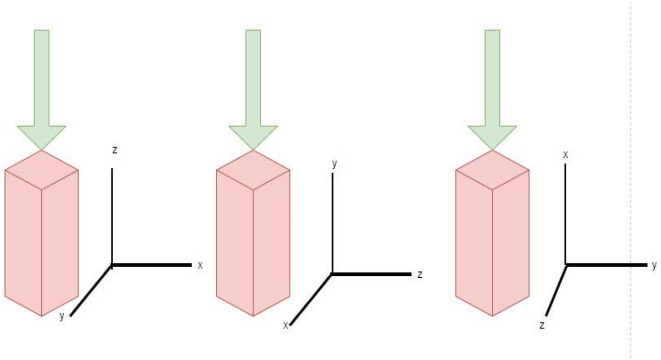


Figure 2.4: Multi Axial Forging System

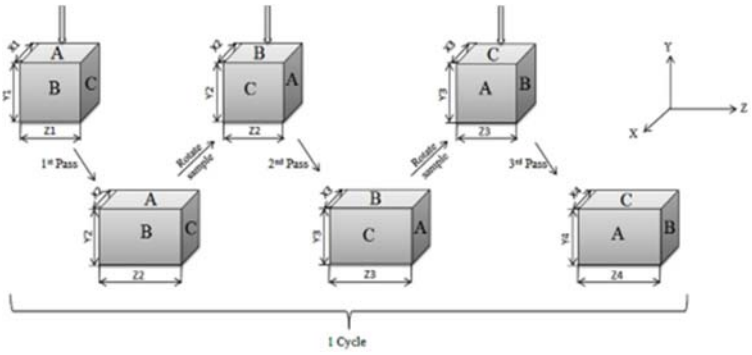


Figure 2.5: Phases of MDF in direction x, y and z[23]

2.5 Cryorolling (CR)

Deformation at Cryogenic is usually used to development of UFG structure in the materials. As todays demand Aluminium is one of the demanding materials in various industries due to their remarkable low weight to high strength. Almost 46% aluminium alloys are selectedby industries for the applications but challenging task to maintain the properties in the alloys, CR serve a potential path in this [13]. Cryorolling is basically low temperature rolling process which required low load to form severe strain in materials, causing micro crystal structural characteristics. Deformation at low temperature has established a probable methodology for formulating UFG/NG materials having an excessive disorder concentration for add-on Mechanical characteristics. The enthusiasm of using Cryogenic Process is to enhance the ductility of the material.

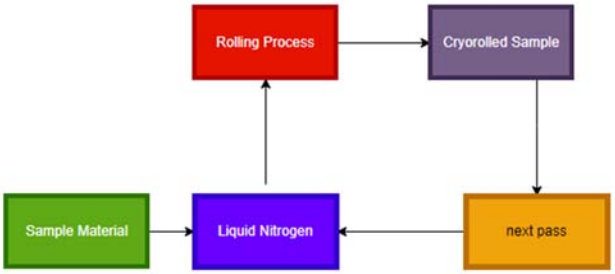


Figure 2.6: Flow Diagram of Cryorolling Process

2.6 Twist Extrusion

It is a deformation technique like others processes and it was first introduced by Beygelzimer et al [24-25]. This process is just like an extrusion process where a billet is extruded over a twist die. The benefit of using this one is its high expanded volume. Unreliable distortion is main limitation of using this process as encountered in HPT, where lesser distortion occurred nearby the extrusion axis. Furthermore it is observed by researcher that this process is not proficient than ECAP or HPT [26].

3. Effects of various SPD techniques on different Materials

As per the discussion it has been clear that SPD techniques are used for imparting the grain boundaries andstrengthening and developing ultrafine grains (UFG) in metals. While doing the experiments it is observed by various researchers that most of the cases when the strength of the materials enhanced in that particular time the ductility decreases but for Age-Hardenable alloys, at the same time there is a edifice enhancement of strength and ductility but in case of Non-Age Hardenable alloys, there is no development of strength and ductility. The observation made from table is while single ECAP process is done by researchers there is no up gradation of strength and ductility but when ECAP processed with Cryo treatment shows a strength and ductility enhancement. Creep behaviour of the material influence by the passes being processed by ECAP as observed by researchers. As ECAP followed by Annealing also shows the transformation of grains which promotes material thermal immovability. Observation made by researcher that on SPD with parallel channel and it is observed by researchers when SPD-PC combined with one matrix of

direct extrusion in non- axisymmetric section that will give an improved microstructure with enhanced mechanical properties [57].

Table 1: Properties of Materials during SPD techniques at several conditions

S.No	Techniques	Material	C.Y.S(MPa)	Ultimate Tensile Strength (MPa)	Elongation	Endurance Limit(MPa)	Ref.
1	HR and ECAP, ST+ECAP+Q	AZ31	115 180	251 286	27 9.5	95 40	27 28
2	ECAP	AA1050	-	-	-	50-55	29
3	ECAP at 150°C ECAP + annealing at 200°C	AA5052	394 350	421 370	9 10.5	- -	31 30
4	ECAP at 150°C ECAP at 110°C	AA5056	280 392	340 442	25 7	116 116	30
5	ST at 350°C and ECAP at 200°C	AA5053	276	352	20		32
6	ECAP and Hot Rolling	AA7075	650 240	720 375	8.4 29	- 150	33
7	ST+ECAP+Ag at 190°C for 4hrs	AA6106d + 0.1Zr + 0.5Sc	570	590	9	225	34
8	CR ECAP at 400°C	Ti (grade 2)	380 640	460 810	26 15	240 380	35
9	CR 87% ECAP at 400°C	Ti (grade 2)	970	1050	8	420	36
10	CR	Ti (grade 4)	530	700	25	350	37
11	ECAP	Cu	375	385	-	170	38
12	ECAP and Aging at	Cu-0.36C	438	454	23	180	39

	500°C , 1hrs	r					
13	ECAP+PC	Al60 603 alloy	210	230	8	-	57

4. Discussion on properties followed by of SPD Techniques

4.1 Strength and Ductility

Table 1 shows, "strength and ductility" are an important property of the materials. Properties of material are actually liable on the grain size because these actually affected by the SPD techniques used, more than any other Mechanical Properties. Besides, a lot of of the properties are unswervingly administered using the ductility and strength [33]. It's very challenging to serve ductility and strength at the same time of the operation. When strength enhances the ductility of the material reduced as most of the cases shown in table 1. Loss of ductility comes from the combined effect of high flow stress and low strain hardening. It is observed by the researchers that for Age hardenable alloys, there is a structure improvement with strength and ductility but in Non-Age hardenable alloys, there is no improvement in strength and ductility [35]. It has been shown that elongation through SPD techniques does not improve however plastic increases flow. While using only ECAP not improves the ductility and strength but ECAP followed by Cryo treatment will tremendously increases the strength and ductility of the material [37-38].

4.2 Fatigue and Creep Behaviour

Fatigue and Creep Behaviour of Material is also important after Strength and Ductility property and to analyse this property of the material is not easy too. By the Previous research it has been observed that the Fatigue behaviour of material does not exhibit the stronger grains [40-42].After doing research it has been observed that when ECAP process combined with other Temperature Mechanical processes that will results the fatigue of UFG metals. There is less research in the field of Creep and Fatigue of UFG material Produced by SPD Techniques [9]. Many Researcher put emphasis on factors which shows the effect of creep on materials processed by ECAP technique [43-45].By the experiment it has been observed that creep Behaviour of the material depends upon the number of passes. It owes to the number of influences counting microstructural deviations,coalesce of Microstructure and Nano porosity prompted through ECAP.

4.3 Thermal stability

Along with strength and ductility, thermal Stability is also important parameter. To maintain various property of the material at the same time is very difficult.According to the needed application of the material, a complete list of properties has to be maintained [46].From the research it has been investigated by researcher that most of the cases SPD treated material i.e.oxygen free copper offers poor thermal stability [47-48].Many experiments has been done to improve the thermal stability with multi-functional properties of the materials. It is clearly shown in Experiments Recovery rate is depends on the number of ECAP passes [48-49]. Copper manufactured from SPD shows a low thermal stability as there is no change in microstructure but when the temperature gets increases normal grain growth takes place and after doing annealing it is shown by experiments there is a transformation of UFG grains into Coarse Grains [50]. To overcome this

limitation and for maintain the multifunctional properties in SPD various processes has been introduced like grain refinement, strain hardening, solid solution hardening and precipitation hardening [48-51].

4.4 Corrosion Resistance

On behalf of future engineering prospects, Corrosion resistance is an important property and to maintain all the property of the materials is challenging task too [52]. The property of corrosion in a single phase in metals is predominantly influence by grain size [53]. Corrosion in a material could be take place in the aspects of pitting, chemical, electrochemical, stress corrosion cracks (SCC) and corrosion fatigue [52-54].From the investigation it is noted ECAP techniques gives a better decision by increasing the Mechanical properties of the material it also enhance the SCC and corrosion fatigue resistance [55-56]. Problem is faced by various researchers while examining the effects of SPD on corrosion behaviour of Magnesium due to the processing temperature while doing SPD [58-64].

5. Conclusion

From the above research it has been perceived that numerous research has been focused on the properties like Strength and ductility but at the same time to maintained the overall properties of the material is very difficult, By using SPD with some post processes the properties of the materials can enhanced. This paper gives a brief history about SPD techniques, their approaches and assets of SPD processed UFG materials. This content will also help the researchers as a summary of SPD techniques.This is also providing a Basic knowledge of problems face by various industrial applications. A large number of concept is developed by various researched but somehow justification is missing in some concepts. However it also gives the relationship between strength ductility balanced by using SPD techniques. The improvement of Corrosion resistance and thermal stability has been carried out by using SPD techniques if the number of ECAP passes will increase. SPD technique mainly extended from Conventional metal working methods as this is further established for handling bulk materials. Now days this technique is further used for various purposes like compaction of powders, for developing alloys etc.

6. REFERENCES

[1]. H. Ng, S. N. M. Yahaya, A. A. A. Majid “Reviews on aluminum alloy series and its applications” Academia Journal of Scientific Research 5(12), December 2017, 708-716.

[2]. John Vivek, Pant Ruby, AggarwalSaurabh&AgrawalPiyush, “Parametric Analysis and Effect of Tool on FSW 6082 al alloy by taguchi method”International Journal of Mechanical and Production Engineering Research and Development (IJMPERD) ISSN (P): 2249-6890; ISSN (E): 2249-8001 Vol. 8, Issue1, Feb 2018, 105-110.

[3]. Hui W, LUO Y, Friedman P, CHEN M, Lin GAO “Warm forming behavior of high strength aluminum alloy AA7075. Trans Nonferrous Met. Soc. China” (2012). 22:1–7.

[4]. Sushanta Kumar Panigrahi, R. Jayaganthan, VivekPancholi, “Effect of plastic deformation conditions on microstructural characteristics and mechanical properties of Al 6063 alloy”, Materials and Design 30 (2009) 1894–1901.

[5]. Mr.ShivkumarPanjab,Mr.JigarPaghadal “To Study about the behaviour of Mechanical Properties of the Material in Cryo-rolling process: A Review” International Journal for Innovative research in multidisciplinary field, Volume - 2, Issue - 8, Aug – (2016), ISSN – 2455-0620.

[6]. Wang JT. Nanomater: "Historic retrospection and present status of severe plastic deformation in China".Severe Plastic Deformat 2006;503–504:363.
[7]. Tanimura H. J: "Development of the Japanese sword". Met 1980;32(2):63–73.
[8]. FatihKahraman, "Surface layer properties of ultrasonic impact–treated AA7075 aluminum alloy",https://journals.sagepub.com/doi/10.1177/0954405416685386 2018,Vol 232, Issue 12.
[9]. Sheik Hassan M, Sanjeev Sharma, Brijesh Kumar, "A Review of Severe Plastic Deformation", International Refereed Journal of Engineering and Science (IRJES) ISSN (Online) 2319-183X, (Print) 2319-1821,Volume 6, Issue 7 (July 2017), PP.66-85.
[10]. Ruslan Z Valiev, Terence G. Langdon, "Principles of equal-channel angular pressing as a processing tool for grain refinement", (September 2006), Progress in Materials Science 51(7):881-981.
[11]. Parshikov R.A.,Rudskoy A.I.,ZolotovA.M.,Oleg V. Tolochko, "Technological problems of equal channel angular pressing Reviews" on Advanced Materials Science, (October 2013), 34(1):26-36.
[12]. S. SenthilMurugan, "Equal Channel Angular Pressing: A Novel Technique for the Production of Ultra Fine Grained Structure in Materials– A Mini Review", International Journal of Modern Studies in Mechanical Engineering (IJMSME) Volume 3, Issue 1, (2017), PP 6-14 ISSN 2454-9711.
[13]. KavehEdalati and ZenjiHorita, "A Review on High-Pressure Torsion (HPT) from 1935 to 1988", (January 2016)Materials Science and Engineering A 652:325-352.
[14]. Z. Horita (ed.), "Production of Multifunctional Materials Using Severe Plastic deformation, Proceedings of the International Symposium on Giant Straining Process for Advanced Materials", (Kyushu University Press, Fukuoka, 2010).
[15]. R.Z. Valiev, private communication, (2014)].
[16]. Zhilyaev AP, Langdon TG. "Using high-pressure torsion for metal processing: Fundamentals and applications".Prog Mater Sci 2008;53:893.
[17]. Sheik Hassan M, Sanjeev Sharma, Brijesh Kumar, "A Review of Severe Plastic Deformation" International Refereed Journal of Engineering and Science (IRJES) ISSN (Online) 2319-183X, (Print) 2319-1821 Volume 6, Issue 7 (July 2017), PP.66-85.
[18]. Vorhauer A, Pippan R: "On the homogeneity of deformation by high pressure torsion". Scripta Mater 2004;51:921.
[19]. Saito Y, Utsunomiya H, Tsuji N, Sakai T: "Novel ultra-high straining process for bulk materials development of the accumulative roll-bonding (ARB) process". Acta Mater 1999;47:579.
[20]. Valiahmetov OR, Galeyev RM, Salishchev GA: "Mechanical properties of VT 8 Ti-alloy of sub microcrystalline structure VT 8", Fiz Met Metalloved1990;10:204 (in Russian).
[21]. Kaibyshev OA: "Grain refinement in commercial alloys due to high plastic deformations and phase transformations". J Mater Process Technol 2001;117:300.
[22]. Beygelzimer Y, Varyukhin VN, Synkov SG, SapronovAN,Synkov VG: "New schemes of large plastic deformations accumulating with using of hydro extrusion". PhysTechnol High Press 1999;9:109.
[23]. Manjunath G.A., Shivakumar S. , Russell Fernandez , Nikhil R. , Sharath P.C. , "A review on effect of multi-directional forging/multi-axial forging on mechanical and microstructural properties of aluminum alloy", Materials Today: Proceedings journal homepage: www.elsevier.com/locate/matpr.

[24]. Beygelzimer Y, Varyukhin VN, Synkov SG, SapronovAN,Synkov VG: “New schemes of large plastic deformations accumulating with using of hydro extrusion” PhysTechnol High Press 1999;9:109.
[25]. Beygelzimer YY, Orlov DV: “High pressure effects in chemistry”. BiolMaterSci 2002;208-2:311.
[26]. Orlov D, Beygelzimer Y, Synkov S, Varyukhin V, Tsuji N, Horita Z: “Plastic flow, structure and mechanical properties in pure Al deformed by twist extrusion” .Mater SciEng A 2009;519:105.
[27]. Kim H-K, Lee Y-I, Chung C-S: “Fatigue properties of a fine-grained magnesium alloy produced by equal channel angular pressing”. Scripta Mater 2005;52:473.
[28]. Vinogradov A, Orlov D, Estrin Y: "Improvement of fatigue strength of a Mg–Zn–Zr alloy by integrated extrusion and equal-channel angular pressing”. Scripta Mater 2012;67:209.
[29]. Tsai TL, Sun PL, Kao PW, Chang CP: “Microstructure and tensile properties of a commercial 5052 aluminum alloy processed by equal channel angular extrusion”. Mater SciEng A2003; 342:144.
[30]. Patlan V, Vinogradov A, Higashi K, Kitagawa K:” Overview of fatigue properties of fine grain 5056 AlMg alloy processed by equal-channel angular pressing”. Mater SciEng A2001; 300:171.
[31]. Raab GJ, Valiev RZ, Lowe TC, Zhu YT: “Continuous processing of ultrafine grained Al by ECAP” Conform. Mater SciEng A2004; 382:30.
[32]. sochowski A, Olejnik L: “Incremental equal channel angular pressing for grain refinement”. Mater Sci Forum 2011; 674:19.
[33]. Markushev M. Vinogradov A. In: Altan BS, Miskioglu I, PurcekG,Mulukov R, Artan R, editors: “Room-Temperature Mechanical Properties of Sub microcrystalline Commercial Aluminum Alloys Processed by Severe Plastic Deformation",USA: Nova Science Publishers; 2006. p. 233.
[34]. Cavaliere P, Cabibbo M: Effect of Sc and Zr additions on the microstructure and fatigue properties of AA6106 produced by equal-channel-angular-pressing”. Mater Charact 2008; 59:197.
[35]. Stolyarov VV, Alexandrov IV, Kolobov YR, Zhu M, Zhu T, LoweT. In: Wu XR, Wang ZG, editors. Fatigue ’99, vol. 3. PRChina: Higher Education Press; 1999. p. 1345.
[36]. Semenova I, Valiev R, Yakushina E, Salimgareeva G, Lowe T: “Strength and fatigue properties enhancement in ultrafine-grained Ti produced by severe plastic deformation”. JMaterSci 2008;43:7354.
[37]. Han BQ, Lavernia EJ, Mohamed FA: “Dislocation structure and deformation in iron processed by equalchannel-angular pressing”. Metall Mater Trans A2004; 35A:1343.
[38]. Jacques P, Furnemont Q, Pardoen T, Delannay F: “On the role of martensitic transformation on damage and cracking resistance in TRIP-assisted multiphase steels”. Acta Mater2001; 49:139.
[39]. HoppelHW, May J, Eisenlohr P, Goken A, “Strain-rate sensitivity of ultrafine-grained materials”. Z Metall 2005; 96:566.
[40]. Vinogradov AY, Agnew SR, Schwarz JA, Contescu CI, PutyeraK,: “Nanocrystalline materials: fatigue and nanotechnology”,vol. 1.1. London: Taylor & Francis; 2005. p. 1.
[41]. Mughrabi H, Hoppel HW: “Cyclic deformation and fatigue properties of very fine-grained metals and alloys”. Int J Fatigue 2010; 32:1413.
[42]. Valiev R: “Principles of producing bulk nanostructured metals with unique properties by severe plastic deformation techniques”. In: 17th International offshore and polar

engineering conference (ISOPE 2007). Lisbon: International Society Offshore &Polar Engineers; 2007. p. 2858.

[43]. Sklenicka V, Krai P, Illucova L, Saxi I, Dvorak J, Svoboda M: “Inhomogeneity of microstructure and creep of ECAP Aluminium. Nanomater Severe Plastic Deform” 2006; 503–504:245.

[44]. Sklenicka V, Dvorak J, Kvapilova M, Svoboda M, Krai P, Saxi I: “Effect of equal-channel angular pressing (ECAP) on creep in aluminium Alloys”.Adv Mater Res 2008; 15–17:2904.

[45]. Hellmig RJ, Janecek M, Hadzima B, Gendelman OV, Shapiro M,Molodova X, et al: “A portrait of copper processed by equal channel angular pressing”. Mater Trans 2008; 49:31.

[46]. Islamgaliev RK, Chmelik F, Kuzel R: “Thermal stability of submicron grained copper and nickel” Mater SciEng A 1997; 237:43.

[47]. Cizek J, Prochazka I, Cieslar M, Kuzel R, Kuriplach J, Chmelik F: “Thermal stability of ultrafine grained copper”. Phys Rev B 2002; 65:094106.

[48]. Molodova X, Gottstein G, Winning M, Hellmig RJ: “Thermal stability of ECAP processed pure copper”. Mater SciEngA 2007;460–461:204.

[49]. Park KT, Kim YS, Lee JG, Shin DH: “Thermal stability and mechanical properties of ultrafine grained low carbon steel”. Mater SciEng A2000; 293:165.

[50]. Hoseini M, Hamid Pourian M, Bridier F, Vali H, SzpunarJA,Bocher P: “Thermal stability and annealing behaviour of ultrafine grained commercially pure titanium”. Mater SciEng A 2012; 532:58.

[51]. Hoseini M, Bridier F, Bocher P, Vali H, Szpunar JA: “Thermal stability and grain growth behavior of ultrafine grained commercially pure titanium fabricated by equal channel angular pressing”. Science Press 2012;1:666.

[52]. Yamasaki T, Miyamoto H, Mimaki T, Vinogradov A, Hashimoto S: “Corrosion Fatigue of Ultra-Fine Grain Copper Fabricated by Severe Plastic Deformation” .In: TMS; 2002. p. 361.

[53]. Yamasaki T, Miyamoto H, Mimaki T, Vinogradov A, Hashimoto S: “Stress corrosion cracking susceptibility of ultra-fine grain copper produced by equal-channel angular pressing” .Mater SciEng A 2001;318:122.

[54]. Miyamoto H, Harada K, Mimaki T, Vinogradov A, Hashimoto S: “Corrosion of ultra-fine grained copper fabricated by equal-channel angular pressing.” CorrosSci 2008; 50:1215.

[55]. Hadzima B, Janecek M, Hellmig RJ, Kutnyakova Y, Estrin Y: “Microstructure and corrosion behaviour of ultrafine-grained copper”. Mater Sci Forum 2006; 503–504:883.

[56]. Son IJ, Nakano H, Oue S, Kobayashi S, Fukushima H, Horita Z: “Effect of equal-channel angular pressing on the pitting corrosion resistance of Al alloy” .Mater Sci Forum 2006;503–504:487.

[57]. E. V. Bobruk, A V Botkin, Maxim Yu. Murashkin, G I Raab, “A new combined SPD technique to improve mechanical properties and electrical conductivity of long-sized billets”,January 2021, IOP Conference Series Materials Science and Engineering 1014(1):012004 DOI: 10.1088/1757-899X/1014/1/012004.

[58]. M.Kasaeian-Naeini, M.Sedighi, R.Hashemi, “Severe plastic deformation (SPD) of biodegradable magnesium alloys and composites: A review of developments and prospects”,8 December 2021,Journal of Magnesium and Alloys.

[59] Malik, Praveen, Singh, R., Gehlot, A., Akram, S. V., & Das, P. K. (2022). Village 4.0: Digitalization of Village with Smart Internet of Things Technologies. Computers & Industrial Engineering, 107938.

[60] Bagwari, S., Gehlot, A., Singh, R., Priyadarshi, N., & Khan, B. (2022). Low-Cost Sensor-Based and LoRaWAN Opportunities for Landslide Monitoring Systems on IoT Platform: A Review. IEEE Access, 10, 7107–7127. https://doi.org/10.1109/ACCESS.2021.3137841

[61] Malik, P, Gehlot, A., Singh, R., Gupta, L. R., & Thakur, A. K. (2022). A Review on ANN Based Model for Solar Radiation and Wind Speed Prediction with Real-Time Data. Archives of Computational Methods in Engineering. https://doi.org/10.1007/s11831-021-09687-3

[62] Singh, R., Gehlot, A., Rashid, M., Saxena, R., Akram, S. V, Alshamrani, S. S., & Alghamdi, A. S. (2021). Cloud server and internet of things assisted system for stress monitoring. Electronics (Switzerland), 10(24). https://doi.org/10.3390/electronics10243133

[63] Karunanidy, D., Ramalingam, S., Dumka, A., Singh, R., Rashid, M., Gehlot, A., ... & AlGhamdi, A. S. (2022). JMA: Nature-Inspired Java Macaque Algorithm for Optimization Problem. Mathematics, 10(5), 688.

[64] Sharma, V. K., Singh, R., Gehlot, A., Buddhi, D., Braccio, S., Priyadarshi, N., & Khan, B. (2022). Imperative Role of Photovoltaic and Concentrating Solar Power Technologies towards Renewable Energy Generation. International Journal of Photoenergy, 2022.

Human Computer Intelligent Interaction Based on sEMG Signals for Upper Limb Movements

[1]Dr. Kapil Joshi, [2]Dr Ritesh Rastogi, [3]Dr. Gunjan Chhabra, [4] Sheeba Praveen, [5]Rajesh Singh, [6]Rohit Kanauzia, [7]Michael G. Albino
[1]Department of CSE, UIT, Uttaranchal University, Dehtadun, India
[2](NIET) Noida institute of engineering and technology, Greater Noida, India
[3]Department of Systemics, school of computer science, UPES, Dehradun, India
[4]Computer Science Engineering, Integral University, Lucknow, India
[5]UIT, Uttaranchal University, Dehradun, India
[6]Department of CSE, College of Engineering Roorkee, India
[7]College of Communication and Information Technology President Ramon Magsaysay State University, Philippines

Abstract.

The role of human computer interaction in our lives has grown manifold in recent times. There is a requirement for more efficient, intelligent and concise interaction between humans and computers. It is significant to optimize this interaction by using efficient methods of calculation. For eliminating the data of interference for biometric recognition, and the simplification of the data processing, whilst improving efficiency of working of usual equipments for sEMG signals, the gesture recognition process for biometric scanning uses an approach based on the theory of variance. The action signals in all five groups become redundant and are visualized and divided into sixteen levels. The result comparison of recognizing the motion of thumb after removal of various channels which are redundant, the combination of optimal channels is acquired. Lastly, two types of classifiers apt for fields of sEMG signals are chosen and the outcomes of classification are weighed against each other, out of which the most efficient recognition pattern for thumb motion is concluded.

Keywords. Computer Intelligent System, sEMG signals, Human Intelligent.

1. INTRODUCTION

The process of simplification of the interaction between humans and computers is a hotspot in research for intelligent control Since human gestures are diverse, the human computer interaction development technology on the basis of recognizing gestures becomes increasingly significant [1,2]. sEMG (Surface Electromyography) is developed steadily in the human computer interaction wave. To perceive human motion, the reflection is on extension and flexion functions of the joints of hand and the information which is dynamic is reflected too, like the intensity and position of the limbs to complete certain actions. For thumb motion recognition, the issue caused by the electrode redundancy through sEMG is solved with the use of statistical Variance

Theory. This paper innovates the usage of statistical theory for eliminating the interfering signals formed by the channels which are redundant while acquiring certain actions specifically by universal sEMG signal usage through the equipment for their acquisition. Taking the research object of sEMG acquisition equipment with 16 channels, the electrodes that are redundant commonly are obtained with identification of the fine motion of the thumb for nine targets [3,4,5]. This research reduces the hardware cost and gives a reference to optimize the distribution of electrodes. It reduces the data processing overheads and saves time.

2. RELATED RESEARCH

Currently the researches in the field of recognition of human gestures focuses majorly on recognizing and classifying of gestures on large scale [6]. In everyday life, hand movements that are subtle with subtle arm, wrist and finger motions [7,8]. The thumb functions more than rest of the fingers do and in daily chores, it has the most functions. It is important to grasp at things and affects the gestures, position and motion of the hand on the whole. For rehabilitation sciences, the thumb movements can symbolize the complete hand posture and the hand movements to a limit and the thumb movements need to be studied deeply to understand the characteristics of motions for the control of the whole hand [9,10]. For human – computer interaction, a number of movements in a subtle manner of the thumb can facilitate the interactions of humans with computer [11,12,13,14]. Surface sEMG based gesture recognition is used for mainly collecting electrical signals which are generated on skin's surface by the equipment used for EMG acquiring. These are then classified and the data is recognized after extracting features. Recently, the myoelectric signals have been studied by many scholars in the human computer interaction field [15]. The influence of classification and feature algorithm was studied by Xun Chen et. al for accuracy of recognition. This experiment included the use of four channel acquisition equipment for sEMG and the signals for forearm sEMG were used for collecting ten gestures of the Chinese language. With the combination of the traditionally used features with the newer algorithm for classification, the accuracy of recognition of the motions of the hand has been 95% more improved [16]. sEMG signals were used by Jongin Kim et. al. for identifying the space scaling between user's thumb and the index finger on the screen of a device [17]. Four features were extracted by Chengcheng Li et. al out of the 9 gestures and they used an SVM classifier for recognizing gestures to a 98% accuracy [18].

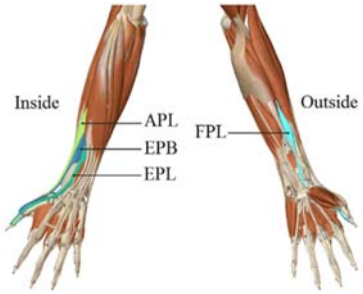


Figure 2.1. Muscle group on Outside [8]

A tactile sensor was used by G.F. Li et. al for the process of recognition of EMG gestures [19]. A device was designed by Nor Anija Jalaludin et. al. for detection of the sEMG strength and the sEMG model was established to decipher the dependability of the force of thumb and its angle [20]. The hand tools were optimized by Essam Odah et. al. for testing of the strain with various positions of sEMG signals comparatively [21, 22]. The thumb is connected to the muscle groups which can be divided into the external and internal groups of muscles [23, 24]. The hand includes most of the inner adnexal group of muscles and the outer one is distributed in the arm mainly. The inside group of muscles constitutes of the abductor pollicis bevis, abductor pollicis brevis, opponent pollicis and first dorsal interossus. The outside muscle group includes the extensor pollicis brevis, flexor pollicis longus and the abductor pollicis longus, as seen from Figure 2.1.

3. METHODS AND MATERIALS

3.1. Acquisition Equipment Selection

Equipment to acquire sEMG function mainly for storing and filtering the electrical signals which are generated by the surface of the skin. As EMG signals are weak and can be interfered with easily along with the bioelectric signals, the signal quality relates closely with the acquisition equipment generally [25,26]. The equipment selected in this research includes an 18 channel technology with a technique for spatio temporal resolution sampling, which is compatible with both wet and dry electrodes and communication technology based on Bluetooth wireless signals, used to analyze gait, muscles' fatigue, gesture recognition, rehabilitating treatment, etc. These equipments are very easily available in the market for collection of the sEMG signals of the upper limbs. Most research tasks are met on the basis on EMG signals. From figure 2 one can see the equipment used for the research. The system for signal acquisition used in this paper has 17 dry electrodes used in a fabric form, which makes it flexible and 15 sEMG bipolar channels are thus formed, thus reducing the artifact use, one of these electrodes is used for grounding and the other electrode is used for reference. Rest of the 15 electrodes are distributed evenly in the sleeve.

Figure 3.2. Acquisition Equipment [16]

3.2. Process Design for Experiment

For fingers, there is a well established international standard which judges the flexibility and the standard's core is if the thumb is able to touch easily rest of the 4 fingers. Hence, five gestures which are dynamic, are dynamic, are selected on the basis of the norms [32,33]. These gestures are UP (thumb moves upward – relax), DOWN (Thumb down – relax), LEFT, RIGHT and PRESS. Figure 3.2 shows the actions in all the five gestures. Ten subjects were used, with 3 women and 7 men between ages of 25 to 35 , who were healthy, with no motor nerve issue history and the activities were of higher intensity in the time preceding the research. For acquiring data, the test needs the subjects to be in a sitting position, with elbows over the table and naturally suspended forearms. Every subject measure six data sets, with six gestures types, each repeated 12 times. This data was collected over a week, forming the dataset.

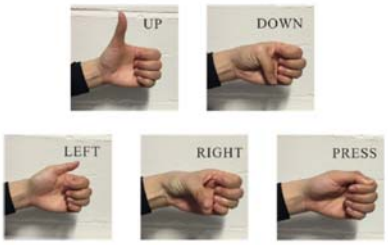


Figure 3.2. Types of Gestures [9]

4. WAVE FILTERING

This operation filters out the regularity in a specifically selected band of the gesture and it is significant to gauge and suppress the obstruction. The discrete output signal is converted by the digital filter into another one, based on the real time need, as shown in figure 4.1.

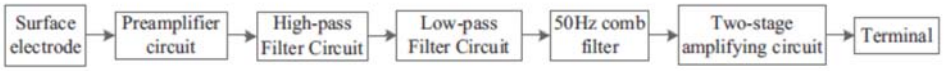


Figure 4.1 Flow of Signal Processing

Since the sEMG indicator is weak, it is imperative to strengthen the electric signals assembled by the first electrode all the way through the circuit for pre-amplification. The next step includes using a high pass filter and using of low pass filter for the noise for wake in 20 Hz and above 500 ambient noise. The filtered signal is finally amplified doubly for fitting in the equipment used for acquisition.

4.1. Detecting Active Segment

All the gestures in the experiment are dynamic and the time used for relaxing while the data is being collected is not very precise. So it is important to get rid of the signals which are redundant with the relaxation time, and the signals of the gestures that are intercepted should also consist of the complete process of movement, and the job of the detection of the segment of activity is determining the beginning and conclusion of the gesture [37, 38, 39]. The method used for moving the average has a defined width of the window for sliding of the signal and comparison of the sequence of immediate energy with a threshold value found in real time for determining the starting and ending points of the signal for gestures [40]. The sEMG signal value alters from +ve to –ve in considerably shorter span and the random fluctuations are very large, so that the method for moving average is suitable for elimination of the fluctuation of interference and getting the signals's overall trend.

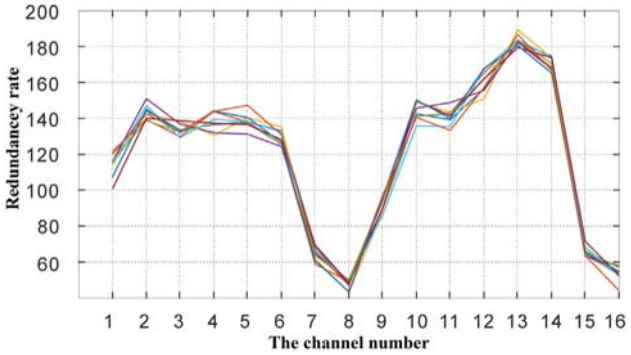


Figure 4.2. Weighted Redundancy of Channels

4.2. Extracting Features

The best features are selected for the methods using time and domain, which are the feature with the absolute mean value, zero crossing feature, feature with the change in slope sign, autoregressive mode, slope sign change feature and the root mean square feature [42,43,44]. Through the validation of clustering, one finds that the amalgamation of RMS, AR and SSC, gives the best effect for classification.

5. RESULTS AND DISCUSSIONS

For achieving better results, there is the need for a classifier to be selected, which has certain advantageous [45] features. Presently, the classifiers used for mainstream classification of limb motions, include LDA or Linear Discriminant Analysis, ANN or Artificial Neural Network, GMM [46-48] or Gaussian Mixture Model, SVM or Support Vector Machine, GRNN or Generalized Regression Neural Network, HMM or Hidden

Markov Model, etc. Out of all these, SVM gives a much better performance and the meticulous mathematical theory it has and the performance of good classification renders it to be used widely for electromyography [49]. Hence the preferred choice for this experiment is SVM. GRNN has a strong ability for both [50] approximation and classification, making the learning rate better than RBF Neural Network (Radial Phase Foundation) and Neural Network with a Back Propagation [51-55] Algorithm (BP), and performance is better. The noise data classification problem uses a GRNN as the second choice classifier.

6. CONCLUSION

To begin with, the dynamic gestures are defined; the raw data collected and subsequently is filtered. Futher, the detection of the active segment is performed, and the feature selection is done. This was followed by the calculation of, the redundancy for every channel's electrode, using dispersion theory. Using a simple coding method, three kinds of overlapping combinations for the electrodes generated. The reference classifiers were taken as the SVMs and 12, 13, and 14 channels were eliminated and the best channels were removed. Finally, the comparison between the effects of classification of GRNN and SVM were done in 16 and 13 channels, proving that the overlapping electrode which was removed, is the electrode which overlaps commonly.

7. REFERENCES

[1] G.F. Li, H. Wu, G.Z. Jiang, S. Xu, H.H. Liu, Dynamic gesture recognition in the internet of things, IEEE Access 7 (2019) 23713–23724.

[2] M. Balouchestani, S. Krishnan, Robust compressive sensing algorithm for wireless surface electromyography applications, Biomed. Sign. Process. Control 20 (2015) 100–106.

[3] Z. Jiang, Y. Jiang, Y. Wang, H. Zhang, H.J.C.G.D. Tian, A hybrid approach of rough set and case-based reasoning to remanufacturing process planning, J. Intell. Manuf. 30 (1) (2019) 19–32.

[4] G.R. Naik, S.E. Suviseshamuthu, M. Gobbo, Principal component analysis applied to surface electromyography: a comprehensive review, IEEE Access 4 (2016) 4025–4037.

[5] H. Wang, Z.G. Jiang, H. Zhang, Y. Wang, Y.H. Yang, Y. Li, An integrated MCDM approach considering demands-matching for reverse logistics, J. Clean. Prod. 208 (2019) 199–210.

[6] D.S. Chen, G.F. Li, Y. Sun, J.Y. Kong, G.Z. Jiang, H. Tang, Z. J. Ju, H. Yu, H.H. Liu, An interactive image segmentation method in hand gesture recognition, Sensors 17 (2) (2017) 253.

[7] S. Mustafa, I. Elghandour, M.A. Ismail, A machine learning approach for predicting execution time of spark jobs, Alexandr. Eng. J. 57 (1) (2018) 3767–3778.

[8] M. Tavakoli, C. Benussi, J.L. Lourenco, Single channel surface EMG control of advanced prosthetic hands: a simple, low cost and efficient approach, Exp. Syst. Applic. 79 (2017) 322–332.

[9] J.B. Hu, Y. Sun, G.F. Li, G.Z. Jiang, B. Tao, Probability analysis for grasp planning facing the field of medical robotics, Measurement 141 (2019) 227–234.

[10] G.F. Li, D. Jiang, Y.L. Zhou, G.Z. Jiang, J.Y. Kong, Gunasekaran Manogaran, Human lesion detection method based on image information and brain signal, IEEE Access 7 (2019) 11533–11542.
[11] A.R. Siddiqi, S.N. Sidek, A. Khorshidtalab, Signal Processing of EMG Signal for Continuous Thumb-Angle Estimation, IECON, 2015, pp. 374–379.
[12] J.X. Qi, G.Z. Jiang, G.F. Li, Y. Sun, B. Tao, Intelligent human computer interaction based on surface EMG gesture recognition, IEEE Access 7 (2019) 61378–61387.
[13] Y.F. Fang, D.L. Zhou, K.R. Li, H.H. Liu, Interface prostheses with classifier-feedback-based user training, IEEE Trans. Biomed. Eng. 64 (11) (2017) 2575–2583.
[14] R. Menon, C.G. Di, H. Lakany, L. Petropoulakis, B. Conway, J. Soraghan, Study on interaction between temporal and spatial information in classification of EMG signals in myoelectric prostheses 1–1, IEEE Trans. Neural Syst. Rehabil. Eng. (2017), https://doi.org/10.1109/TNSRE.2017.2687761.
[15] Y. Sun, C.Q. Li, G.F. Li, G.Z. Jiang, D. Jiang, H.H. Liu, Z.G. Zheng, W.N. Shu, Gesture recognition based on kinect and sEMG signal fusion, Mobile Netw. Applic. 23 (4) (2018) 797–805.
[16] X. Chen, Z.J. Wang, Pattern recognition of number gestures based on a wireless surface EMG system, Biomed. Signal Process. Control (2013) 184–192.
[17] J. Kim, D. Cho, K. Lee, B. Lee, A real-time pinch-to-zoom motion detection by means of a surface EMG-based human computer interface, Sensors 15 (1) (2014) 394–407.
[18] C.C. Li, G.F. Li, G.Z. Jiang, D.S. Chen, H.H. Liu, Surface EMG data aggregation processing for intelligent prosthetic action recognition, Neural Comput. Appl. (2018), https://doi.org/10.1007/s00521-018-3909-z.
[19] G.F. Li, L.L. Zhang, Y. Sun, J.Y. Kong, Towards the sEMG hand: internet of things sensors and haptic feedback application, Multim. Tools Applic. 78 (21) (2019) 29765–29782.
[20] N.A. Aranceta, B.A. Conway, Differentiating variations in thumb position from recordings of the surface electromyogram in adults performing static grips, a proof of concept study, Front. Bioeng. Biotechnol. 7 (2019), https://doi.org/10.3389/fbioe.2019.00123.
[21] E. Odah, I. Abu-Qasmieh, N. Al Khateeb, E. Al Matalbeh, S. Quraan, M. Mohammad, Gender considerations in optimizing usability design of hand-tool by testing hand stress using SEMG signal analysis, Alexandr. Eng. J. 57 (4) (2018) 2897–2901.
[22] K. Usha, M. Ezhilarasan, Fusion of geometric and texture features for finger knuckle surface recognition, AEJ – Alexandr. Eng. J. 55 (1) (2016) 683–697.
[23] S. Kim, J. Kim, M. Kim, S. Kim, J. Park, Grasping force estimation by sEMG signals and arm posture: tensor decomposition approach, J. Bion. Eng. 16 (3) (2019) 455–467.
[24] R.R. Muhammad, N.S. Shahrul, F.T. Siti, Portable thumb training system for EMG signal measurement and analysis, in: International Conference on Computer and Communication Engineering, 2016, https://doi.org/10.1109/ICCCE.2016.44.
[25] D. Jiang, G.F. Li, Y. Sun, J.Y. Kong, B. Tao, D. Chen, Grip strength forecast and rehabilitative guidance based on adaptive neural fuzzy inference system using sEMG, Pers. Ubiquit. Comput. (2019), https://doi.org/10.1007/s00779-019-01268-3.

[26] Z.O. Khokhar, Pattern Recognition of Surface Electromyography Signals for Real-time Control of Wrist Exoskeletons, vol. 9, 2010, pp. 24–41

[27] D. Jiang, Z.J. Zheng, G.F. Li, Y. Sun, J.Y. Kong, G.Z. Jiang, H. G. Xiong, B. Tao, S. Xu, H.H. Liu, Z.J. Ju, Gesture recognition based on binocular vision, Clust. Comput. 22 (Supplement 6) (2019) 13261–13271.

[28] M.C. Yu, G.F. Li, D. Jiang, G.Z. Jiang, F. Zeng, H.Y. Zhao, D. S. Chen, Application of Pso-Rbf neural network in gesture recognition of continuous surface EMG signals, J. Intell. Fuzzy Syst. (2019), https://doi.org/10.3233/jifs-179535.

[29] Y.W. Cheng, G.F. Li, J.H. Li, Y. Sun, G.Z. Jiang, F. Zeng, H. Y. Zhao, D.S. Chen, Visualization of activated muscle area based on sEMG, J. Intell. Fuzzy Syst. (2019), https://doi.org/ 10.3233/jifs-179549.

[30] Y. He, G.F. Li, Y.J. Liao, Y. Sun, J.Y. Kong, G.Z. Jiang, D. Jiang, H.H. Liu, Gesture recognition based on an improved local sparse representation classification algorithm, Clust. Comput. 22 (Supplement 5) (2019) 10935–10946.

[31] E.N. Kamavuako, D. Farina, K. Yoshida, W. Jensen, Estimation of grasping force from features of intramuscular EMG signals with mirrored bilateral training, Ann. Biomed. Eng. 40 (3) (2012) 648–656.

[32] X. Zhai, B. Jelfs, R.H. Chan, Short latency hand movement classification based on surface EMG spectrogram with PCA, in: Conf Proc IEEE Eng Med Biol Soc, 2016, pp. 327–330.

[33] M. Kim, K. Kim, W.K. Chung, Simple and fast compensation of sEMG interface rotation for robust hand motion recognition 1–1, IEEE Trans. Neural Syst. Rehabil. Eng. (2018).

[34] Kasiprasad Mannepalli, Panyam Narahari Sastry, Maloji Suman, A novel adaptive fractional deep belief networks for speaker emotion recognition, Alexandr. Eng. J. 56 (4) (2017) 485–497. https://linkinghub.elsevier.com/retrieve/pii/S1110016816302484,https://doi.org/10.1016/j.aej.2016.09.002.

[35] H. Kalani, S. Moghimi, A. Akbarzadeh, SEMG-based prediction of masticatory kinematics in rhythmic clenching movements, Biomed. Sign. Process. Control 20 (2015) 24–34.

[36] T. Al, A. Ahmad, J. Al, Adel, Optimal Feature Set for Finger Movement Classification Based on SEMG, IEEE Engineering in Medicine and Biology Society, 2018, pp. 5228–5231.

[37] M.C. Yu, G.F. Li, D. Jiang, G.Z. Jiang, B. Tao, D.S. Chen, Hand medical monitoring system based on machine learning and optimal EMG feature set, Pers. Ubiquit. Comput. (2019),
https://doi.org/10.1007/s00779-019-01285-2.

[38] M.R. Al-Mulla, F. Sepulveda, Super wavelet for sEMG signal extraction during dynamic fatiguing contractions, J. Med. Syst. 39 (1) (2015) 167.

[39] T.F. Yang, J.W. Wang, Z.G. Hu, Variation of neck electromyography and mechanical characteristics during the process of cervical traction, J. Med. Biomech. 31 (5) (2016) 421–425.

[40] K. Veer, T. Sharma, A novel feature extraction for robust EMG pattern recognition, J. Med. Eng. Technol. 40 (4) (2016) 6.

[41] S. Guo, M. Pang, B. Gao, Comparison of sEMG-based feature extraction and motion classification methods for upper-limb movement, Sensors 15 (4) (2015) 9022–9038.

[42] J.X. Qi, G.Z. Jiang, G.F. Li, Y. Sun, B. Tao, Surface EMG hand gesture recognition system based on PCA and GRNN, Neural Comput. Appl. (2019), https://doi.org/10.1007/s00521-019-04142-8.
[43] O.A. Moura Karina, R.S. Ruschel, A. Balbinot, Fault-Tolerant Sensor Detection of sEMG Signals: Quality Analysis Using a Two-Class Support Vector Machine, IEEE Engineering in Medicine and Biology Society, 2018, pp. 5644–5647.
[44] P. Karthick, G. Venugopal, S. Ramakrishnan, Analysis of Muscle Fatigue Progression using Cyclostationary Property of Surface Electromyography Signals, Plenum Press, 2016,https://doi.org/10.1007/s10916-015-0394-0.
[45] G.F. Li, J.H. Li, Z.J. Ju, Y. Sun, J.Y. Kong, A novel featureextraction method for machine learning based on surfaceelectromyography from healthy brain, Neural Comput. Appl.31 (12) (2019) 9013–9022.
[46] G.F. Li, H. Tang, Y. Sun, J.Y. Kong, G.Z. Jiang, D. Jiang, B.Tao, S. Xu, H.H. Liu, Hand gesture recognition based on convolution neural network, Clust. Comput. 22 (Supplement 2) (2019) 2719–2729.
[47] M. Nisser, S. Derlien, U.C. Smolenski, N. Best, Measurement of muscular load with sEMG in using different computer input devices - a pilot study, Physikalische Medizin Rehabilitations medizin Kurortmedizin 28 (2018) 231–234.
[48] D. Kaneishi, R.P. Matthew, M. Tomizuka, A sEMG Classification Framework with Less Training Data, IEEE Engineering in Medicine and Biology Society, 2018, pp. 1680–1684.
[49] B. Zhou, X.L. Ji, R. Zhang, Design of surface EMG acquisition instrument based on embedded technology, Mod. Electr. Tech.33 (6) (2010) 55–57.
[50] B.W. Luo, Y. Sun, G.F. Li, D.S. Chen, Z.J. Ju, Decomposition algorithm for depth image of human health posture based on brain health, Neural Comput. Appl. (2019), https://doi.org/10.1007/s00521-019-04141
[51] Bagwari, S., Gehlot, A., Singh, R., Priyadarshi, N., & Khan, B. (2022). Low-Cost Sensor-Based and LoRaWAN Opportunities for Landslide Monitoring Systems on IoT Platform: A Review. IEEE Access, 10, 7107–7127. https://doi.org/10.1109/ACCESS.2021.3137841
[52] Malik, P, Gehlot, A., Singh, R., Gupta, L. R., & Thakur, A. K. (2022). A Review on ANN Based Model for Solar Radiation and Wind Speed Prediction with Real-Time Data. Archives of Computational Methods in Engineering. https://doi.org/10.1007/s11831-021-09687-3
[53] Singh, R., Gehlot, A., Rashid, M., Saxena, R., Akram, S. V, Alshamrani, S. S., & Alghamdi, A. S. (2021). Cloud server and internet of things assisted system for stress monitoring. Electronics (Switzerland), 10(24). https://doi.org/10.3390/electronics10243133
[54] Karunanidy, D., Ramalingam, S., Dumka, A., Singh, R., Rashid, M., Gehlot, A., ... & AlGhamdi, A. S. (2022). JMA: Nature-Inspired Java Macaque Algorithm for Optimization Problem. Mathematics, 10(5), 688.
[55] Sharma, V. K., Singh, R., Gehlot, A., Buddhi, D., Braccio, S., Priyadarshi, N., & Khan, B. (2022). Imperative Role of Photovoltaic and Concentrating Solar Power Technologies towards Renewable Energy Generation. International Journal of Photoenergy, 2022.

Biographies

Dr. Kapil Joshi received the Ph.D. degree in computer science & engineering from Uttaranchal University in 2021, the master's degree in Information Technology from Graphic Era University in 2014. He has published various patents, research papers in reputed journals (UGC/Scopus) as well as published two books in international publication. He was also a student member of IEEE UP section in session 2019-20 and editorial member, guest editor in various reputed national/international journals. He has been awarded with “Best Young Researcher" Award in Global Education and Corporate Leadership (GECL-21) received by Life Way Tech India Pvt. Ltd. in 2021, He also appointed as an Asia Coordinator in “International Association for the promotion of Asia-Africa Research” (IAPAAR) from 2020.

A Comparative Analysis of Cardiovascular Disease using Different Machine Learning Techniques

Mukesh Pandey[1], Satish Kumar Mahariya[2], Dr.Shiv Dayal Pandey3, Dr. Kapil Joshi[4], Dev Baloni[5],Jefferson A.Costales[6]
Assistant Professor Uttaranchal University Dehradun[1,2,3,4,5], Eulogio "Amang" Rodriguez Institute of Science and Technology, Manila, Philippines[6]

sitmukesh@gmail.com[1],satishmaharia191095@gmail.com[2],sdpmnnit@gmail.com[3],kapilengg0509@gmail.com[4], devbalini1982@gmail.com[5], jacostales@earist.ph.education

Abstract- Every year, around 18.6 million people die from cardiovascular diseases, the most common of which are myocardial infarctions and heart failure. MI and heart attacks are caused by persistent ischemia and hypoxia, which causes irreparable damage to myocardial tissue. An Effective, accurate, and feasible method for the timely diagnosis of these diseases is therefore needed for proper treatment. Heart disease early state can be identified with the help of multiple methods to use Several data mining, machine learning, deep learning algorithms Techniques. To simplify broad and complicated data analysis, the different algorithms models and methods of Artificial Intelligence can be implemented to different healthcare datasets. Indeed, this review's Details Analysis will be required to provide a methodological framework for the development of additional effective prediction models. The purpose of this paper is to review some of the recent research on cardiovascular disease prediction using various learning techniques. The goal of this study is to illustrate different algorithms of machine learning for cardiovascular disease prediction and find the best possible learning technique to identify the particular objective of cardiovascular disease.

***Keywords*:**Machine Learning, heart disease, accuracy, prediction. Data Mining, Artificial Neural Network

1. INTRODUCTION

Roughly 18.6 million people die each year from cardiac disease. Heart performs the most important part of bloodstream. Under continuous, diligent treatment, diabetes may prompt a collection of blood sugars, that could raise the danger of hazardous inconveniences including cardiovascular sicknesses. About 47% of unexpected deaths occur outside of an emergency department. Which suggests that several cardiovascular disease sufferers do not react to early warning signs. Heart problems have some primary signs and symptoms: chest pain and discomfort. Awe-inspiration. Nausea, flush or cold sweat. Installation's store enormous quantities of information in their directories, which is very complicated and demanding, due to the rapidly growing digital technologies. The techniques and algorithms on a dataset can be used directly to create specific models and identify important conclusions. Machine learning methods are therefore useful in predicting performance from actual data. Therefore, the whole work relates one such machine learning method, called classification, to predict risk from heart disease risk factors. This also seeks to increase the efficiency of cardiac risk assessment using a technique called an ensemble.

2. LITERATURE REVIEW

Numerous studies on disease prediction systems have been performed using various algorithms methods of Artificial Intelligence for the different kind of data sets

K. Polaraju et al, **[7]** Recommended cardiovascular infection study utilizing Multiple Clustering calculation and it exhibits that different straight relapse is suitable for foreseeing coronary illness chance. The analyses are carried out using the training data set, which contains 3000 occurrences of each of the 13 distinct attributes stated above. The data collection is divided into two areas, with 70% of the information used for planning and 30% for research. In light of the test, unmistakably the grouping precision of the Regression calculation is higher than some other AI calculations.

Marjia et al, **[8]** Expectation of coronary illness created utilizing WEKA programming, utilizing impression of Bayes Net, Multilayer K Star, j48 and SMO. The exhibition from one of a kind factor Bayes Net or SMO accomplishes ideal productivity than multilayer vision, KStar, and J48 procedures use k-crease cross approval. The exactness precision got by these calculations is as yet unsuitable. Thus, the proficiency of the precision is upgraded more to give better prescient ailment end.

Purushottam et al, [15] Suggested an effective system for predicting heart disease. This system will help health professionals make meaningful outcome-making dependent on the factor in issue.

S. Seema et al,[9] Focused on methods capable of detecting severe diseases by mining historical health information. A comprehensive algorithm analysis is conducted to test the increased output at an appropriate rate. SVM gives this investigation the most noteworthy exactness rating while Naïve Bayes offers the least accuracy.

Megha Shahi et al, [11] Emerging cardiovascular illness prescient framework using information mining systems; WEKA is utilized for programmed sicknessdistinguishing proof and administration quality conveyance in medicinal services communities The examination utilized different calculations, for example, Naïve Bayes, KNN, Alliance law , ANN, SVM and Tree of Judgment. The examination recommended by SVM is effective and gives more noteworthy accuracy contrasted with anotherparameters.

Chala Beyene et al, [12] Supported Cardiac Disease Detection and Evaluation can be done by Data Mining Techniques. The primary objective is to identify the incidence of cardiovascular disease within a short period of time. The method suggested is also important in many organizations of healthcare who have no more knowledge and skills It offers different clinical highlights, for example age, sex, pulse are a portion of the traits included for deciding if the individual has illness or not. Dataset analysis is computed using WEKAtools.

R. Sharmila et al, [13] Suggested usage of classification algorithm for the analysis of dynamical cardiac disease. It is advised to use bigdata resources including Distributed File System of Hadoop, Map reduction alongside expectation of coronary illness with optimized attribute selection.This research conducted an inquiry into the use of various data mining techniques to predict heart disease It recommends using HDFS to store big data in multiple regions and continuously execute.

Sharma Purushottam et al, [15] New rules c45 and partial tree prediction methods for cardiovascular risk. This research will uncover rules to estimate patients ' risk levels based on their health factor. Output can be measured in measurements of classification of performance, classification of failures, rules generated and the performance. And analysis was performed using partial tree and C4.5. The outcome reveals that prediction is potential and more efficient.

P.Sai Chandrasekhar Reddy et al, [17] Suggested cardiovascular risk prediction taking data mining method ANN. As the expense of diagnosing patients with cardiac failure grew, new technologies needs to be established which can diagnose heart attacks.After measurement, predictor method is being used to evaluate the wellbeing of the patient dependent on various factors, The system is known to be powerful in java.

S. Prabhavathi et al, [23] Suggested Neural Fuzzy System (DNFS) tree-based decision-making technique for study and prediction of various heart diseases. This report discusses work into the treatment of cardiovascular disease. DNFS stands for Decision Neural Fuzzy System based on a Decision Tree. This work is aimed at creating a smart and value-effective system, as well as improving the current system efficiency. In this report in particular, Data mining techniques are being utilised to enhance cardiovascular disease prediction. The conclusion of this study indicates that the neural networks and SVM have a positive way of predicting heart problems. Even data mining tools are ineffective in predicting heartillness.

Sairabi H.Mujawar et al, [24] K-means and naïve coves used for heart attack prediction. This work is conceived to develop the program utilizing a prior heart dataset that offers care. Thirteen qualities were considered for product design. Techniques for data processing, including clustering, Database details may be obtained from the classificationprocesses. Cleveland Heart Registry used 13 attributes for 300 records in total. This system is built to assess how often the individual has heart failure or depending on the 13 values of the attributes.

Sharan Monica.L et al, [25] The study of heart attack has been suggested. This research suggested methods for predicting the disease using data mining. It's also intended to choose the analysis of existing methods to retrieve data from the database,

and will be useful to heath care professionals. The output can be achieved depending on the time it takes for the program to construct the decision tree.

B Gnaneswar et al,[27] explored numerous simple and hybrid data mining methods, which were analysed to determine which one is best for accurately predicting heart disease. They demonstrated the merits and limitations of existing systems such as the (SVM), decision trees Nave Bayes methods, , K-Means algorithms, and Neural Networks

G. Shanmugasundaram et al,[29] looked into the many aspects and their importance in detecting cardiac disorders. They discovered that several predicted features were not taken into consideration by the researchers, and that key aspects such as drinking and smoking could have been added**.** They came to the conclusion that lowering the number of variables did not accurately predict cardiovascular diseases in order to increase accuracy. This statement might not be true if the reduced qualities have no bearing on the forecast. What matters here is the significance of the traits picked, not their quantity

3. OBSERVATION OBTAINED ABOUT CVD

Table I shows the various kind of coronary Illness and their description and Table II shows the observations were analyzed and summarized in terms of technology and Accuracy rate.

Table 1 Various kinds of Coronary Illness

Illness	Descption
Stroke	A stroke is a neurological disorder where insufficient blood circulation through the brain leading to cell death
Cardiac arrest	Cardiac arrest is a whole cessation of blood deliver springing up from the heart's incapability to successfully pump. Signs consist of loss of manage and anomalous or incomplete respiration
Arrhythmia	A Problem with the charge or temper of your pulse. It means that your coronary heart pulsates too rapidly, too gradually, or with a sporadic example. At the factor whilst the coronary heart pulsates faster than ordinary, it's far referred to as tachycardia.
High Blood Pressure	This will happen when your strain augmentations to disastrous levels. In this circulatory process strain estimation thinks about how much blood is experiencing your veins and the proportion of check the blood meets while the heart is siphoning. Confined veins increase deterrent.
Coronary artery disease	Coronary artery disorder is coronary artery narrowing or blockage, typically triggered by atherosclerosis. The accumulation of cholesterol and fatty compounds (called plaques) on lies inner walls of the arteries is atherosclerosis (sometimes named "hardening" or "clogging" of arteries).
Peripheral artery disease	It develops in the arteries that bring blood to the brain, lungs, and limbs. Plaque consists of food, glucose, magnesium, fibrous tissue, and other blood-borne compounds.
Congestive heart failure	Congestive cardiovascular breakdown (CHF) is a relentless powerful condition that impacts the siphoning force of your heart muscles. While routinely insinuated similarly as "cardiovascular breakdown," CHF expressly suggests the stage where fluid creates around the heart and makes it siphon inefficiently. We have four heart chambers.
Congenital heart disease	Congenital cardiac disorder (congenital heart defect) is one or more defects within the anatomy of the heart from which you are born.

TABLE II A STUDY OF DIFFERENT TECHNIQUES USED FOR CVD

S.NO	AUTHOR	OBJECTIVE	TECHNIQUES USED	ACCURACY/CONCLUSION
1	Noura Ajam et al	Artificial Neural Network is Used for Predicting Heart Attack.	Artificial Neural Network	By using ANN we get accuracy of 88%.
2	Sharma Purushottam et al,	Evaluation of the heart attack prediction method using C4.5 principles and a partial list.	Naive Bayes algorithm and C4.5 rules	Naïve Bayes gives less accuracy as compare to C4.5 in predicting heart attack.

3	S. Seema et al,	Data mining is being used to detect chronic diseases.	Naïve Bayes SVM	Naïve Bayes ' average precision in the case of heart attack is 91.56% Best precision reached by SVM was 95.56% for heart disease
4	Ashok al,Kumar Dwivedi et al,	Efficiency evaluation of various machine learning methods for detecting heart attack	Logistic Regression Naive Bayes KNN Classification Tree	85% 83% 80% 77%
5	K. Gomathi et al,	Data processing techniques for multi-disorder analysis.	Naive Bayes J48	Detect Breast Cancer: 82.5% Diabetes: 77.6% Detect Heart Disease: 79% Diabetes: 100% Cardiac disease: 77% Detect Breast cancer: 75.5%
6	Jayamin Patel et al,	Forecasting heart attack utilizing the computer learning and data mining methodologies.	Random Forest Algorithm	LMT gives us 55.75% of accuracy while J48 give us 56.76% accuracy which is better than LMT.
7	Ashwini Shetty A et al,	Similar approaches to data analysis for forecasting heart attack.	Neural Network	By using WEKA tool and Matlab accuracy got 88%
8	Prajakta Ghadge et al,	Intelligent program for forecasting cardiac failure using big data	Naïve bayes.	The computerization of this system makes very favorable
9	S. Prabhavathi et al,	Comparative coronary illness conclusion and expectation utilizing DNFS methods .	Naïve Bayes, Decision Tree	Accurate similar to the forms of heart disorder. Treatment of CVD= around 85% to 99%. CHD Diagnosis= from 82 to 92 %
10	Sharan Monica. L et al,	Mining Techniques prediction analysis for Cardio-Vascular	Simple CART J48 Naïve Bayes	92.2% 91.1% 88.5%
11	Jayami Patel et al,	Mining and ML techniques for predicting heart attacks	LMT, UCI	UCI, compared with LMT, gives better precision.
12	P. Sai Chandrasekhar Reddy et al,	The use of ANN and data mining algorithms to forecast cardiac cases	ANN	The accuracy has been proved in java
13	Chala Bayen et al,	A Review to predict the incidence of heart disease	SVM, Naïve Bayes, J48,	It offers short-term result that helps to deliver quality programs and lower individual costs.

14	R. Sharmila et al,	A statistical method to boost diagnosis computer techniques for cardiac attack.	Parallel Fashion Support Vector Machine	Support Vector Machine offers 82.35% better accuracy. SVM provides greater accuracy in parallel mode than sequential SVM .
15	S. Carolin Jeeva et al,	Ensemble Classification techniques to predict risk-based heart disease and also improving the accuracy	ensemble techniques	By using ensemble techniques, using ensemble classification, a cumulative improvement of 7% accuracy was obtained for poor classifiers.
16	Aakash Chauhan et al,	Determine the criteria for predicting coronary artery disease.	Rule mining algorithm	Rule Mining improves results based on prediction techniques
17	Chaitanya Suvarna et at,	Data mining and optimization approaches were used to create a prediction algorithm for cardiac disease.	particle swarm optimization technique (PSO)	Data mining methods based on particle swarms successfully forecast cardiac disease.
18	Balasaheb Tarle et al,	Propose a medical data classification model for cardiac disease	Artificial Neural Network	Classification accuracy is83 percent when cross-validation and ANNs are usedtogether.
19	Sonam Nikhar et al,	classifier of decision tree and Naive Bayes approach are discussed to predict cardiac disease	Decision Tree, Naive Bayesian Classifier	The DTC is more precise than the Naive Bayes classifier.

4. CONCLUSION AND FUTURE SCOPE

A study is guided in this article on the many ways proposed to predict cardiovascular illnesses. A total of 19 attributes were used. The goals, data source, objectives, and outcomes of each strategy were all presented. The results were discussed, and it was highlighted to different techniques of machine Learning and Data Mining were used for these approaches as well as attributes. In the future, we will conduct a more in-depth investigation of these attributes that are thought to be possible threat factors in order to compile a comprehensive list of applicable attributes that can be used to forecast cardiovascular illnesses and then explore for probable links between them. As a result of this research, we will be able to present a framework for developing a predictive model..

Reference

1. PreetiGulia, SumitaSihag.**Enhance Security in MANET using Bacterial Foraging Optimization Algorithm. International Journal of Computer Applications.**2013 December, 84(1), pp. 32-35 (International Journal of Computer Applications (0975 – 8887)) .

2. Animesh Hazra, Arkomita Mukherjee, Amit Gupta, Asmita Mukherjee, “Heart Disease Diagnosis and Prediction Using Machine Learning and Data Mining Techniques: A Review”, Research Gate Publications, July 2017, pp.2137-2159.

3. V. Krishnaiah, G. Narsimha, N. Subhash Chandra, “Heart Disease Prediction System using Data Mining Techniques and Intelligent Fuzzy Approach: A Review”,International Journal of Computer Applications, February 2016

4. Guizhou Hu, Martin M. Root, “Building Prediction Models for Coronary Heart Disease by Synthesizing Multiple Longitudinal Research Findings”, European Science of Cardiology, 10 May 2005.

5. T.Mythili, Dev Mukherji, Nikita Padaila and Abhiram Naidu, “A Heart Disease Prediction Model using SVM- Decision Trees- Logistic Regression (SDL)”, International Journal of Computer Applications, vol. 68, 16 April 2013.

6. Devangi Kotak “Data Mining Classification Techniques for Cardiovascular Disease Prediction”I nternational Journal of Engineering Research&Technology IJERT) ISSN: 2278-0181 IJERTV8IS110412

7. Nimai Chand Das Adhikari, Arpana Alka, and rajat Garg, “HPPS: Heart Problem Prediction System using Machine Learning”.

8. K. Polaraju, D. Durga Prasad, “Prediction of Heart Disease using Multiple Linear Regression Model”, International Journal of Engineering Development and Research Development, ISSN:2321-9939, 2017.

9. Marjia Sultana, Afrin Haider, “Heart Disease Prediction using WEKA tool and 10-Fold cross-validation”, The Institute of Electrical and Electronics Engineers, March 2017.

10. Dr.S.Seema Shedole, Kumari Deepika, “Predictive analytics to prevent and control chronic disease”, January 2016.

11. Ashok kumar Dwivedi, “Evaluate the performance of different machine learning techniques for prediction of heart diseaseusing ten-fold cross-validation”, Springer, 17 September 2016.

12. Megha Shahi, R. Kaur Gurm, “Heart Disease Prediction System using Data Mining Techniques”, Orient J. Computer Science Technology, vol.6 2017, pp.457- 466.

13. Mr. Chala Beyene, Prof. Pooja Kamat, “Survey on Prediction and Analysis the Occurrence of Heart Disease Using Data Mining Techniques”, International Journal of Pure and Applied Mathematics, 2018.

14. R. Sharmila, S. Chellammal, “A conceptual method to enhance the prediction of heart diseases using the data techniques”, International Journal of Computer Science and Engineering, May 2018.

15. Jayami Patel, Prof. Tejal Upadhay, Dr. Samir Patel, “Heart disease Prediction using Machine Learning and Data mining Technique”, March 2017.

16. Purushottam, Prof. (Dr.) Kanak Saxena, Richa Sharma, “Efficient Heart Disease Prediction System”, 2016, pp.962-969.

17. K.Gomathi, Dr.D.Shanmuga Priyaa, “Multi Disease Prediction using Data Mining Techniques”, International Journal of System and Software Engineering, December 2016, pp.12-14.

18. Mr.P.Sai Chandrasekhar Reddy, Mr.Puneet Palagi, S.Jaya, “Heart Disease Prediction using ANN Algorithm in Data Mining”, International Journal of Computer Science and Mobile Computing, April 2017, pp.168- 172.

19. Ashwini Shetty A, Chandra Naik, “Different Data Mining Approaches for Predicting Heart Disease”, International Journal of Innovative in Science Engineering and Technology, Vol.5, May 2016, pp.277- 281.

20. Jaymin Patel, Prof. Tejal Upadhyay, Dr.Samir Patel, “Heart Disease Prediction using Machine Learning and Data Mining Technique”, International Journal of Computer Science and Communication,vol.7, September 2015- March 2016, pp.129-137.

21. Boshra Brahmi, Mirsaeid Hosseini Shirvani, “Prediction and Diagnosis of Heart Disease by Data Mining Techniques”, Journals of Multidisciplinary Engineering Science and Technology, vol.2, 2 February 2015, pp.164- 168.

22. Noura Ajam, “Heart Disease Diagnoses using Artificial Neural Network”, The International Insitute of Science, Technology and Education, vol.5, No.4, 2015, pp.7-11.

23. Prajakta Ghadge, Vrushali Girme, Kajal Kokane, Prajakta Deshmukh, “Intelligent Heart Disease Prediction System using Big Data”, International Journal of Recent Research in Mathematics Computer Science and Information Technology, vol.2, October 2015 - March 2016, pp.73- 77.

24. S.Prabhavathi, D.M.Chitra, “Analysis and Prediction of Various Heart Diseases using DNFS Techniques”, International Journal of Innovations in Scientific and Engineering Research, vol.2, 1, January 2016, pp.1-7.

25. Sairabi H.Mujawar, P.R.Devale, “Prediction of Heart Disease using Modified K-means and by using Naïve Bayes”, International Journal of Innovative research in Computer and Communication Engineering, vol.3, October 2015, pp.10265-10273

26. Sharan Monica L, and Satheesh Kumar B (2016), “Analysis of Cardiovascular Heart Disease Prediction Using Data Mining Techniques”, International Journal of Modern Computer Science (IJMCS), ISSN: 2320-7868(Online), Vol. 4, Issue 1, February 2016, pp. 55-58.

27. C. Beulah Christalin Latha, S. Carolin Jeeva “Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques” Informatics in Medicine Unlocked 16 (2019) 100203

28. G. B et E. J. M.R, “A Review On Prediction And Diagnosis Of Heart Failure,” International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), 2017.

29. J. Maiga, G. G. Hungilo and Pranowo, "Comparison of Machine Learning Models in Prediction of Cardiovascular Disease Using Health Record Data," 2019 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS), 2019, pp. 45-48, doi: 10.1109/ICIMCIS48181.2019.8985205.

30. S. G, M. S. V, R. Saravanan et S. Balaji, “An Investigation of Heart Disease Prediction Techniques,”IEEE, vol. 10, 2018

31. M. A. P et S. P. Shaji, “Prediction and Diagnosis of Heart Disease Patients using Data Mining Technique,” International Conference on Communication and Signal Processing, 2019.

32. S. N. Induja et C. G. Raji, “Computational Methods for Predicting Chronic Disease in Healthcare Communities,”chez International Conference on Data Science and Communication (IconDSC), 2019

33. A. Gavhane, I. Pandya, G. Kokkula et P. K. Devadkar, “Prediction of Heart Disease Using Machine Learning,” chez ProInternational conference on Electronics, Communication and Aerospace Technology, 2018.

34. A. Rairikar, V. Kulkarni, V. Sabale, H. Kale et A. Lamgunde, “HEART DISEASE PREDICTION USING DATA MINING TECHNIQUES,” chez International Conference on Intelligent Computing and Control, 2017.

35.

Mukesh Pandey is an Assistant Professor in CSE Department at Uttaranchal University, Dehradun. He is SimpliLearn Certified Trainer in Machine Learning. He holds B.Tech inInformation Technology and an M.Tech in Computer Science.He is also doing his Ph.D in ComputerScience from UTU, Dehradun. He has more than 9.8-year Experience in Government as well asPrivate Engineering Institutes in Academics’. He has a good number of Research Papers in ReputedJournals and Conferences.

OVERVIEW OF VARIOUS STUDIES ON TECHNICAL ASPECTS AND SIGNIFICANCE OF A STIRLING ENGINE INCORPORATED CSP SYSTEM

Kota Krishna Koundinya[1], Karthik Pavan Kumar Gunturu[2], Dr. Madhu Sharma[3]

[1,2] Masters of Technology, School of Engineering, University of Petroleum and Energy Studies
[3] Sr. Associate Professor, School of Engineering, University of Petroleum and Energy Studies

Abstract

Solar energy has been playing a crucial role as a clean energy source for power generation without any carbon emissions and many countries have been diversified and adopted the technologies such as solar PV and solar thermal technology, while these methods are still having limitations such as lower efficiencies, challenges in the thermal energy storage, requirement of larger areas. Stirling engine is a type of heat engine that uses an external combustion mechanism for the operation, while the thermal energy from the external heat source is converted into mechanical work and then to electricity, whereas the working fluid is used in this type of heat engine will be chemically inert. This paper overviewed various beneficial aspects of a Stirling engine such as this heat engine is eco-friendly, quiet, efficient and the most noticeable advantage of Stirling engine is the heat absorbed by the working fluid can also be stored by various phase changing materials using regenerators. This paper describes the studies on off-grid power generation for micro-scale applications, design aspects of the CSPP-Stirling engine system, and also various studies on the working parameters of the system.

Key Words: Carbon Emissions, Power generation, Regenerators, Solar Thermal Energy, Stirling Engine, and Thermal Energy Storage.

1. Introduction

Power generation in recent times has become an important per capita energy efficiency development factor globally as all the countries are diversifying to utilize renewable energy which has been playing a crucial role in the sustainable development of the power sector as the sources of this energy kind are abundant, while some renewable sources like solar, wind, hydro and biomass can be harnessed with reduced emissions of greenhouse gases which could be caused due to the utilization of fossil fuels for energy generation [1]. The hybridization of renewable sources can improve the energy efficiency by overcoming the limitations of power generation from standalone systems and also provides a wide scope in installing off-grid hybrid renewable energy systems based on the resources available especially in rural areas because of various factors such as area, climatic conditions, population, hard terrains and revenues [2]. Thus, the integration of various power generation systems with renewable energy can reduce the gap between energy demand and supply in rural areas. Moreover, solar energy is one of the effective alternatives that has been replacing the traditional and conventional power generation using fossil fuels such as coal, oil, diesel, and other gases like LPG, CNG, etc., for power generation. This replacement also protects the environment and ecological conditions [3]. Many advancements have been made in solar power plants for concentrating and capturing solar energy to achieve an appropriate output efficiency, while various equipment such as concentrators (LFRs) and collectors (HFCs, CPCs, PDCs, and PTCs) is used to fulfill the objective of concentrating solar radiation [4-9]. Adoption of a cost-effective technology is a barrier as cost per square meter accounts for the total cost of installation, while the concentration ratio should be appreciable to attain a good range of temperatures to ensure optimum efficiency of the respective system. The Concentration of solar energy can be done at much lower costs when compared to the electricity-generating using solar PV systems [10]. Usually, LFRs and PTCs are used for medium-temperature applications ($\leq\approx$500°C), while PDCs and HFCs are used for higher temperature applications (>500°C) [11]. Another barrier that has been affecting the efficiency of a CSPP is energy storage. These concentrators are arranged in a manner by which the sun rays are converged and focused by mirrors onto a certain point at which various power generation systems such

as water boilers, Stirling motors, and gas turbines are installed. The power generated should be consistent and regulated because renewable energy sources are uncontrollable and unpredictable as they completely depend on climatic conditions such as humidity, solar radiation, wind speed, temperature, etc. There are many methods to utilize and also store the solar thermal energy and one of the most effective methods is the integration of solar thermal energy with a Stirling engine as this external combustion system has an appreciable range of efficiencies and also this heat engine could be more eco-friendly as any heat medium can be used to drive this Stirling engine, while this also reduces the dependency on the fossil fuels and the power generation from this system can be done without any carbon emissions. The basic representation of power generation using a Solar-dish Stirling engine is shown in figure 1.

Moreover, an energy storage system in a power generation plant is recommended as it stores the excess energy available from the renewable resources in the form of thermal energy and electricity, such that the stored energy will be utilized whenever it is required to fulfill the energy demand. Various energy storage methods such as organic Rankine cycles, Stirling cycles, Brayton cycle, etc., are implemented to recover the excess energy in many combined heat and power applications, while Stirling engines can be significantly used for energy storage and medium temperature applications, these type of heat engines can be integrated with CSPPs for energy storage and power generation which would be a viable alternative to the other thermodynamic cycles as any heat source can operate this kind of heat engine. The Stirling engine incorporation with a CSPP as it can overcome the environmental challenges such as carbon emissions, global warming, habitat loss, and can also protect the eco-system [12].

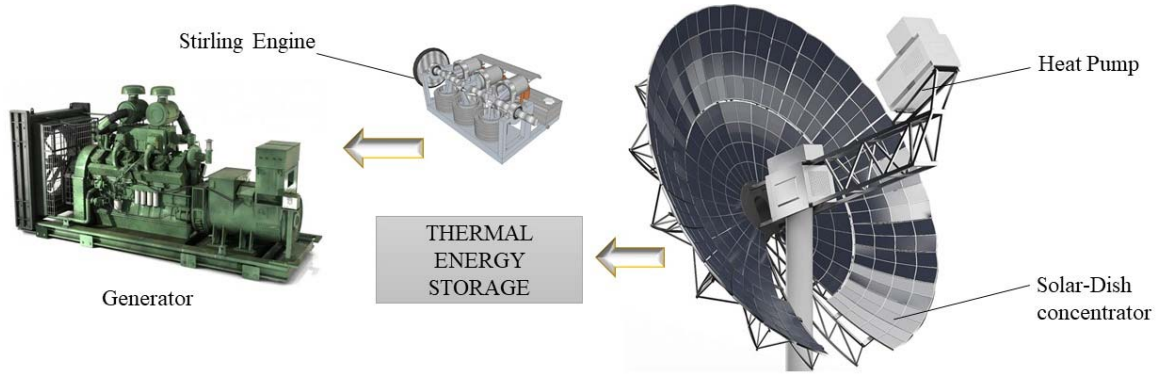


Figure 1: A Solar-dish Stirling engine incorporated with a Thermal Energy Storage

2. Working of a Stirling Engine

Stirling engine is an external combustion engine that was invented by Dr. Robert Stirling in 1816. These Stirling engines have a unique operating cycle that exhibits a closed regenerative thermodynamic process known as “Stirling Cycle” which includes the combination of various cyclic expansion and compression processes such as isothermal compression in stage 1, isochoric heating in stage 2, isothermal expansion in stage 3 and isochoric cooling in stage 4 as shown in figure 1 and the working of the displacer is shown in table 1 [13, 14]. PV and TS representation of the Stirling cycle along with the process flow sheet describing the principle have been shown in figure 2. This regenerative thermodynamic cycle has an appreciable efficient of about 40% which is greater than both otto and diesel engines which have their respective efficiencies of about 25% and 35 % [15]. The Stirling engine works on the temperature difference of the working fluid which can be achieved by any thermal energy source such as concentrated solar energy, biomass combustion energy, etc., that can heat the working fluid which operates the engine, whereas various gases such as hydrogen, helium, air, and nitrogen can be used as working fluids and thus the system will be safer to utilize as the flow by the continuous volume changes in the working fluid and temperature raise at the cold sie is regulated by the installation of various extended surface systems for cooling the engine.

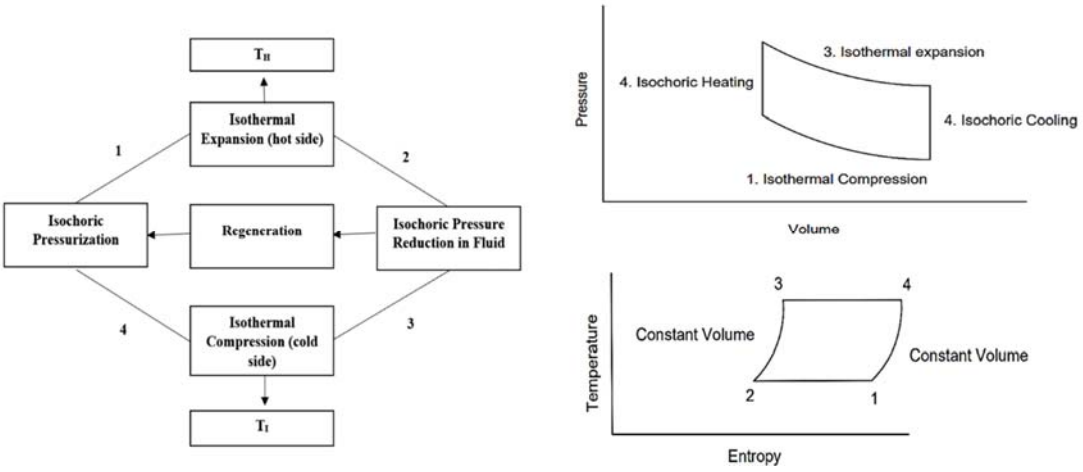
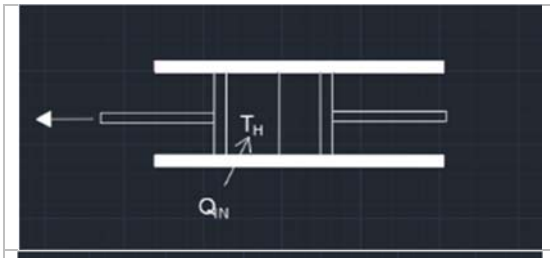
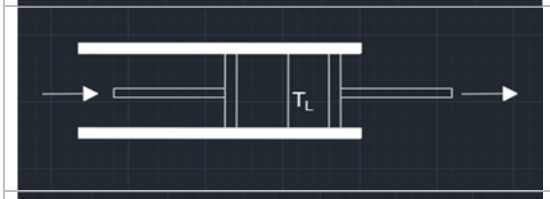
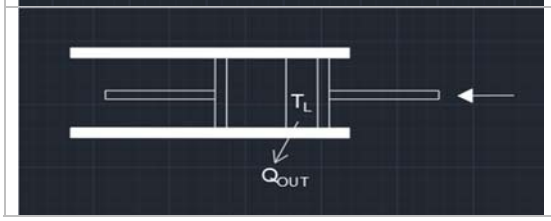


Figure 2: Flowsheet of a Stirling Cycle

Table 1: Summary on the operation of a Stirling engine

	**Step I:** 1. Isochoric or Isometric. 2. Increase in enthalpy of the working fluid. 3. Follows the ideal gas law. 4. Pressure of the working fluid increases.
	**Step II:** 1. Isothermal expansion step 2. Work is done due to the expansion of fluid. 3. Heat transfer takes place between the working fluid and external heat source at a constant temperature.
	**Step III:** 1. Isometric process 2. Enthalpy of the working fluid decreases. 3. Displacement of the working fluid to the cold side from the hot side.
	**Step IV:** **1.** Isothermal contraction of the working fluid. **2.** Work is done due to the compression of fluid.

Considering the environmental aspects and energy utilization to meet the demand, the Stirling engine can be able to deliver an appropriate output with high heat efficiencies, and silent operation with a wide selectivity of operational methods which make it more advantageous in giving an effective low range power outputs by converting solar thermal energy into electricity [12]. However, the power generation from a Stirling engine incorporated CSPP can be a consistent, reliable, and efficient as this combined heat and power system can be suitable for the off-grid power generation in a range of 1-50 kW in residential scale along with building energy demand achieving heat loads [16, 17].

3. Technical Approach to the Solar-Stirling Engine

Though the Stirling engines are the most effective devices in converting solar thermal energy into low range power outputs, still they have some disadvantages as their performance depends on their geometrical and physical characteristics which require large volume, while the other disadvantages such as the probability of working fluid leakage, low compression ratios, slow process and thermal limitations in operating conditions of the heater and cooler temperatures became the barriers of engine's performance [18, 19]. Thus, several methods have been adopted for enhancing the performance of a Stirling engine as it has more advantages when compared to ICEs.

Houda H et al. [20], elaborated the technological challenges and various factors affecting the performance of a Stirling engine which included geometric parameters like compression ratio, phase angle between power piston and displacer, dead volume, swept volume, engine speed, oscillations, and the characteristics of regenerators such as the material of construction, porosity, design, lubrication, and temperature distribution. However, for micro/small-scale power generation purposes, tubular heat exchangers are preferred as heaters and coolers, while for low-temperature applications extended surface heat exchangers or fins are sufficient. A. Asnaghi et al. [21] performed thermodynamic analysis on the performance of solar Stirling engines under ideal conditions which included the regenerator effectiveness for hot and cold sides by determining the effective temperature of the regenerator using various approaches such as arithmetic mean approach, logarithmic mean approach, and half hot space-half cold space approach, while considering piston's motion to be simple harmonic, volumes of the of pistons are determined and by assuming the working fluid to be an ideal gas, the pressure equations and mass of the working fluid is determined. Whereas the thermal efficiency of the entire system is determined from the work diagram of the Stirling cycle.

Du Marchie [22] summarized the calculation of various parameters and characteristics of a solar concentrator system such as the area required for the mirrors for utility-scale solar energy, the efficiency of the solar thermal power plant, and determined the other parameters such as path length of solar radiation through atmosphere, zenith angle, average horizontal irradiance on the mirror field affecting the absorbing factor for four different days. This system was economically integrated with the desalination of water and the cost analysis was done by comparing the results of output power with utility-scale solar PV plants. Ana Cristina Ferreira et al. [23], assessed the performance of the Stirling engine by comparing solar energy and biomass by modelling the CSPP and biomass plant, while the solar dish diameter was taken as 8m, and 45° rim angle was assumed at which the heat flux has been increased as the solar radiation will be focused to a smaller receiving area because at 45° angle the focal distance and aperture ratio are lower [24-25]. A two-step numerical model has been developed for an alpha Stirling engine small scale power generation where solar energy and biomass as external sources.

Analysis of the economy of Stirling engine, pollution generation, and efficiency is very important as the Stirling engines are quiet and pollution less with good efficiency of a minimum of 80% and the cost of energy generation will be cut off for sure as long as there is sunshine. The materials and manufacturing costs of the Stirling engine as its efficiency increases with the increase in temperature and so the materials should withstand the higher temperatures. The exploitation of the technology like Stirling engine will reduce the environmental impacts and dependency on the fossil fuels.

4. Conclusion:

This paper reviewed various technical, environmental and economic aspects of a Stirling engine which can be employed for various applications in power generation and co-generation technologies. The versatility of this heat engine makes it a unique model for both commercial and micro-scale purposes, as it is highly efficient, quiet, safe and simple in the operation perspective. The design must be done by considering the leakage probabilities of the working fluid, and effect of lubrication on the working fluid. Even though, there are some limitations such as high costs, low reliability, and low production rates which became the major barriers for Stirling engines to succeed commercially in market. The PCMs used for thermal energy storage also makes the entire system expensive. But still, globally there are

many research studies are being done to improve the characteristics of this heat engine as it has a large scope in the energy sector, because besides its limitations Stirling engine when integrated with CSPP can reduce the CO_2 emissions of nearly 60% and the performance can be improved in co-generative operation of this system especially in winters. The Solar-Stirling engine systems can be the better option for off-grid power generation. Lastly, the Stirling engine has a potential to be competitive enough with the other renewable technologies with some necessary technical modifications to make this heat engine a feasible and effective option for power generation in the commercial space such that the penetration of this technology into market would improve the energy efficiency and also quality of the environment by decarbonization. However, still further research is needed to overcome the challenges and making the Stirling engine commercially viable and cost effective.

5. REFERENCES

[1] Aditya A et al. 2018, Design and development of Solar Stirling Engine for power generation, IOP Conf. Series: Materials Science and Engineering 376 (2

[2] 018) 012022 doi:10.1088/1757-899X/376/1/012022.

[3] Vendoti Suresh et al. 2020, Modelling and optimization of an off-grid hybrid renewable energy system for electrification in rural areas. Energy Reports, Volume 6, 2020, Pages 594-604, ISSN 2352-4847, https://doi.org/10.1016/j.egyr.2020.01.013.

[4] Abdelaziz E.A., Saidur R., Mekhilef S., A review on energy savings strategies in industrial sector, Renawable and Sustainable Energy Reviews 15 (2011) 150-168.

[5] Winter C. J., Sizmann R.L., Vant-Hull L.L., Solar Power Plants, ed. Springer-Verlag, 1991.

[6] Barlev D., Vidu R., Stroeve P., Innovation in concentrated solar power, Solar Energy Materials & Solar Cells 95 (2011) 2703-2725

[7] Zhang H.L., Baejens J., Degreve J., Caceres G., Concentrated Solar Power Plants: Review and design methodology, Renewable and Sustainable Energy Reviews 22 (2013) 466-481

[8] Siva Reddy V., Kaushik S.C., Ranjan K. R., Tyagi S.K., State-of-the- art of solar thermal power plants – A review, Renewable and Sustainable Energy Reviews 27 (2013) 258-273

[9] Morin G., Dersch J., Platzer W., Eck M., Haberle A., Comparison of linear fresnel and parabolic trough collector power plants, Solar Energy 86 (2012) 1-12

[10] Heliofocus web site,http: //www.heliofocus.com, last access 16/06/2015.

[11] IEA 2004, World Energy Outlook 2004, International Energy Agency (IEA), Paris, page 233.

[12] A Giovannelli 2015, State of the Art on Small-Scale Concentrated Solar Power Plants, State of the Art on Small-Scale Concentrated Solar Power Plants (core.ac.uk)

[13] Union of Concerned Scientists, 2013. Environmental Impacts of Renewable Energy Technologies, available at https://www.ucsusa.org/resources/environmental-impacts-renewable-energy-technologies.

[14] Mike Miao He, 2018, Stirling Engine for Solar Thermal Electric Generation, Technical Report No. UCB/EECS-2018-15 http://www2.eecs.berkeley.edu/Pubs/TechRpts/2018/EECS-2018-15.html

[15] Hongling Chen et al. 2014, Design of a Stiring Engine for Electricity Generation.

[16] Ali et al. 2016, Investigation of the effect of design parameters on power output and thermal efficiency of a Stirling engine by thermodynamic analysis, International Journal of Low-Carbon Technologies, Volume 11, Issue 2, May 2016, Pages 141-156, https://doi.org/10.1093/ijlct/ctu030

[17] D. G. Thombare and S. K. Verma, "Technological development in the Stirling cycle engines," Renewable and Sustainable Energy Reviews, vol. 12, no. 1, pp. 1–38, 2008

[18] K. Alanne, N. Soderholm, K. Sir € en, I. Beausoleil-Morrison, Techno-economic assessment and optimization of Stirling engine micro-cogeneration systems in residential buildings, Energy Convers. Manag. 51 (2010) 2635e2646, https://doi.org/10.1016/j.enconman.2010.05.029.

[19] M. Badami, F. Camillieri, A. Portoraro, E. Vigliani, Energetic and economic assessment of cogeneration plants: a comparative design and experimental condition study, Energy 71 (2014) 255e262, https://doi.org/10.1016/ j. energy.2014.04.063.

[20] M.H. Ahmadi, M.A. Ahmadi, A. Mellit, F. Pourfayaz, M. Feidt, Thermodynamic analysis and multi objective optimization of performance of solar dish Stirling engine by the centrality of entransy and entropy generation, Int. J. Electr. Power Energy Syst. 78 (2016) 88e95, https://doi.org/10.1016/ j.ijepes.2015.11.042.

[21] Houda H et al. 2018, Technological Challenges and optimization efforts of the Stirling machine: A review; available at https://doi.org/10.1016/j.enconman.2018.06.042

[22] A. Asnaghi et al. 2012, Thermodynamics performance analysis of solar Stirling engines, International Scholarly Research Network ISRN Renewable Energy Volume 2012, Article ID 321923, 14 pages doi:10.5402/2012/321923

[23] Evert Du Marchie Van V, 2007, Large-scale concentrating solar power (CSP) technology, DOI:10.1007/1-4020-4604-9_3. In book: Macro-Engineering (pp.31-44)

[24] Ana C Ferreira et al. 2020, Assessment of the Stirling engine performance comparing two renewable energy sources: Solar energy and biomass. https://doi.org/10.1016/j.renene.2020.03.020 0960-1481/© 2020 Elsevier Ltd. All rights reserved.

[25] D.J. Shendage, S.B. Kedare, S.L. Bapat, Numerical investigations on the DisheStirling engine system, Int. J. Ambient Energy 40 (2019) 274e284, https://doi.org/10.1080/01430750.2017.1388840.

[26] R. Beltran, N. Velazquez, A.C. Espericueta, D. Sauceda, G. Perez, Mathematical model for the study and design of a solar dish collector with cavity receiver for its application in Stirling engines, J. Mech. Sci. Technol. 26 (2012) 3311e3321, https://doi.org/10.1007/s12206-012-0801-0

Biographies

Kota Krishna Koundinya received the bachelor's degree in Chemical Engineering from Maharaj Vijayaram Gajapathi Raj College of Engineering in 2020, pursuing the master's degree in Renewable Energy Engineering from the department of Electrical and Electronics, University of Petroleum and Energy Studies. The research areas include biofuels, sustainable energy development, waste to energy and renewable energy.

Gunturu Karthik Pavan Kumar received the bachelor's degree in Mechanical Engineering from Koneru Lakshmayya Educational Foundation in 2020, pursuing the master's degree in Energy Systems from the department of Electrical and Electronics, University of Petroleum and Energy Studies,. The research areas include, automobile development, sustainable energy, and energy recovery.

Dr. Madhu Sharma received the bachelor's degree in Electrical Engineering from Rani Durgawati Vishwavidyalaya, Jabalpur, master's degrees in Power System from J.R.N. RV University, Udaipur and Energy Management from Devi Ahilya University, Indore and Doctorate in Renewable Energy from University of Petroleum and Energy Studies. Dr. Madhu Sharma PhD in renewable Energy and a Certified Energy Auditor has over 25 years of rich experience in Renewable Energy, Power Management, Energy Auditing, Energy Management and Academics. She has 34 publication in various international and national journals and 7 book chapters. She has conducted various Energy Audits of industries, buildings etc, organized many conferences, Workshops, Training programs & Seminars, Presently working as Sr. Associate Professor at UPES

A Novel Region-Based Image Segmentation Method Using GLCM Algorithm

Mukesh Kumar[1], Pooja Joshi[2], Ravi Dhaundiyal [3], Arzoo Chamoli[4],

Gaurav Singh Negi[5], Siddharth Shankar Bhatt[6], Michael G. Albino[7]

[1, 2,3,5,6] *Assistant Professor, UIT, Uttaranchal University, Dehradun, India*

[6]*College of Communication and Information Technology President Ramon Magsaysay State University, Philippines*

E-mail[1]: mukeshkumar.cse@uttaranchaluniversity.ac.in,

E-mail[2]: poojajoshi@uttaranchaluniversity.ac.in, E-mail[3]: ravi.kec@gmail.com,

E-mail[4]: arzoo.chamoli@gmail.com, E-mail5: gauravnegi@uttaranchaluniversity.ac.in

E-mail[6]: siddharthbhatt@uttaranchaluniversity.ac.in

Abstract.

An image is a medium to transfer information. An image includes lots of valuable information. In the technological domain of digital image, a very important task is to understand the image and extract information from the image for accomplishing some operations. Image segmentation is the initial step towards gaining the understanding regarding the image. The technique of GLCM is applied with the region based segmentation. The proposed technique is implemented in MATLAB and results are analysed in terms of certain parameters which shows high performance.

Keywords. Image segmentation, Region Segmentation, GLCM

1. INTRODUCTION

An image is a medium totransfer information. An image includes lots of valuable information. In the technological domain of digital image, a very important task is to understand the image and extract information from the image for accomplishing some operations. Image segmentation is the initial step towards gaining theunderstanding regarding the image. In fact, People generally not interested in all segments of the image. They just show interest for some certain regions withthe similar features. Image segmentation is a major hotspot in image processing and computer vision. It also provides a significant base for the image recognition. Image segmentationdepends on certain criteria for dividing an input image into a number of the similar nature of the category so that the areas appealing to people can be extracted. Image segmentation contributes significantly for image analysis and generates understanding regarding image feature extraction and recognition [1].

There are mainly two objectives of image segmentation. The first one is to decompose the image into segments for more analysis. In simple case, the environment might be reasonably well controlled with the purpose to the make the extraction possible of just those regions by segmentation process that requires further analysis. In complicated conditions thatinclude the extraction ofanoverall road network from a greyscale aerial image, it may be quite difficult to implement the segmentation process. Therefore, the application of a great deal of areafor building knowledge might be required in these conditions. Performing a change of representation is the second purpose of image

segmentation. It is essential to arrange the pixels of the image into higher-level units. These units should be either more expressive or more competent or both for advance analysis.It is imperative and at the same time complex to implement segmentation in the domain of image processing. In this process, the partitioning of an image is carried out into a number ofexpressivesegments of alike attributes. Image segmentation aims to simplify an image. This means this process presents a picture in such a way that it can be understood easily.There are many commonly used images segmentation algorithms. Some of these algorithms have been described below:

a. Threshold Method:One of the most fundamental approaches of image segmentation is threshold.In this approach, the division of data processing about grayscale image is carried out on the basis of the gray level of dissimilarobjects in direct manner.Threshold segmentation is a very popular segmentation technique in region-based segmentation algorithms. The key idea of this approach is to robotically determine the optimal threshold on the basis of a definitestandard, and usage of these pixels in accordance with the graylevel for cluster analysis. The picture that have less contrast as compared to the backdrop get treated by this approach. These approaches can be selected either manually or automatically on the basis of predefined knowledge of the picture attributes [2].

i) Local Threshold: In the local threshold approach, manifold segmentation thresholds should be selected. Once the selection takes place, this approach makes use of many thresholds for dividing the picture into many object areas and backdrops.

ii) Global Threshold: The global threshold approach makes use of a single threshold for dividing the picture into two parts comprising target and the backdrop.

The largest interclass variance method or Otsu is one of the most frequently used threshold segmentation algorithmic approach.In this algorithm, an optimal threshold is selected globally by maximizing the variance among classes. Moreover, some other well-known threshold-based approaches that are used in the integration with other techniques [3]. The threshold-based segmentation offers a major benefit of the non-complex computation and faster operational speed. Especially, it is possible to obtain the segmentation effect when both the target as well the background has high contrast. The coreweakness of this approach is that the accurate results for image segmentation problems cannot be detected easily in the absence of a noteworthy gray scale variancewithin the image. This segmentation approach showssensitivity towards noise and grayscale inconsistency as it just pays attention to the gray knowledge of the picture without taking into account the spatial image knowledge. This is the reason that this approach is generally used in combination with other segmentation approaches.

b.Region Based Segmentation Method:Region based segmentation focuses on the direct determination of the region. The keymotive of this approach is to do the classification of a particular picture into multiple regions or segments. There, it is required to determine or evaluate that a pixel within an image relates to which class [4]. At present, a number of approaches are available to perform region-based segmentation. The segmentation algorithms based on region operate in iterative manner by grouping neighbouring pixels of similar values in the same group and the pixels with different values in another group. The two popular categories of these approaches are region growing and split and merge. Both of these approaches have been described below:

i) Region growing methods: These segmentation techniques come in the category ofarchetypalsuccessive region segmentation. In these methods, the pixels with have similar features are combined together for creating a region. These methods initially make selectionof a seed pixel. After that, the similar pixels in the vicinity of the seed pixel are merged together into the locative region of the seed pixel. The key benefit of this approach isthat it generallydoes the separation of the linked areas with the alikeproperties. In this way, these methods generate satisfactory outcomes and give high quality knowledge of the image edge.

ii) Region splitting and merging methods: Split and merge is another popular approach of region-based image segmentation. The key idea of this approach is taken from quad tree data representation. In such representation, the partitioning of a square image is carried out into four quadrants in case the real image segment is uneven in terms of feature. The segmentation methods based on split and merge tries to partition a picture into even regions.

2. RELATED WORK

Yubing Li, et.al (2018) suggested an advanced algorithm which was planned on the basis of the grab cut. The efficiency and accuracy of this algorithm had demonstrated in the results of experiments. However, the time cost was increased through this algorithm [5]. The research on the association among the pixels fixed the imperfection of grab cut algorithm which was the importance of this paper. One of the difficult points for the image segmentation was the complex background. The grab cut algorithm had helped enhancing the performance in complex background using this recommended algorithm. The outcomes of experiments indicated that this algorithm had some imperfection and thus it required further enhancement in future research. In particular, the result of segmentation attained was poor if it had colourful object and the complex background.

Ozan Oktay,et.al (2018) recommended a comprehensive training approach in which a previous knowledge was integrated into CNNs using a new model that was subsequently trained successfully [6]. The novel approach persuades models that the worldwide anatomical features of the underlying anatomy were followed by them through the non-linear representations that were learnt for different shapes. The adaptation of the recommended approach was easy for various analysis tasks and it proved in this paper. The predictive accuracy of the modern models was enhanced using this approach. The multi-modal cardiac datasets and public benchmarks were employed to represent the capability of this approach. Furthermore, it was indicated that the interpretation and utilization of deep models of 3 dimensional shapes was done as biomarkers for classification of cardiac pathologies.

Haigang Sui, et.al (2017) presented a new stable shape image registration technique which was based on feature after matching the stable region with a set of rotations, scale invariant features. The matching areas were achieved using multi-scale image segmentation [7]. First of all, the images were converted into image objects through convexity model restriction and multi-scale segmentation in this algorithm. Afterward, the utilization of these reliableand steady image regions had completed as matching units instead of points and/or lines. It was indicated in the experiments that the presented algorithm was insensitive to rotations and/or resolutions distortionwhich causes the image registration accomplishment in automatic manner.

Shih-Hsiung Lee, et.al (2017) intended an image pre-processing technique for the segmentation of various parts of nail: lunula and nail plate [8]. The data representation of lunula was not clear due to poor image quality. The microscope was carried out for capturing the nail image in this paper so as the nail image quality was maintained. The images that had capture through microscope were represented the various details of the nails instead of lunula and nail plate. Demonstration of more details and the maximization of meaning related to the complexity of image processing had done in this paper. The lunula was separated from the nail plate by executing a series of image pre-processing stages at the present step. It had prominent impact was proved in the results of experiments. In the future, it would adaptable in various biometric applications.

Nishant Jain, et.al (2017) suggested a new and a different technique for automatic segmentation of diaphragm from ultrasound images. The difference in echogenicity of the diaphragm is depending on two facts included neighbouring tissues and the location of the

diaphragm in the ultrasound images. These facts were carried out in the suggested technique while segmenting the Diaphragm [9]. Various threshold values assigned with maximum and minimum intensity values available in the ultrasound image had employed in this technique for generating a set of binary images. Difference between any two threshold inputs was presented in the multiples of predetermined constant, α. The detection of all possible diaphragm which were same to curved objects had performed from each binary image at first. To end with, the exact position of the diaphragm was achieved automatically on the basis of positions of the detected objects. Fattyliver ultrasoundimages of healthy volunteers were employed for the testing of suggested technique and it was proved that this technique hadpotential for the segmentation of diaphragm from ultrasound images and obtained high precision without any help of user input.

Annegreet Van Opbroek, et.al (2019) advocated a new image weighting techniquewhich minimize the MMD between training and testing data which further assisted in the optimization of image weighting and kernel [10]. The results of experiments on hippocampus segmentation demonstrated that the performance on heterogeneous data was enhanced using the recommended methods. There was equality in the performance of earlier presented image weighting schemes and the recommended MMD. The small additional enhancement in performance had obtained when the image weighting was integrated with kernel learning optimizing whether in individual or joint way.

Tiechao Jiang, et.al (2018) focused on developing the vascular calcification segmentation process that assisted to obtain the image segmentation pre-treatment, to complete the CV model level set related to segmentation algorithm. The mathematical morphometric was implemented for processing the segmentation of an image [11]. At last, it was indicated in the outcomes of experiments that the vascular calcification image was segmented efficiently comparatively to conventional technique and the size and location of vascular calcification was achieved with accuracy through segmentation of CV model level set. The clear and smooth calcification area edge had achieved in the segmentation. In this way, an effectual basis for clinical diagnosis had obtained. However this technique was inadequate when there was a great computation. Therefore, the next work would be focused on the reduction of computational complexity. Further experiment can be done through the fusion technique [16] [17] [18] in image segmentation.

3. METHODOLOGY

This research is related to image segmentation and techniques for this purpose are generallygrouped into region based and threshold-based techniques. The region-based segmentation techniques are those which can segment the image based on the textual features of the images. The images have various type of textural features like energy, entropy etc. The threshold-based segmentation techniques are those which can segment the image into two segments. The pixels which have value above the threshold will be segmented into one segment and other into the second. Here, the region and threshold-based segmentation will be merged together to form hybrid image segmentation technique. In the proposed methodology the threshold-based segmentation techniques will be applied which can form two segments of the original image. The one segment will be proposed further and region-basedapproach will be implemented for partitioning image into N number of regions based on the pixel similarity. The phases of research work are given below:

1. Pre-processing: -The cleaning of input picture will be carried out in this stepby means of a de-noising algorithm.Also, the obtained picture will undergo more processing for the local segmentation.

2. HybridSegmentation: -This stage applies the approach of local segmentation for the image into two segments. These segments are called foreground and background. Further, the RGB scanned picture will be segmented into same and different parts by implanting the region-based segmentation. This work applied south's segmentation algorithm. A large number of benefits are provided by the segmented images obtained by this process. This work makes use of a gray scale image for threshold. The conversion of RGB image is carried out into binary image. The image achieved by this process is in B&W format. In the next step of feature extraction, the features are extracted by implementing GLCM. This algorithm is used in this work just for the extraction of texture attributes. This algorithm takes out total thirteen image attributes so that the tumor can be detected.

- Energy: This parameter computes how many times the pixel pairs are repeated. It computes the evenness of a picture. In case of extremely tiny pixels, the energy level will be big.

  Energy = $\sqrt{\sum_{i,j=0}^{N-1} p_{i,j}^2}$ (1)

- Entropy: Entropy measures the average information content in every source sign.

  Entropy=$\sum_i p_i log_x i$ (2)

- Contrast: This parameter measures the brightness of a pixel and its neighbour in the picture. The difference in the colour and intensity of the object and other objects w.r.t to the similar field is measured in the realistic manner.

  Contrast=$\frac{I_{max}-I_{min}}{I_{max}+I_{min}}$ (3)

K-Mean is a popular region-based segmentation approach. The k-mean segmentation technique will segment input image into certain classes based on the GLCM factors.

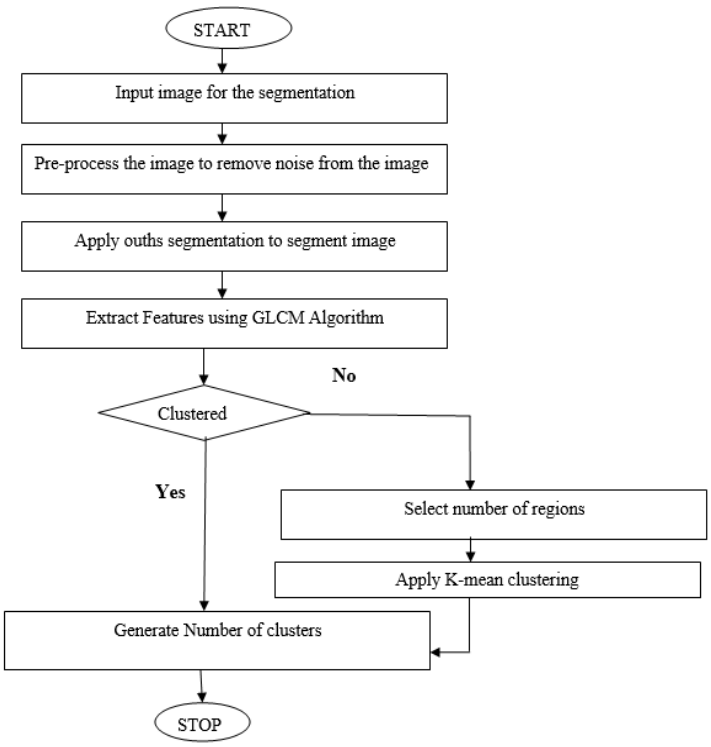


Figure 1. Experimental Flowchart

4. RESULTS AND DISCUSSION

The American Psychopathological Society demonstrates several thousand of technically peer-reviewed pictures presenting infection indications, pests, and other anarchies related to the crops and plants by its books and journal-issuing plans. APS images are beneficial for education and training purposes. Over these many years, APS figure has been available via APS PRESS slide suites, record tools, picture CDs, and these days by the online APS picture record. The pictures involving subtitles during their availability together with supplier/photographer identification and credentials are also discoverable. The APS picture folder is assisted through the APS PRESS Editorial Board and is possible to stiff into tens of thousands of infection, nuisance, and chaos pictures obtainable on Net.

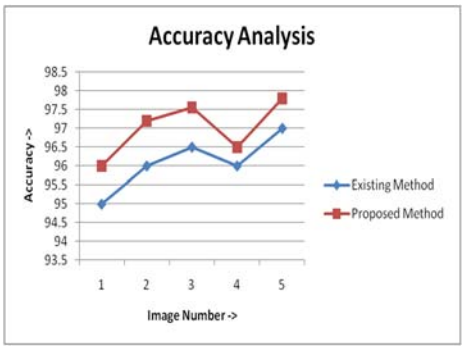


Figure 2. Accuracy Analysis

Figure 3. Execution Analysis

Figure 2 shows the accuracy-based comparison amid the existing and proposed technique. The accuracy of new technique for the segmentation is analysed in this figure. Figure 3 shows execution time-based comparison amid the new and earlier technique. The execution time of new technique is low than the earlier technique.

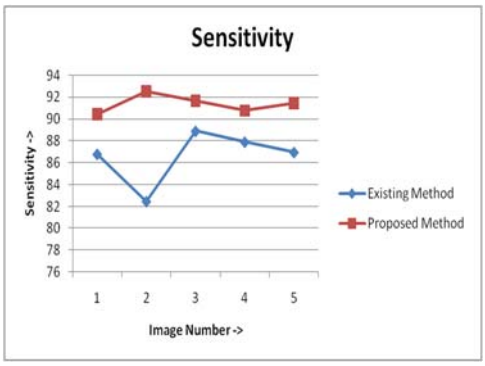


Figure 4. Sensitivity Analysis

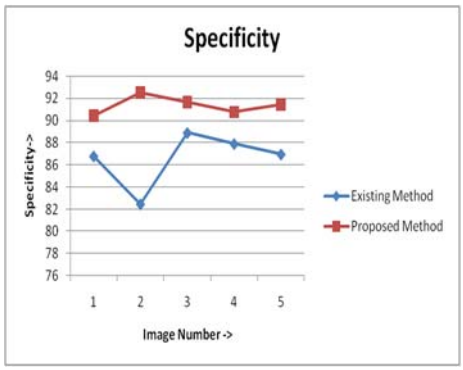


Figure 5. Specificity Analysis

Figure 4 shows the sensitivity based comparison amid the new and earlier technique. The sensitivity of proposed technique for the segmentation is analysed in this figure. Figure 5 shows the specificity-based comparison amid the existing and proposed technique. The specificity of new technique for the fruit quality prediction is analysed in this figure.

5. CONCLUSION

The technique used for the improvement of unprocessed or raw pictures is called imageprocessing. The image capturing tools installed on different objects can click images

for multiple applications. These images are mostly rough in nature. Over the time, researchers have devised various image processing algorithms.It is required to convert an image clicked by an image capturing tool into a suitable format prior to carry out any sort of image processing. Image segmentation comes at the second place in image processing. This work implements k-mean segmentation for segmenting the pre-processed image. The region-based segmentation will segment the data dependent on the taken-out features using GLCM algorithm. The performance of existing and proposed algorithms is analysed with regard to several metrics. The analytic outcomes reveal that new algorithm performs superior than the earlier algorithmic approach

6. REFERENCES

[1] A. P. James and B. V. Dasarathy, "Medical image fusion: A survey of the state of the art", Information Fusion, vol. 19, no. Supplement C, pp. 4-19, 2014.

[2] Y. Yang, "Multimodal Medical Image Fusion through a New DWT Based Technique", in 2010 4th International Conference on Bioinformatics and Biomedical Engineering, 2010, pp. 1-4.

[3] X. Zhu, H.-I. Suk, S. W. Lee, D. Shen and I., "Subspace Regularized Sparse MultiTask Learning for Multi-Class Neurodegenerative Disease Identification", IEEE transactions on bio-medical engineering, vol. 63, no. 3, pp. 607-618, 2016.

[4] S. Klein, U. A. van der Heide, I. M. Lips, M. van Vulpen, M. Staring, and J. P. W. Pluim, "Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information", Medical Physics, vol. 35, no. 4, pp. 1407-1417, 2008.

[5] Yubing Li, Jinbo Zhang, Peng Gao, Liangcheng Jiang, Ming Chen, "Grab Cut Image Segmentation Based on Image Region", 2018, IEEE 3rd International Conference on Image, Vision and Computing (ICIVC).

[6] Ozan Oktay, Enzo Ferrante, Konstantinos Kamnitsas, Mattias Heinrich, Wenjia Bai, Jose Caballero, Stuart A. Cook, Antonio de Marvao, Timothy Dawes, Declan P. O'Regan, Bernhard Kainz, Ben Glocker, Daniel Rueckert, "Anatomically Constrained Neural Networks (ACNNs): Application to Cardiac Image Enhancement and Segmentation", IEEE Transactions on Medical Imaging, 2018, Volume: 37, Issue: 2.

[7] Haigang Sui, Zhina Song, Dongsheng Gao, Li Hua, "Automatic Image Registration Based on Shape Features and Multi-scale Image Segmentation", 2017 2nd International Conference on Multimedia and Image Processing (ICMIP).

[8] Shih-Hsiung Lee, Chu-Sing Yang, Ting-Wei Hou, Chien-Hui Yeh, "An image preprocessing method for fingernail segmentation in microscopy image", 2017 IEEE 2nd International Conference on Signal and Image Processing (ICSIP).

[9] Nishant Jain, Vinod Kumar, "Approach towards automatic segmentation of diaphragm from ultrasound images", 2017 Fourth International Conference on Image Information Processing (ICIIP).

[10] Annegreet Van Opbroek, Hakim C. Achterberg, Meike W. Vernooij, Marleen De Bruijne, "Transfer Learning for Image Segmentation by Combining Image Weighting and Kernel Learning", IEEE Transactions on Medical Imaging, 2019, Volume: 38, Issue: 1.

[11] Tiechao Jiang, Xiaoqiang Ji, "The Vascular Calcification Image Segmentation Based on CV Model", 2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI).

[12] MojtabaSeyedhosseini, TolgaTasdizen, "Semantic Image Segmentation with Contextual Hierarchical Models", 2016, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume: 38, Issue: 5.

[13] Lingyun Wu, Yang Xin, Shengli Li, Tianfu Wang, Pheng-Ann Heng, Dong Ni, "Cascaded Fully Convolutional Networks for automatic prenatal ultrasound image segmentation", 2017, IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017).

[14] X. Xu, D. Shan, G. Wang, and X. Jiang, “Multimodal medical image fusion using PCNN optimized by the QPSO algorithm”, Appl. Soft Comput., vol. 46, no. C, pp. 588-595, 2016.

[15] M. Vallières, C. R. Freeman, S. R. Skamene, and I. E. Naqa, “A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities”, Physics in Medicine & Biology, vol. 60, no. 14, p. 5471, 2015.

[16] Joshi, K., Diwakar, M., Joshi, N. K., & Lamba, S. (2021). A concise review on latest methods of image fusion. Recent Advances in Computer Science and Communications, 14(7), 2046–2056. https://doi.org/10.2174/2213275912666200214113414

[17] Joshi, K., Joshi, N. K., & Diwakar, M. (2019). Image fusion using cross bilateral filter and wavelet transform domain. International Journal of Engineering and Advanced Technology, 8(4C), 110–115. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85067989461&partnerID=40&md5=58061068d8045b6b28dbf8b47a7cf091

[18] Joshi, K., Joshi, N. K., Diwakar, M., Tripathi, A. N., & Gupta, H. (2019). Multi-focus image fusion using non-local mean filtering and stationary wavelet transform. International Journal of Innovative Technology and Exploring Engineering, 9(1), 344–350. https://doi.org/10.35940/ijitee.A4123.119119

DATA BREACH IN HEALTHCARE INDUSTRY: AN INFRINGEMENT OF RIGHT TO PRIVACY; A MODERN ERA EFFECT

Ummul Waraah[1], Adwitiya Prakash Tiwari[2], Dr. Poonam Rawat[3] & Dr. Radhey Shyam Jha[1]

1. Presidency University, Bangalore, uzmaamu7@gmail.com
2. Advocate, Allahabad High Court, adwitiyaprakash@gmail.com
3. Law College Dehradun, Uttaranchal University Dehradun Uttarakhand, hodlaw@uttaranchaluniversity.ac.in
4. Law College Dehradun, Uttaranchal University Dehradun Uttarakhand, radhey.shyamjha@gmail.com

ABSTRACT

This research paper explores the impact of innovation and technology in the field of healthcare sector. In the era of globalization, ICT has highly impacted in every sector of human agency worldwide by introducing new dimensions and opportunities. Healthcare sector is one of those sectors which gains rapid shift from the conventional practice such as data protection. With the advent of ICT in the healthcare sector apart from data protection it paved ways for many collateral benefits in the industry, this impact simultaneously resulted into incidents of data breach i.e., breach of privacy. In this research paper, researcher has elucidated the convergence of technology and right to privacy.

KEYWORDS: Healthcare, Data, Breach, Right to Privacy, ICT, Globalization.

1. INTRODUCTION

From past several decades' globalization introduced us with new innovative technologies, which was duly accepted and appreciated in countries worldwide. Due to which in past decade India has evidently seen paradigm shift from conventional practice to modern approach in every sector. One of the most promising advents of past century is ICT. In the present every household is equipped with some form of technology in one or other forms. ICT impacted us in numerous ways and made its prominent place in the growth of economy. Even though India is a developing country but it welcomed ICT in pursuance of the recommendations of UNCITRAL and enacted, Information and technology Act, 2000, with due respect to Article 51 clause c of the constitution of India. ICT has provided us ease in our tedious works with less consumption of time. Use of mobile phones, e-commerce transactions, payment through QR mechanism and sui-generis things have become a part of a welfare state. With the acceptance of technology people started building their beliefs in third party with regard to security of their credentials. Innovation demands sustainability for wider acknowledgement but contingent transitional barriers are inevitable. Data breach is one of the persistent failures of ICT, whereas regulatory mechanisms are stringently taking this issue in consideration. Every sector is dealing with this issue in some or other manner. Healthcare is one of the vulnerable sectors which is impacted by such breach. In the constitution of India, right to health and privacy both are enumerated impliedly which demands protection and

give rise to claim in case of breach. This issue raises an alarming concern towards economic development, as data breach is a pitfall in building a fiduciary relationship between ICT and consumers.

2. HEALTHCARE SECTOR AND DATA PRIVACY

India being a developing country is highly impacted by globalization which elated cash and services inflow in the territory. Health is a primary concern for human beings it made the healthcare sector have maximum outreach and profit due to diversification from the earlier conventional methods. Hospitals, equipments & devices related to medical, health insurance, clinical trials, telemedicine, medical tourism, and medical outsourcing are all part of India's healthcare industry. Due to its expansion, people started building trust with respect to their data such as clinical reports, consultations, and Ors. With the impact of globalization, India is also moving towards a digital India mission. The emergence of digitalization in India leads the pathways for various sectors such as online payment transactions or e-commerce. The Healthcare sector also adopted the notion of digital India by addressing data protection in digital form. The expansion speed of healthcare sector is extremely fast. For this rapid growth related in healthcare sector, either in terms of extended coverage of services or investment, the role of both of the public and private sector is appreciable. Among the fastest growing industries, Healthcare sector of India has an important position which emancipate the idea of belief by the consumers. The concept of data privacy can be traced back to Indian history, which emphasizes the need not to divulge topics of worship, sex, and family matters to others. In the 21st century, we are living in a digital era where data privacy is a major concern, especially with reference to the healthcare sector. Data protection in the health care sector is associated with information privacy which requires handling of data with due care and diligence as it is sensitive and confidential data. In India in past encroaching on privacy is noticeable whereas the healthcare sector is also affected by the same. Healthcare data breach leads to the discovery of personal information which reveals their interest and necessity-based requirement in the public domain. Breach of healthcare data facilitates the healthcare industry to easily hunt their targeted consumers to enhance their economic benefits. It is apparent that for economic benefits industries demote the idea behind the right to privacy.

3. RIGHT TO PRIVACY

The idea of privacy is an intrinsic part of a person’s life. In the year 1215 AD by the introduction of the Magna Carta, it was acknowledged that a person’s liberty and freedom is an inalienable right of a human being. Afterward, when World War I and World War II shook the belief in humanity, in the year 1945 United Nation was introduced to re-establish the belief and collective responsibility of every nation in the protection and security of humanity. In pursuance of that in 1948 Universal Declaration of Human Rights was made with respect to giving a list of protection to every human irrespective of their identity. Article 12 of Universal Declaration of Human Rights explicitly mentions every form of encroachment into someone’s privacy is a violation of his basic human rights. The right to privacy existed in ancient India from the time of the Dharmashastra, according to which privacy was linked to virtuous morals, and the concept was well established at the time. In India, though the right to privacy is traceable from ancestors it gained recognition in India due to membership in the UN. In India right to privacy is a controversial matter for a long time. Indian judiciary has

looked into various cases related to privacy from Maneka Gandhi (1978 SCR (2) 621) to Naz Foundation case (160 Delhi Law Times 277, 2009). In 2011 the Privacy bill was also introduced to address innumerable privacy issues. Finally, in the year 2017, a nine judges' bench of Supreme Court in Puttaswamy versus Union of India (AIR 2017 SC 4161) stated that the right to privacy, though not in express words, is a part of fundamental right under the phrase 'Personal Liberty' used under Article 21 of the Indian Constitution.

4. DATA PROTECTIONS LAWS

India with the passage of time emancipated different laws in form of rules and regulations to protect data. India does not have any specific legislation on the concept of data protection. The credibility of data protection can be inferred from the Constitution of India, Information technology Act, 2000, and Indian Contract Act 1872. Article 21 of the Indian Constitution gives the right to a person to claim the benefit of the sanctity of his privacy. Information Technology Act 2000, in various sections deals with the issue of data privacy. In sec.72 of Information Technology Act 2000, it deals with the penalty for breach of confidentiality and privacy. Section 72A also penalizes any kind of disclosure of information without the concerned person's consent punishable. Other sections 65, 66, 10A also hold the validation of data privacy issues. E-Contracts being administered by the Indian Contract Act give acknowledgment to the non-disclosure of information provided to the parties by one another. The Government has notified the Information Technology (Reasonable Security Practices and Procedures and Sensitive Personal Data or Information) Rules, 2011. It deals with the protection of "Sensitive personal data or information of a person". In aggregate India has implied laws dealing with the issue of a data breach. The point of concern is its application of it. Though India has a sufficient number of Laws the lacuna is in the execution and administration of it.

5. HEALTHCARE DATA BREACH

The right to health is a fundamental right provided in the constitution of India along with the right to privacy. Data breach of the healthcare sector is a convergence of two rights simultaneously. Health care is a private and sensitive part of human life which deems to be inevitable in occurrence. Healthcare data breach is one of the most violative acts towards a person's vulnerability. Data breaches in the health care sector not only affect the individuals but also the organizations. Healthcare sector digitalization paved way for opportunities along with cyber security concerns. A recent healthcare data breach of around 120 Million Indians data was hacked where Mumbai's high-end Breach Candy Hospital and Utkarsh Scans were majorly impacted. Worldwide data breaches are occurring. Eskenazi Health and many others got affected by such incidents. This domain needs to be adjudged and scrutinized closely for maintaining fiduciary relationship between patients and hospitals.

6. CONCLUSION AND RECOMMENDATIONS

Healthcare is a vulnerable aspect of every individual from cradle to grave. In a lifetime everyone has gone through some or other form of health care but we never bothered about breaches of our data. This issue is highlighted after the advent of the digital era in our community. Digital acceptability is important for economic benefits, perpetuity, avoidance of human fallacy, and many other reasons. But there is a grey area between acceptability and efficiency in execution. India has adapted but now we have to make our mechanism more

definite and clearer. We have introduced digitalization of data but we lack in spreading basic digital literacy. On the one hand, we have introduced data collection in digital form but a protection firewall against hacking from the dark web that we have not yet figured out. India as a developing nation needs to propose one stringent law for the protection of data privacy taking into consideration the requirements of different sectors and targeted consumers. India can take reference from other countries which are efficiently dealing with data breach issues.

7. REFERENCES

[1]. Biswas U.N., Allard K., Pousette A., Härenstam A. (2017) The Healthcare Sector. In: Understanding Attractive Work in a Globalized World. Springer, Singapore. https://doi.org/10.1007/978-981-10-6133-2_9

[2]. Batar S,, Chaudhary A., Jha R, S., Tyagi. N., Jain V., 'Protection of Software- Patent or Copyright?' (2020) 'Our Heritage' ISSN: 0474-9030 Vol-68-Issue-1-January

[3]. Batar S,, Tyagi.N., Jha R, S. & Chaudhary A., Tyagi.N., 'A Short History of Human Rights (Ancient Period- LPG Era).' (2020) 'Our Heritage' ISSN: 0474-9030 Vol-68-Issue-1-January

[4]. "Copyright Law in the Digital World", Springer Science and Business Media LLC, 2017

[5]. https://www.archives.gov/exhibits/featured-documents/magna-carta (Last accessed on 22 February 2022)

[6]. Team Inc 42, Healthcare Data Leak: Over 120 Mn Medical Images Of Indian Patients Left Exposed, 04/02/20, India Healthcare Data Leak: Over 120 Mn Medical Images Exposed, inc42.com, (Last accessed on 23 February 2022).

[7]. Vijay Pal Dalmia, India: Data Protection Laws In India - Everything You Must Know, 13/12/2017, Data Protection Laws In India - Everything You Must Know - Privacy – India, mondaq.com, (Last accessed on 22 February 2022).

THERMO-HYDRAULIC PERFORMANCE ENHANCEMENT OF A RIBBED CHANNEL DUCT USING ANN & TAGUCHI APPROACH

Saurabh Aggarwal[a],Sanjeev Kumar Joshi[b],Vivek John[c], Shushat Singh[d], Dr. Van Cuong Nguyen[e]

[a,b,cd] Department of Mechanical Engineering, Uttaranchal Institute of technology, UttaranchalUniversity

[e]Healthcare Big data Research Center, Hanyang University, Korea

Abstract

Thermo-Hydraulic Performance ("THP", η) waspredicted using an ANN model for a fixed channel with the turbulators fixed at 90 degrees and compared by CFD analysis conducted by applying the Taguchi approach.The ribs were arranged in a straight line alternately, and the numerical values of the coefficient of friction and the coefficient of convective heat transfer were evaluated. The L_{16} orthogonal array was adopted to evaluate the thermal performance. The relative rib pitch to height (p / e) varies from 3 to 12 in steps of three. The rib vertical angle converges from the plane of the square ribs (i.e. 90°) 80°, 70° to 65° (trapezoidal ribs). Heat transfer & friction coefficient aredetermined using Reynolds numbers of 4000, 8000, 12000, 16000. The ANN model predicted the values close to the CFD Analysis and it was found that Equilateral triangular ribs with $(\alpha)_1(Re)_1(p/e)_3$have the best THP.

Keywords: ANN, Taguchi, Thermo- Hyrdraulic Performance, CFD

1. INTRODUCTION

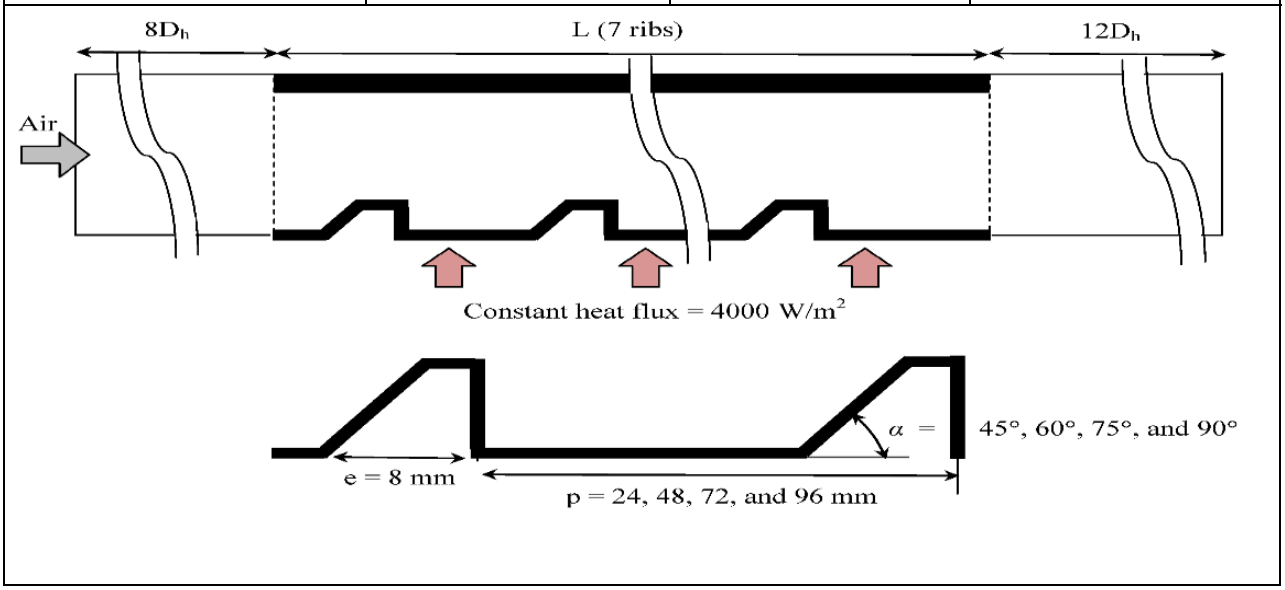
Thermal applications in Industries employ Heat exchangers and major design challenge in a heat exchanger is to make compact, high heat transfer rate equipment using minimum power[1]. To improve heat transfer performance in ducts, Channels and other heat transfer applications, techniques like fin, protrusion, dimple , vortex generators, groove and ribs are used [2]. Heat transfer and friction coefficientsof rectangular channel having inclined ribs were discretized throughgeneration of not inline gaps [3]. Different cross sections air flow paths at sides of parallel pass in the absorber plates are roughened with inclined shape ribs [4]. Significant output parameters like Nu, friction coefficient and thermal parameters like Re,P/e and e/Dh was analyzed for T shaped Ribs [5]. The THP of triangular duct with ribs of varying inclination in turbulent flow fieldhas been studied [6]. Computational Fluid Dynamics along with Taguchi design has been performed for investigation of thermo-hydraulic performance [7].The Fluid Dynamics can be performed with the help of input images[8]. The cubic Interpolation method can be used for Images to investigate Thermo -hydraulic[9].Thermal Application can be employed with self compacting concrete methods[10].

2. METHODS AND METHODOLOGY

In the present study, a double walled duct with upper wall insulated and lower wall with a series of Trapezoidal ribs (with equilateral angle varying from 65˚, 70˚, 80˚ and 90˚) and other input parameters are taken as shown in table 1..The totallength of the duct was taken 1500mm, entry length of 320 mm, test length of 700mm,exit length of 480mm,duct width(W) of 160 mm, duct height, (H) 40mm, hydraulic diameter (D_h) of 64mm, uniform heat flux (I) and Prandtl Number (Pr.) of 4000 W/m^{2}and 0.707 respectively.

Table 1.Input parameters and investigated geometry of ribbed channel

Parameters Symbol	p/e	α	Re
Range of values	3 – 12 (step size of 3)	65˚, 70˚, 80˚, 90˚	4000 – 16000 (step size of 4000)



2.1. ANN Modelling

The modelling for input and output was done through ANN toolbox through MATLAB software. The simulated MLFF-ANN architecture consisting 3 neurons in input layer one for each considered process inputs and the output layer comprising of 3 neurons foreach output characteristics. Table 2. Shows the CFD analysis results along with ANN predictions.

Table 2. Main Experimental Table with ANN Modelling

S.No.	Input Parameters			CFD Analysis			ANN Predictions		
	Re	P/e	A	Nu	F	η	Nu	f	η
1	4000	3	65°	43.35	0.0104947	2.73717	43.3	0.011	2.7223
2	4000	6	70°	40.48	0.0132207	2.36649	40.52	0.0135	2.3354
3	4000	9	80°	40.97	0.0153632	2.27790	41.02	0.01501	2.1223

4	4000	12	90°	41.01	0.0175015	2.18363	40.98	0.01657	2.1789
5	8000	3	70°	62.24	0.0099913	2.16550	62.26	0.0103	2.1872
6	8000	6	65°	65.53	0.0112298	2.19283	65.43	0.01212	2.2020
7	8000	9	90°	65.44	0.0118445	2.15129	65.32	0.011527	2.65
8	8000	12	80°	67.44	0.0141818	2.08784	67.58	0.019654	2.39
9	12000	3	80°	74.19	0.0086417	1.89350	74.41	0.008268	1.59
10	12000	6	90°	76.58	0.0101059	1.85531	76.68	0.009085	1.37
11	12000	9	65°	95.51	0.0122588	2.16949	99.45	0.0.12310	2.73
12	12000	12	70°	91.67	0.0129090	2.04672	93.24	0.012782	2.22
13	16000	3	90°	83.63	0.0082362	1.68231	83.45	0.008685	1.52
14	16000	6	80°	96.67	0.0099908	1.82340	93.71	0.0099	1.72
15	16000	9	70°	113.59	0.0120500	2.01278	114.11	0.12380	2.08
16	16000	12	65°	113.09	0.0125956	1.97456	115.07	.013056	1.91

3. RESULT AND DISCUSSION

The current investigation data is presented in terms of contour/vector plots of velocity distribution superposed with mean streamlines in the inter-rib region. The streamline patterns superposed on dimensionless mean streamwise velocity are shown in Fig1.for rib p/e ratio varying from 3 to 12. It is observed that the flow characteristics in the inter-rib regions after 3rd rib are matching well irrespective of p/e, and therefore, the flow can be considered as periodic.

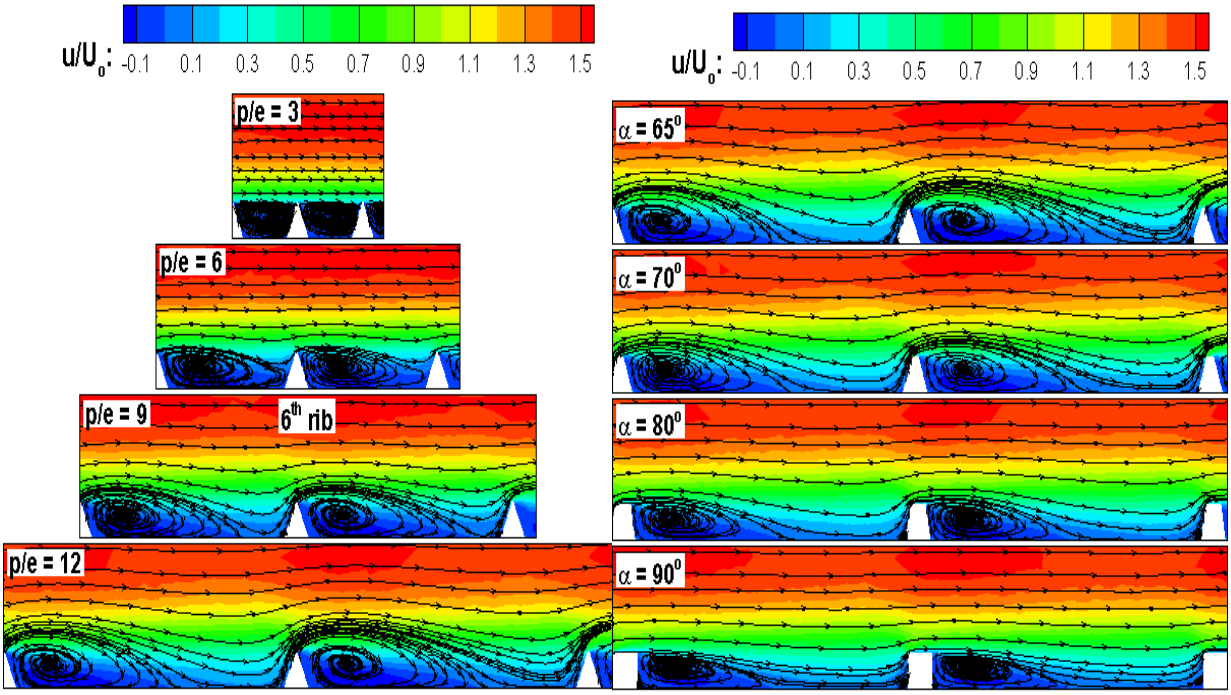


Fig.1. Effect of rib p/e ratio on flow structures and Effect of rib geometry on flow structures

Table 3. Signal to Noise ratio for Nu, f and THP(η)

Input Parameters			CFD Analysis			Signal to Noise Ratio		
Re	P/e	A	Nu	f	η	Nu	f	η
4000	3	65°	43.35	0.0104947	2.73717	32.74	39.5806	8.75
4000	6	70°	40.48	0.0132207	2.36649	32.15	37.5749	7.48
4000	9	80°	40.97	0.0153632	2.27790	32.25	36.2704	7.15
4000	12	90°	41.01	0.0175015	2.18363	32.26	35.1385	6.78
8000	3	70°	62.24	0.0099913	2.16550	35.88	40.0075	6.71
8000	6	65°	65.53	0.0112298	2.19283	36.33	38.9925	6.82
8000	9	90°	65.44	0.0118445	2.15129	36.32	38.5297	6.65
8000	12	80°	67.44	0.0141818	2.08784	36.58	36.9654	6.39
12000	3	80°	74.19	0.0086417	1.89350	37.41	41.2680	5.55
12000	6	90°	76.58	0.0101059	1.85531	37.68	39.9085	5.37
12000	9	65°	95.51	0.0122588	2.16949	39.60	38.2310	6.73

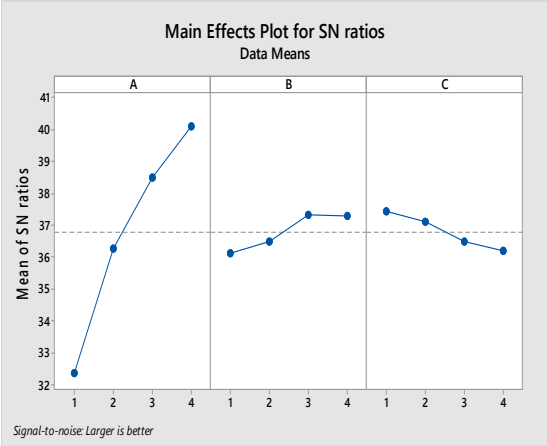
12000	12	70°	91.67	0.0129090	2.04672	39.24	37.7821	6.22
16000	3	90°	83.63	0.0082362	1.68231	38.45	41.6854	4.52
16000	6	80°	96.67	0.0099908	1.82340	39.71	40.0080	5.22
16000	9	70°	113.59	0.0120500	2.01278	41.11	38.3802	6.08
16000	12	65°	113.09	0.0125956	1.97456	41.07	37.9956	5.91

After conducting the CFD analysis based on the recommended experiment set by Taguchi orthogonal array the results of signal-to-noise ratio (S/N) were analyzed. Signal and noise depicts controllable and noncontrollable variables in aphenomenon. Assigned implementations with their corresponding S/N values for Nur, f and THPare presented in Table 3.

Table 4. Response table of S/N ratios & Main effects plot for S/N ratio of Nusselt number (a), Friction Factor (b) and Thermo-Hydraulic performance (c)

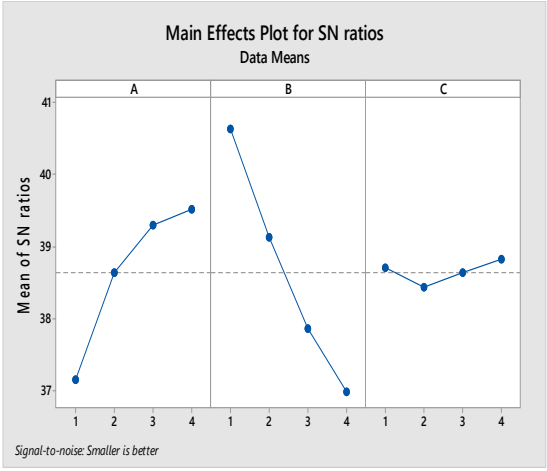
(a)

Level	Re	p/e	α
1	32.35	36.12	37.43
2	36.28	36.47	37.09
3	38.48	37.32	36.48
4	40.08	37.29	36.18
Delta	7.73	1.2	1.26
Rank	1	3	2
CR, (%)	75.86	11.78	12.37



(b)

Level	Re	p/e	α
1	37.14	40.64	38.7
2	38.62	39.12	38.44
3	39.3	37.85	38.63
4	39.52	36.97	38.82
Delta	2.38	3.66	0.38
Rank	2	1	3
CR, (%)	37.07	57.01	5.92



(c)

Level	Re	p/e	α

1	7.541	6.38	7.051	
2	6.645	6.222	6.623	
3	5.965	6.652	6.077	
4	5.43	6.327	5.831	
Delta	2.11	0.43	1.22	
Rank	1	3	2	
CR, (%)	56.12	11.44	32.45	

The difference in Max and Min S/N ratio is Delta while the ratios of these Delta indicatesthe contribution of each of the parameter under investigation which are being represented in Table 4. along with the Main effects plot for S/N ratios for output characteristics under investigation. Nusselt number andthermohydraulic performance follows larger SNR is better, whereas Friction Factorfollows smaller SNR is better. the Nu increases as Re increases, with increasing p/e, as the flow reattaches resulting inincreased heat transfer. Whereas itdecreases as α increases. It is alsoobserved thatf increases asRe increase and decrease p/e increases. The values of 'f' slightly decreaseas α increases.

4. CONCLUSIONS

The THPwas computed for stationary channels with ribs and four levels of the three input parameters.From the CFD analysis based onL_{16} orthogonal array of Taguchi approach. The subsequent conclusions have been drawn:

1. Nu increased with p/e ratio having optimum value at $Re_4p/e_3\alpha_1$ the outcomes indicates that major contributor was Reynold Number (Re).
2. p/e ratio is major contributor for Friction and optimum parameters Friction factor are at $Re_4p/e_1\alpha_4$.
3. Overall THP decreases with increase in Re values and are almost unaffected by change in p/e values and decreases as α increases. Overall maximum performance was obtained at $Re_1p/e_3\ \alpha_1$.
4. It was also observed that the ANN model predicted the values close to those of the CFD investigation.

5. REFERENCES

[1] S. J. Thikane, "A Comprehensive Review of Various Heat Transfer Enhancement Techniques used to improve Thermo-hydraulic Performance of Rectangular Duct," vol. XII, no. Vi, pp. 227–232, [Online]. Available: http://jctjournal.com/gallery/41-june2019.pdf.

[2] A. Kumar and N. Tenguria, "Computational Investigation on Thermo-Hydraulic Performance Characteristics of Ribbed Passage with Turbulence Generators," *Int. J. Sci. Res. Eng. Dev.*, vol. 2, [Online]. Available: www.ijsred.com.

[3] M. K. Solanki and K. R. Aharwal, "Thermo-Hydraulic Performance of a Rectangular Duct With Staggered Inclined Discrete Rib Arrangement," no. July 2020, 2019, doi: 10.35940/ijitee.A4697.119119.

[4] C. Sivakandhan, T. V Arjunan, and M. M. Matheswaran, "Thermohydraulic performance enhancement of a new hybrid duct solar air heater with inclined rib roughness," *Renew. Energy*, vol. 147, pp. 2345–2357, 2020, doi: 10.1016/j.renene.2019.10.007.

[5] Y. Mahanand and J. R. Senapati, "Thermal enhancement study of a transverse inverted-T shaped ribbed solar air heater Thermal enhancement study of a transverse inverted-T shaped ribbed solar air heater," no. November 2020, 2021, doi: 10.1016/j.icheatmasstransfer.2020.104922.

[6] K. Nidhul, S. Kumar, A. K. Yadav, and S. Anish, "Computational and experimental studies on the

development of an energy-efficient drier using ribbed triangular duct solar air heater," *Sol. Energy*, vol. 209, no. March 2021, pp. 454–469, 2020, doi: 10.1016/j.solener.2020.09.012.

[7] S. Singh and S. Lakhera, "CFD & Taguchi Analysis for Optimization of Geometrical and Flow Parameters in A Ribbed Channel Duct," vol. 2, no. 4, pp. 140–154, 2018.

[8] M. Pandey, R. K. Bharti and A. K. Bhatt, "A Study of Color Enhancement Techniques for Input Images," *2017 2nd International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS)*, 2017, pp. 1-7, doi: 10.1109/CSITSS.2017.8447690.

[9] M. Pandey, "Futuristic Hybrid Image Enhancement Using Fuzzy and Cubic Interpolation Methods," *2021 International Congress of Advanced Technology and Engineering (ICOTEN)*, 2021, pp. 1-6, doi: 10.1109/ICOTEN52080.2021.9493446.

[10] Singh, P., Usman, M., Chandramauli, A., & Kumar, D. (2018). Brief experimental study on self compacting concrete. *International Journal of Civil Engineering and Technology*, *9*(5), 77–82. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85047820420&partnerID=40&md5=a97197db8ae8a6e388c889b43314a542

Economic impacts of artificial intelligence in India

Reeta Rautela[1], Ashutosh Sharma[2], Gaurav Thakur [3], Vinod Balmiki [4], Shravan Kumar[5], Shruti Goel[6]

[1]*School of Liberal Arts, Uttaranchal University, Dehradun, 248007, India; reetarautela@uttaranchaluniversity.ac.in*

[2]*School of Liberal Arts, Uttaranchal University, Dehradun, 248007, India; ashutosh.sharma@uttaranchaluniversity.ac.in*

[3]*Department of Civil Engineering, Uttaranchal Institute of Technology, Uttaranchal University, Dehradun, 248007, India; gauravthakur@uttaranchaluniversity.ac.in*

4Department of Civil Engineering, Uttaranchal Institute of Technology, Uttaranchal University, Dehradun, 248007, India; vinodbalmiki@uttaranchaluniversity.ac.in

[5]*School of Liberal Arts, Uttaranchal University, Dehradun, 248007, India; deansla@uttaranchaluniversity.ac.in*

[6]*Research Scholar, University of California San Diego, California, USA*

Abstract.

The application of knowledge and technology has proved to be a new horizon in economics also. The current economy is technology based economy. In today's globalized and digitalized economy, there has been a progressive advancement in artificial intelligence. From the perspective of economics, artificial intelligence is cost efficient as along with decreasing the cost it increases the efficiency as well. It has opened the doors with ample opportunities for trade both at national and international level. This paper is an attempt to identify the focus areas of artificial intelligence and explores the economic impacts of artificial intelligence in India and also discuss about the Policy initiatives to support the AI segment in India.

Keywords.artificial intelligence, economy, economic impact, knowledge and technology, AI segment

1. INTRODUCTION

In the recent years, there has been an increasing trend towards the knowledge based technology that is the result of innovation. Artificial Intelligence deals with the innovations that act like human being [1]. In today's life artificial intelligence is all around us. It has become a part of our daily life.

AI has its roots in many disciplines including Math, Psychology, Biology, Logics. Although the field of logic and math, philosophy and psychology are interrelated.

The impacts of artificial intelligence [2] can be seen in all economic sectors. The integration of economies during the process of globalization has promoted international trade. The use of artificial intelligence in all the economic sectors have shown tremendous

growth in terms of international trade. India is 6th largest economy in the world. According to the report published by [3] World Bank, India's trade as a percentage of GDP in 2020 will be 37.87%. Being the largest exporter of technology-based services, the service sector has contributed higher to its economic growth. Artificial intelligence has transformative impact on international trade. It has encouraged the skill based employment along with significant investments to make the economies more developed, enhancing the productivity by accelerating the international trade [4].

2. MAJOR FOCUS AREAS OF ARTIFICIAL INTELLIGENCE

All the leading technological based industries are using artificial intelligence to expand their size in market. There are exciting potentials of economic boost due to reshaping of the industries with artificial intelligence. Even during the recent COVID-19 pandemic, the economies throughout the world were blessed [5] for not facing the stationary state just because of artificial intelligence. According to the report of Artificial Intelligence Index 2021, produced by Human-Centered Artificial Intelligence [6] of Stanford University, there was a tremendous growth of conferences based on artificial intelligence as the conferences held on virtual mode and the participation number was overwhelming. As per this report the private artificial intelligence investment has shown [7] significant increase in drug design and discovery (i.e. USD 13.8 billion).

2.1. Sub-fields of Artificial Intelligence

The sub field of artificial intelligence include Robotics, techniques using time series analysis for predictions, algorithms, expert support systems, Natural Language Processing, Planning, Machine learning, Vision, speech recognition, brain modelling, vision space learning, image based, gaming techniques and many more [8].

2.1.1 Focus areas of AI in India:

According to a published report on Discussion paper on the National [9] Artificial Intelligence Strategy (AI for All) by Niti Ayog, the focus areas of AI in India are:

- Health care
- Agriculture
- Education
- Smart Cities and Infrastructure
- Smart mobility and Transportation

3. ARTIFICIAL INTELLIGENCE AND ECONOMICS

All Artificial Intelligence driven technology is fundamentally based on innovations [10] that have economic incentives too. Like natural Science, economics is not a pure science because the theories and models are based on assumptions. Economic forecasting has many challenges.

In economics, there are many challenges in business forecasting as it is all based on assumptions and also in the absence of evidence based data, the forecasting is not authentic. AI uses algorithms to predict, so this issue of business forecasting can be [11] better sorted out. For Business purpose, data require predictions based on human behaviour and first information is based on qualitative analysis, thereby implementing AI techniques in behavioural economics can be better judged to find out human perceptions. The use of statistical data will be more authenticated as well. Government authorities may use this data to frame better policies by a proper analysis that is based on evidences and correct predictions in behavioural economics. The economic downturn can be better predicted by the use of algorithms that can track recession, predict business cycles, predict the changes in consumer behaviour and many more. This may be helpful to frame [12] Monetary and Fiscal Policies accordingly. The emergence of AI in economics will be helpful for the government to formulate and analyse the policies and these policies will be evidence based economic policies and with the involvement of AI, economics will come closer to natural science.

4. ARTIFICIAL INTELLIGENCE'S ECONOMIC IMPACT IN INDIA

As per the Report of 'The India Express', around 20 million jobs are to be added by 2025 around artificial intelligence. Due to recent innovations in the technology based industries all the spheres of mankind will be touched and explored by AI. There are many spheres like mental health, agriculture, planning of smart cities, disaster management, counselling on education and health, planning about traffic system, AI will develop all. Government of India had set a target to achieve 20 million jobs by 2025.

The major impacts of Artificial Intelligence on economy can be seen from the Artificial Index Report 2021, where it was reported that the highest private AI investment has been found to be of more than USD 13.8 billion in the areas of Drugs, Cancer, Molecular, Drug Discovery.

It was also reported that during recent pandemic, the AI [13] hiring have shown the remarkable growth in the countries like Brazil, India and Singapore.

- The major industries to get benefits from AI are: Information and Communication Technology, Professional, Scientific and tech services, Manufacturing, hunting, Public Administration, Education, Banking and Finance, Mining.

- As per the LinkedIn data published in AI Index Report, 2020 India is second in terms of highest growth in AI hiring (Fig. 1) and second in terms of AI skill penetration (Fig. 2).

The use of artificial intelligence will be significant in analysing the GDP over the years. As the predictions of macroeconomic variables may be useful to know the causes behind the slow or fast growth rate of the economy or the GDP as a percentage and it will also enable the economies to make a reliable and unbiased comparisons in the economies. The application of macroeconomic theories can be better applied if based on predictions made by artificial intelligence. The relationship among economic variables can be useful for policy makers and the errors in forecasting macroeconomics and other areas may be minimised. More technologies like IoT, cloud [14-16] will work for the development for the development of india

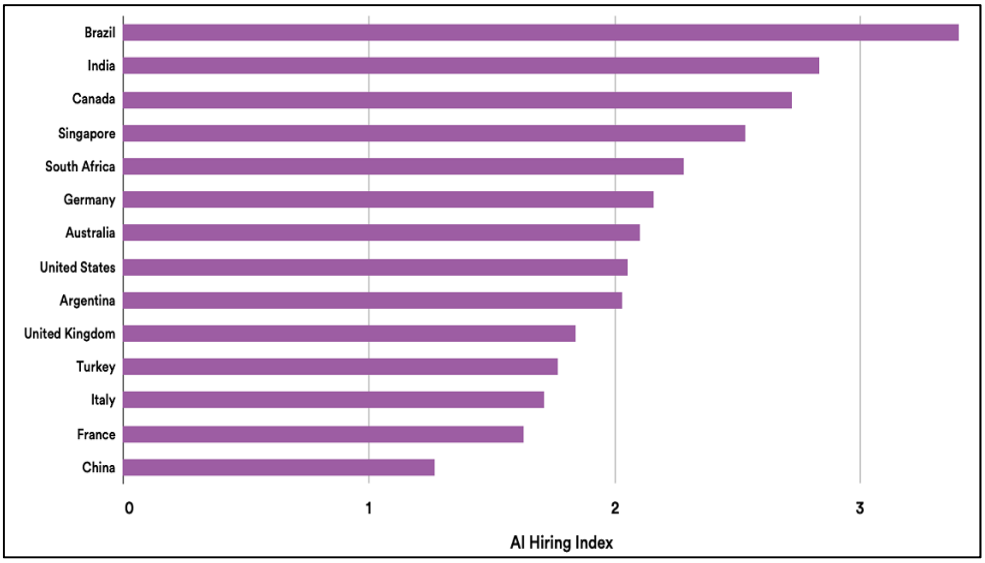


Fig. 1: Country wise AI employment index, 2020
Source: LinkedIn, 2020 | Chart: 2021 AI Index Report
(Retrieved from-https://datascience.fm/2021-ai-index-report-newsletter/)

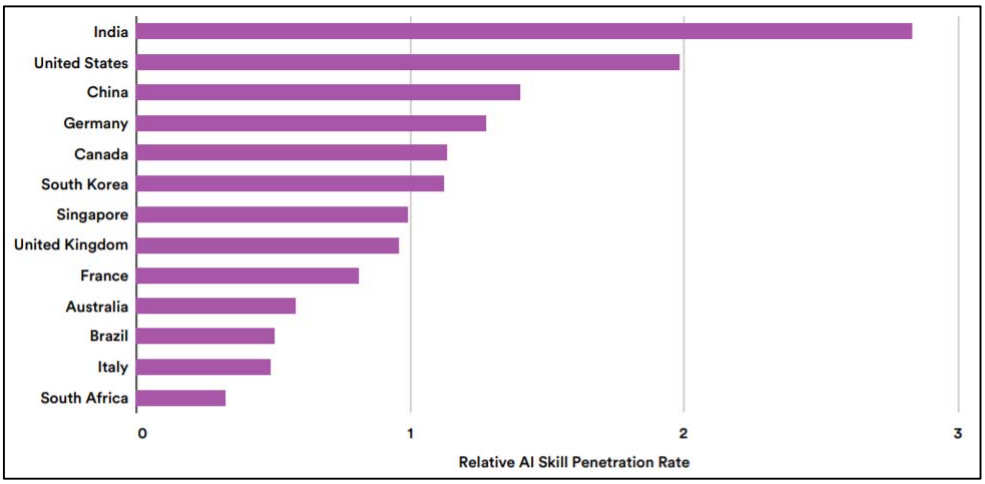


Fig. 2: Relative AI Skill Penetration Rate by Country, 2015-20

Source: LinkedIn, 2020 | Chart: 2021 AI Index Report
(Retrieved from-https://digiscorp.com/blog/industries-technologies/state-of-artificial-intelligence-in-europe- statistics-2021/)

INDIAai is the pivot for all small details AI in India and even beyond. It is a collaboration between MeitY, NeGD, and NASSCOM. According to a report by NASSCOM and INDIAai, In India, AI patents account 6% of all developing technology patents. Some of the key trends in AI patent filing in India are:

- From 2015-20, appx. 92% patents are from AI
- MNCs have contributed appx 60% of patents that have been granted
- It was also reported that India has shown a growth in patents (about 64% newly originated)
- Personal health care and Consumer Electronics have shown remarkable progress in patents.
- The report explored that major sector that is leading in patents belong to IT sector.

The development of AI has macroeconomic impacts. With increased AI investments the skill based service sector is the emerging sector and it is transforming the economy towards services economies. AI based industries have opened the doors for small businesses to enter in to trade at international level via digital platforms.

5. POLICY INITIATIVES TO SUPPORT THE AI SEGMENT IN INDIA

According to Government of India Report, to share the ideas and experiences, US and India have launched, US-India Artificial Intelligence initiative on March 18, 2021 and to have application based research, on the same day it was also announced to create Centres of Excellence at national level for Transformation AI.

To encourage and motivate students and make them compatible with AI, Government of India has adopted the incorporation of AI in the new National Education Policy (NEP) 2020 and National Research Foundation has been constituted for the same. To encourage research and development there is a strong need to bring together academia and industry. To incorporate AI in the syllabus, Secondary level education has been opted for the same. To promote AI, the Indian government increased the Expenditure to $477 million. Although AI has some challenges also that includes, ethical issues, environmental issues, misuse of social media, cyber security, data privacy is another significant issue, but we can't deny the fact that the next age is of AI.

In India, AI can play a significant role in Social Development with inclusive growth. Inclusive growth will provide a quality life to the individuals and also improve the productivity of factors thus enhance the industrial growth as well. Artificial intelligence in education will create interest among students to come forward with new ideas to make innovations and self-reliance. Although the potential of artificial intelligence in the Indian economy is very high but along with opportunities it has many challenges as well. Lack of ethical policy norms, financial support, Strong Research and Development, technology transfer and many such. But this can be made possible by the government support to the private sectors. The framework of opportunities should not be limited to only a limited institutes or sectors rather it should be openly accessible to all.

6. REFERENCES

[1]Goldfarb, A & Trefler, D. (2018) "AI and International Trade". NBER Working Papers 24254, National Bureau of Economic Research, Inc

[2]Telecommunication Engineering Center- Future Networks (FN) Division. (August 2020). Study Paper on Artificial Policies in India-A Status Paper.

[3]IAI: The Roots, Goals and Sub-fields of AI © John A. Bullinaria, 2005. Retrieved from: https://www.cs.bham.ac.uk/~jxb/IAI/w2.pdf

[4]Niyazov, S. November 10, 2019. How AI will Redefine Economics. Retrieved from https://towardsdatascience.com/how-ai-will-redefine-economics-ec305e3cb

[5]Saxena, P. Dec. AI impact on India: Jobs and employment. Retrieved from https://indiaai.gov.in/article/ai-impact-on-india-jobs-and-employment

[6]https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-applications-and-value-of-deep-learning

[7]AI index report 2021. Retrieved from: https://ai100.stanford.edu/2016-report/section-i-what-artificial-intelligence/ai-research-trends

[8] Gazala, N. AI in India: Initiatives from the Indian Government in 2021. Retrieved from: https://indiaai.gov.in/article/ai-in-india-initiatives-from-the-indian-government-in-2021

[9]Shanthi S. What are the key AI initiatives for Indian Government. Retrieved from: https://analyticsindiamag.com/what-are-the-key-ai-initiatives-of-indian-government/

[10]AI Patents: Driving emergence of India as an AI innovation hub. Retrieved from:https://nasscom.in/knowledge-center/publications/ai-patents-driving-emergence-india-ai-innovation-hub

[11] Malik, Praveen, Singh, R., Gehlot, A., Akram, S. V., & Das, P. K. (2022). Village 4.0: Digitalization of Village with Smart Internet of Things Technologies. Computers & Industrial Engineering, 107938.

[12] Bagwari, S., Gehlot, A., Singh, R., Priyadarshi, N., & Khan, B. (2022). Low-Cost Sensor-Based and LoRaWAN Opportunities for Landslide Monitoring Systems on IoT Platform: A Review. IEEE Access, 10, 7107–7127. https://doi.org/10.1109/ACCESS.2021.3137841

[13] Malik, P, Gehlot, A., Singh, R., Gupta, L. R., & Thakur, A. K. (2022). A Review on ANN Based Model for Solar Radiation and Wind Speed Prediction with Real-Time Data. Archives of Computational Methods in Engineering. https://doi.org/10.1007/s11831-021-09687-3

[14] Singh, R., Gehlot, A., Rashid, M., Saxena, R., Akram, S. V, Alshamrani, S. S., & Alghamdi, A. S. (2021). Cloud server and internet of things assisted system for stress monitoring. Electronics (Switzerland), 10(24). https://doi.org/10.3390/electronics10243133

[15] Karunanidy, D., Ramalingam, S., Dumka, A., Singh, R., Rashid, M., Gehlot, A., ... & AlGhamdi, A. S. (2022). JMA: Nature-Inspired Java Macaque Algorithm for Optimization Problem. Mathematics, 10(5), 688.

[16] Sharma, V. K., Singh, R., Gehlot, A., Buddhi, D., Braccio, S., Priyadarshi, N., & Khan, B. (2022). Imperative Role of Photovoltaic and Concentrating Solar Power Technologies towards Renewable Energy Generation. International Journal of Photoenergy, 2022.

Advancements in IoT Based Intelligent Systems for Bridge Health Monitoring

Gaurav Thakur [1], Vinod Balmiki [2], Anita Gehlot[3], Sanjeev Kumar Shah [4], Shiv Dayal Pandey [5], Mohammed Ismail Iqbal[6]

[1] *Department of Civil Engineering, Uttaranchal Institute of Technology, Uttaranchal University, Dehradun, 248007, India; gauravthakur@uttaranchaluniversity.ac.in*

[2] *Department of Civil Engineering, Uttaranchal Institute of Technology, Uttaranchal University, Dehradun, 248007, India; vinodbalmiki@uttaranchaluniversity.ac.in*

[3] *Division of Research & Innovation, Uttaranchal Institute of Technology, Uttaranchal University, Dehradun, 248007, India; eranita5@gmail.com*

[4] *Department of Electronics & Communication Engineering, Uttaranchal Institute of Technology, Uttaranchal University, Dehradun, 248007, India; principal.ucp@uttaranchaluniversity.ac.in*

[5] *Department of Mechanical Engineering, Uttaranchal Institute of Technology, Uttaranchal University, Dehradun, 248007, India; deanuit@uttaranchaluniversity.ac.in*

[6]*University of Technology and Applied Sciences, Nizwa, Oman; mohammad.iqbal@nct.edu.om*

Abstract

With the advancements in the field of Internet of Things (IoT) and its implementation to Structural Health Monitoring (SHM), IoT has garnered attention from the academic as well as the industrial community. Health monitoring of bridge structures is consequential because of their security concerns and strategic significance. IoT has proved to be relevant for bridge health monitoring because of its ability to monitor health in real-time, collect, transmit, store, analyze as well as generate warnings. Research has also been conducted and improvised cost-effective, as well as energy-efficient IoT systems, have been proposed. In this paper, a detailed study of the latest advancements in the field of applications of IoT in bridge health monitoring have been studied based on several parameters such as advancements in cloud architecture, advanced techniques for fusion with IoT, and cost and energy effective IoT systems, and suggestions have been made.

Keywords. Structural Health Monitoring (SHM), Cost-Effective Monitoring System, Energy Efficient Monitoring System, Bridge Health Monitoring, Internet of Things (IoT), IoT fused with Advanced Techniques

1. INTRODUCTION

Bridges have long been the backbone of the transportation industry as well as the development of a nation by facilitating the connection between important locations and

hence serve as infrastructures of security concerns and strategic significance. With the advent of rapid development, bridges have been tested with the passage of time due to changes in loading pattern, weather conditions, and degradation of bridge material often leading to fatal collapse of the bridge structures leading to huge loss of lives and economy. Thus making bridge health monitoring indispensable. With the traditional monitoring in place, it was very difficult to investigate the entire bridge structures manually; not sooner till Internet of Things (IoT) advent into the field of Structural Health Monitoring (SHM).

IoT is a system of physical objects embedded with sensors used to collect, store, analyze and exchange the data with other devices which are connected with other systems over a communication network. IoT has long been successfully implemented for health monitoring and as early warning systems to improve operational efficiency and level of safety in bridges [1]. The IoT systems used for bridge health monitoring often comprise of sensors connected by IoT, monitoring system, data storage and a data processing module that evaluates data in real-time [2].IoT has also been used to assess inclined angles, bends of various parts of the bridge, monitor the looseness of the bolt joints, and project this data to a smartphone app, generating warnings by flashing LEDs [3]. Health monitoring of a cluster of bridges has also been made possible with the use of wireless gateway and cloud platform [4].

Much of the advancements, related to IoT in Bridge Monitoring Systems, have been done in the area of cloud architecture and cost and energy effective IoT enabled monitoring systems. Improvised IoT enabled bridge monitoring systems fused with advanced techniques such as machine learning, artificial intelligence etc. have also developed. Section 2 discusses the advancements in all of these areas.

2. ADVANCEMENTS IN INTELLIGENT SYSTEMS EMBEDDED WITH IOT FOR BRIDGE HEALTH MONITORING

This section covers the advancements in the field of IoT enabled bridge health monitoring systems. Three criterions, namely, advancements in cloud architecture, fusion of advanced techniques in IoT enabled health monitoring systems and cost and energy efficient IoT enabled health monitoring systems have been covered.

2.1 Advancements in Cloud Architecture for IoT Based Bridge Health Monitoring

Data collection and exchange have been very important aspects of IoT. Research works have been conducted to develop effective systems for data exchange. For monitoring bridge health, a cloud-based cyberinfrastructure platform based on peer-to-peer distributed database architecture has been built [5]. It allows for scalable data administration, sharing, and use. Amazon Web Services (AWS), a cloud computing platform, has also been used to link IoT devices as well as to store and analyse data. Sensors in AWS-integrated systems relay data to nodes via the Xbee protocol (raspberry pi). Data is then sent to the cloud using message protocols [6].

2.2IoT Fused with Advanced Techniques Systems for Bridge Health Monitoring

Advanced techniques such as machine learning, neural networks, artificial intelligence, soft computing, etc. are employed to identify and locate the damaged area in a bridge structure [7]. Machine Learning (ML) and Artificial Intelligence (AI) are employed for bridge health monitoring including irregularity or crack detection. Traditional prediction algorithms from regression modelling and data mining may encompass problems of data scaling and difficulty adapting to dynamic changes in the data input. To overcome this drawback, bridge health is monitored by integrating predictive analytics based on artificial intelligence into the IoT sensors [8].

Stable systems comprising of Convolutional Neural Network (CNN) fused with AI and Raspberry Pi 4 embedded computer integrating IoT has been developed to monitor bridge health by Google Firebase cloud database based apps and websites [9]. Novelty detection and location algorithms based on machine learning have also been created for monitoring data. Using the collected data such as external conditions, interior temperature, or data from other sensors and training set, neural networks assisted multiple regression models are built for each sensor [10]. Deep neural networks have also been used to effectively monitor the health of bridges. Auto-encoder-based methods inspired by the deep neural network's ability to automatically extract feature have proved to be effective for bridge health monitoring [11]. In other systems, bridge crack photographs collected and processed by neural network model-simulated and trained by MATLAB have also been shown to have overall accuracy greater than 90% [12]. IoT systems are also used in conjunction with a Fiber Bragg Grating (FBG) optical sensing technology to sense the load and strain distribution on bridges in order to keep track of unanticipated conditions caused by load variations. At each data source, various techniques such as random forest (RF) and K-nearest neighbor (KNN) were used [13].Advancements have also been done in the area of reducing the latency response and extending network lifetime while guaranteeing desired coverage of the IoT systems by bringing computing close to the source of data production. SHM Fog and Edge computing architectures have been developed for this purpose [14]. Learning automata (LA) is a reinforcement learning model that provides sensor scheduling strategy which is energy-efficient by integrating Confident Information Coverage (CIC) model in an IoT-based bridge health monitoring system to ensure network coverage and extend network lifetime [15].

2.3Cost & Energy Effective IoT Systems for Bridge Health Monitoring

Cost-effectiveness and energy-effectiveness have always been among the major deciding factors for the adoption and implementation of a system. A structural health monitoring system based on Internet of Things (IoT) have proved to be low-cost, energy efficient, secure and reliable system [16]. With the advancements in the field of IoT, research works in this area have also taken place, which allows long-distance monitoring and is immune to Electromagnetic (EM) radiations. Self-powered IoT systems have been developed consisting of sensors that derive energy from slight strain variations in the structure [17]. Energy harvesting systems, incorporated in IoT networks in Sunshine Skyway Bridge, consisting of bimorph piezoelectric cantilever beams and used to provide powers to the wireless sensors have also been developed [18]. Integrated distributed sensing systems consisting of IoT, Fiber Bragg grating sensors, Big Data technology and Kafka have been developed to develop long-distance and cost-effective bridge health monitoring systems

[19]. Low-cost systems incorporating multiple regression models based on neural networks have also been developed. Learning automata (LA) is a reinforcement learning model that has been combined with the Confident Information Coverage (CIC) model to considerably increase network lifetime as well as its coverage. For the monitoring of structures, an efficient and effective data management platform called PIERS that is compatible with the IoT method has also been given [20]. Low-cost distributed systems for bridge health monitoring systems with Butterworth filter for efficient signal processing for IoT and cross-correlation for damage detection have also been designed [21]. Fog computing has also been emerging as an efficient approach in terms of power consumption to structural health monitoring [22].

Strive is now made in the direction of ultra-low-cost wireless, energy-efficient, and high-performance SHM system with IoT connectivity and such a system has been installed at Itztal bridge, Germany [23]. Monitoring systems is also one of the key factors in deciding the cost associated with the operation of the system. Instead of cabled monitoring systems, open-source IoT-based frameworks have been developed to support Wireless Structural Health Management Systems thus making the monitoring systems cost-effective [24]. Studies have also been conducted for the development of low-cost and low-power wireless acceleration sensors [25, 26, 27].

3. RECOMMENDATIONS AND DISCUSSIONS

The utilization of drone technology, robots, Virtual Reality (VR), Augmented Reality (AR), and digital twin, an emerging technology, in the field of IoT-enabled bridge health monitoring systems is, however, seems to be absent from the literature and needs rigorous research. Drones integrated with IoT and machine learning can be used to detect faults and irregularities in different parts of the bridges. Augmented reality can also be integrated to IoT systems for visualization of data in real-time and can also be implemented to details of structural faults on the bridge model. Digital Twin is an emerging technology and can also be incorporated in the field of bridge health monitoring to monitor real-time health monitoring even from places far off the bridge location. Therefore, suitable efforts can be made to venture into the field of IoT-enabled bridge health monitoring systems incorporating the aforementioned untouched technologies which could lead to further more efficient monitoring systems.

4. CONCLUSIONS

IoT has recompensed in the area of bridge health monitoring with a whole lot of new opportunities for better monitoring of critical structures. As reviewed in the above sections, wireless real-time monitoring IoT systems have been upgraded with better cloud infrastructure leading to efficient management and better exchange of data. Many improvements have also been made in the integration of modern technologies such as blockchain, Machine Learning (ML) and Artificial Intelligence (AI). The use of neural networks, a kind of machine learning, allows for easy monitoring of bridges via websites or mobile apps. Efforts have been made in the direction of power-efficient and cost-efficient IoT enabled monitoring systems including edge/fog assisted devices, self-powered IoT systems, energy harvesting systems, Fiber Bragg grating sensors, Kafka, Big Data technology Research works, effective data management platforms, and effective data

management platform thus leading to more power and cost-efficient bridge monitoring IoT systems. However, integrating new technologies along with IoT and venturing further into the well-established technologies can lead to unexplored pathways.

5. REFERENCES

[1] Sui, L. Y., Chen, Z. H., Li, W. E. I., & Wang, L. X. (2014). Study on monitoring and safety early warning technology of bridge health based on Internet of things technology. In Applied Mechanics and Materials (Vol. 556, pp. 5994-5998). Trans Tech Publications Ltd.

[2] Sun, S., Ma, Y., Wei, D., Li, M., & Liang, L. (2022). Performance Evaluation of Truss Structure via Cloud Matter Element Fusion. Journal of Performance of Constructed Facilities, 36(2), 04022007.

[3] Paul, P., Dutta, N., Biswas, B. A., Das, M., Biswas, S., Khalid, Z., &Saha, H. N. (2018, November). An internet of things (IoT) based system to analyze real-time collapsing probability of structures. In 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON) (pp. 1070-1075). IEEE.

[4] Tong, X., Ye, Z., Liu, Y., Yang, H., Hou, Y., & Wang, L. (2018, February). The health monitoring system design for bridge based on internet of things. In Transportation Research Congress 2016: Innovations in Transportation Research Infrastructure (pp. 685-696). Reston, VA: American Society of Civil Engineers.

[5] Jeong, S., Hou, R., Lynch, J. P., Sohn, H., & Law, K. H. (2019). A scalable cloud-based cyberinfrastructure platform for bridge monitoring. Structure and Infrastructure Engineering, 15(1), 82-102.

[6] Naraharisetty, V., Talari, V. S., Neridu, S., Kalapatapu, P., &Pasupuleti, V. D. K. (2021, August). Cloud Architecture For IOT Based Bridge Monitoring Applications. In 2021 International Conference on Emerging Techniques in Computational Intelligence (ICETCI) (pp. 39-42). IEEE.

[7] Ahmed, Z., Ali, J. S. M., Rafeeq, M., &Hrairi, M. Application of Machine Learning with Impedance Based Techniques for Structural Health Monitoring of Civil Infrastructure.

[8] Vasavi, S., Sravanthi, G. L., Ram, B. S., &Gokhale, A. A. (2021). Predictive analytics of bridge safety for intelligent transportation system using ensemble model. Materials Today: Proceedings, 45, 5608-5616.

[9] Dang, X. K., Ho, L. A. H., Nguyen, X. P., & Mai, B. L. (2022). Applying artificial intelligence for the application of bridges deterioration detection system. Telkomnika, 20(1).

[10] Manzini, N., Mar, N., Schmidt, F., Bercher, J. F., Orcesi, A., Marchand, P., ...& Thom, C. (2021, August). An Automated Machine Learning-Based Approach for Structural Novelty Detection Based on SHM. In International Conference of the European Association on Quality Control of Bridges and Structures (pp. 1180-1189). Springer, Cham.

[11] Chalapathy, R., &Khoa, N. L. D. (2021, November). Comparison of unsupervised shallow and deep models for structural health monitoring. In Proceedings of the Institution of Civil Engineers-Bridge Engineering (pp. 1-11). Thomas Telford Ltd.

[12] Zhang, L., Zhou, G., Han, Y., Lin, H., & Wu, Y. (2018). Application of Internet of Things technology and convolutional neural network model in bridge crack detection. Ieee Access, 6, 39442-39451.

[13] Mohapatra, A. G., Talukdar, J., Mishra, T. C., Anand, S., Jaiswal, A., Khanna, A., & Gupta, D. (2022). Fiber Bragg grating sensors driven structural health monitoring by using multimedia-enabled iot and big data technology. Multimedia Tools and Applications, 1-21.

[14] Martín, C., Garrido, D., Llopis, L., Rubio, B., &Díaz, M. (2022). Facilitating the monitoring and management of structural health in civil infrastructures with an Edge/Fog/Cloud architecture. Computer Standards & Interfaces, 81, 103600.

[15] Yi, L., Deng, X., Yang, L. T., Wu, H., Wang, M., & Situ, Y. (2020). Reinforcement-Learning-Enabled Partial Confident Information Coverage for IoT-Based Bridge Structural Health Monitoring. IEEE Internet of Things Journal, 8(5), 3108-3119.

[16] Malik, H., Khattak, K. S., Wiqar, T., Khan, Z. H., &Altamimi, A. B. (2019, November). Low cost internet of things platform for structural health monitoring. In 2019 22nd International Multitopic Conference (INMIC) (pp. 1-7). IEEE.

[17] Aono, K., Lajnef, N., Faridazar, F., &Chakrabartty, S. (2016, May). Infrastructural health monitoring using self-powered internet-of-things. In 2016 IEEE international symposium on circuits and systems (ISCAS) (pp. 2058-2061). IEEE.

[18] Farhangdoust, S., Mederos, C., Farkiani, B., Mehrabi, A., Taheri, H., &Younesian, D. (2020, November). A creative vibration energy harvesting system to support a self-powered internet of thing (IoT) network in smart bridge monitoring. In ASME International Mechanical Engineering Congress and Exposition (Vol. 84492, p. V02BT02A041). American Society of Mechanical Engineers.

[19] Mohapatra, A. G., Khanna, A., Gupta, D., Mohanty, M., & de Albuquerque, V. H. C. (2020). An experimental approach to evaluate machine learning models for the estimation of load distribution on suspension bridge using FBG sensors and IoT. Computational Intelligence.

[20] Aihara, K., Takasu, A., Kawakatsu, T., Kinoshita, A., & Adachi, J. (2021). A data management platform for efficient monitoring of infrastructures. In Bridge Maintenance, Safety, Management, Life-Cycle Sustainability and Innovations (pp. 2970-2977). CRC Press.

[21] Mahmud, M. A., Abdelgawad, A., Yelamarthi, K., & Ismail, Y. A. (2017, December). Signal processing techniques for IoT-based structural health monitoring. In 2017 29th International Conference on Microelectronics (ICM) (pp. 1-5). IEEE.

[22] Sinha, D., Doshi, K., &Babu, M. R. (2018). An efficient approach to civil structures health monitoring using fog computing as clusters through 5g network environment. Advances in Systems Science and Applications, 18(3), 123-143.

[23] Smarsly, K., Mthunzi, E. M., Hahn, O., & Planer, J. (2019). Validation of an ultra-low-cost wireless structural health monitoring system for civil infrastructure. In Proceedings of the 12th International Workshop on Structural Health Monitoring (IWSHM). Stanford, CA, USA (Vol. 9, No. 10, p. 2019).

[24] MALIK, H., & ZATAR, W. (2019). An IoT Enabled Framework to Support Structural Health Monitoring Applications. Structural Health Monitoring 2019.

[25] Tong, X., Yang, H., Wang, L., & Miao, Y. (2019). The development and field evaluation of an IoT system of low-power vibration for bridge health monitoring. Sensors, 19(5), 1222.

[26] Hou, S., & Wu, G. (2019). A low-cost IoT-based wireless sensor system for bridge displacement monitoring. Smart Materials and Structures, 28(8), 085047.

[27] Malik, H., &Zatar, W. (2020, July). A real-time and low-cost flash flood monitoring system to support transportation infrastructure. In 2020 IEEE 44th Annual Computers, Software, and Applications Conference (COMPSAC) (pp. 1111-1112). IEEE.

[28] Malik, Praveen, Singh, R., Gehlot, A., Akram, S. V., & Das, P. K. (2022). Village 4.0: Digitalization of Village with Smart Internet of Things Technologies. Computers & Industrial Engineering, 107938.

[29] Bagwari, S., Gehlot, A., Singh, R., Priyadarshi, N., & Khan, B. (2022). Low-Cost Sensor-Based and LoRaWAN Opportunities for Landslide Monitoring Systems on IoT Platform: A Review. IEEE Access, 10, 7107–7127. https://doi.org/10.1109/ACCESS.2021.3137841

[30] Malik, P, Gehlot, A., Singh, R., Gupta, L. R., & Thakur, A. K. (2022). A Review on ANN Based Model for Solar Radiation and Wind Speed Prediction with Real-Time Data. Archives of Computational Methods in Engineering. https://doi.org/10.1007/s11831-021-09687-3

[31] Singh, R., Gehlot, A., Rashid, M., Saxena, R., Akram, S. V, Alshamrani, S. S., & Alghamdi, A. S. (2021). Cloud server and internet of things assisted system for stress monitoring. Electronics (Switzerland), 10(24). https://doi.org/10.3390/electronics10243133

[32] Karunanidy, D., Ramalingam, S., Dumka, A., Singh, R., Rashid, M., Gehlot, A., ... & AlGhamdi, A. S. (2022). JMA: Nature-Inspired Java Macaque Algorithm for Optimization Problem. Mathematics, 10(5), 688.

[33] Sharma, V. K., Singh, R., Gehlot, A., Buddhi, D., Braccio, S., Priyadarshi, N., & Khan, B. (2022). Imperative Role of Photovoltaic and Concentrating Solar Power Technologies towards Renewable Energy Generation. International Journal of Photoenergy, 2022.

[34] Joshi, K., Diwakar, M., Joshi, N. K., & Lamba, S. (2021). A concise review on latest methods of image fusion. Recent Advances in Computer Science and Communications, 14(7), 2046–2056. https://doi.org/10.2174/2213275912666200214113414

A Descriptive Statistical analysis of the effectiveness of Lockdown in India

Manoj Kumar Sharma[1], Vineet Kishore Srivastava[2] , Shalu Chaudhary [3], Sanjeev Kumar Shah[4], Satya Jeet Singh[5] ,Mukesh Kumar[6]

[1,2,3]Department of Mathematics, Uttaranchal University, Dehradun, Uttarakhand-INDIA

[4]Department of ECE, Uttaranchal University, Dehradun, Uttarakhand-INDIA

[5,6]Department of Mathematics, Graphic Era Deemed to be University, Dehradun, Uttarakhand-INDIA

Abstract

Like all global worldwide, India has been afflicted by the increase of COVID-19, which began out in a Chinese town in December 2019. The spreading effect may be estimated in distinct stages, loss of life because India has previously noted more than four lakhs covid cases. Also, the influence of COVID-19 has impacted on economy, training systems, schooling, banking, and lots of different sectors. Our purpose is to observe and examine the impact of lockdown days on the spread of the coronavirus in India country. We have gathered India's COVID-19 cases and lockdown days records that carried out the before, during, and after lockdowns among precise dates. Using Statistical Measures, inferences are drawn primarily based totally on COVID records on day-by-day new and daily death cases.

Keywords COVID-19, lockdown, data analysis, correlation.

1. INTRODUCTION

COVID-19 is not the first severe pandemic; pandemics have transpired on a regular basis throughout human history. In December 2019, a disease similar to pneumonia hit Wuhan, China. According to the investigation, the covid cases emerged a new type of Corona Virus that had not been adequately described because it first arrived in Dec 2019; this variant is known as coronavirus 2019 or COVID-19. Viruses are transmitted from animals to humans and transmitted from human to human. The molecular mechanism of human-to-human transmission of COVID-19 is still unknown, but the principles of respiratory disease transmission are similar in the prevailing Noymer and Garenne [1]. Respiratory illness is repast by splash sprinkling. In this form of spread, sick men or women are exposed to this microorganism by the people around them through coughing and sneezing. The outbreak of COVID-19 is spreading very rapidly daily. Since the virus has actively infected more than 4 million people, COVID 19 restrictions are being enforced in almost every region where Oxford et al. [2]. Adherence to hygiene rules is essential to containing the coronavirus pandemic and prevent infection. The most important thing is regular hand washing. As a result, the spread of this virus is engrossed in washing the arms and slows down in societies that are aware of simple hygiene rules. Our aim is to study and evaluate the impact of infestation on the coronavirus pandemic in India. Many reputable institutions participate in the "household ban." Scientists advise that the COVID-19 virus can attain any age

institution speedily. Alternations in influenza infection rates were spread speedily all over the world. So, there are some steps taken by the Indian government like short-term Lockdown; Stay at home; Hand washes regularly; Wear the mask and maintain a distance of at least one meter. The flu pandemic spread widely in 1918. (January 1918 – December 1920) Five hundred million people were infected in that pandemic, and 50 to 100 million people of them have killed Collier [3]. This is the deadliest natural disaster [4] in human history. The covid-19 [5] killed in every corner of the globe. Atlan [7] investigated the study assessed lockdown effects on psychological, environmental, and economy. Das and Kumari [8] studied the statistical analysis of COVID cases in India. The effect of lockdown policies on the basic multiplicative quantity of coronavirus cases in Saudi Arabia has been investigated by S. Ahmad Alajlan et al. [9], and Saleh Alrashed et al. [10] presented the impact of lockdowns on the spread of COVID-19 in Saudi Arabia. We have taken the data for this investigation from [4-7].

In the proposed study, our purpose is to observe and examine the impact of lockdown days spread of the coronavirus in India. We have collected the COVID-19 cases data and lockdown days records for India that carried out the before, during, and after lockdowns among particular dates. Using [11] Statistical Measures, inferences are drawn primarily based totally on COVID records on daily new and daily death cases.

2. MATERIALS AND METHODS

Corona cases data of country India considered and collected from www.worldometer.com [6]. The descriptive statistical analysis analyzed the relations between daily new cases and daily death cases of COVID-19 by India during the first and second lockdown. The relationship of the lockdown(s) on the number of Corona cases is calculated as unrestricted (normal distribution). The distribution is based on (Shapiro-Wilk Test). This distribution is stated mathematically in Equation (1).

$$Z = \frac{\left(\sum_{i=1}^{n} a_i x_i\right)^2}{\left(\sum_{i=1}^{n} x_i - \bar{x}\right)^2} \tag{1}$$

$$a_i = \frac{m_v^T V^{-1}}{N}, m_v = \left(m_{v_1}, m_{v_2}, \ldots, m_{v_n}\right)^T \tag{2}$$

$$N = \|V^{-1} m_v\| \tag{3}$$

where x_i is sample, a_i is coefficient sample, N is vector norm, V is the covariance matrix of those normal order statistics and $\bar{x}$ is sample mean.

2.1. RESULTS AND DISCUSSIONS

The following research is based on Covid-19 new cases and death rates before, during, and after the first lockdown in India. With the help of table 1 & table 2, we will analyze the impact of lockdown on daily [12] new and death cases. There were few assumptions regarding p-value and t-test; before lockdown, it was assumed that the p-value would lie between 0 and 1, and the value of the t-test was assumed to be [-2.0262: 2.0262]. The conclusions of the descriptive analysis were queried as 95 percent confidence intervals before, during, and after lockdown days. The statistical examination was two-sided, and a value of the probability of success is less than 0.05 measured for model and parameter statistically significant. The data set used is not suitable for normal distribution according to Shapiro-Wilk [p-values (1st lockdown) of Shapiro-Wilk test are $6.128e^{-7}$ (before lockdown), 0 (during lockdown), and 0.00001601 (after lockdown)] normality

tests. The data set used is not suitable for normal distribution (2nd lockdown) according to Shapiro-Wilk [p-values (2nd lockdown) of Shapiro-Wilk test are 0 (before lockdown), $2.844e^{-7}$ (during lockdown), and $2.22e^{-16}$ (after lockdown)] normality tests.

Correlation analysis is performed between daily new cases and daily deaths in the first lockdowns period of the COVID-19. The correlation value lies between -1 and +1. The correlation values of the factors indicate that they have a negative relationship as they approach -1 a positive relationship as they approach +1. And no relationship if the correlation [16] coefficient is equal to zero. The statistical process was performed by transforming the COVID 19 data.

In table 1we taken the data of Covid-19 cases came in India before lockdown (15-02-2020 to 24-03-2020), during lockdown (25-03-2020 to 31-05-2020) and after lockdown (01-06-2020 to 30-07-2020). In which we have studies the statistical parameters on new cases came per day and death per day before lockdown, during lockdown and after lockdown and impact of lockdown in respect of COVID-19 cases per day and death per day.

Table 1 shows that the average daily new cases were around 14 before lockdown. The average daily death rate was around 0.28 during the lockdown. The number of cases per day increased up to 2795, and the number of death cases also increased. Further, after lockdown number of cases still increased, and the number of death cases also increased but critically examining the facts, it is brought that the correlation coefficient (r) decreased after lockdown, which clearly shows that lockdown had an impact. T-test value before the lockdown was 6.00, and the same was increased during the lockdown and decreased after lockdown.

Therefore, table 1 indicates that lockdown could not make any effect in respect of numbers of cases per day and death per day but it had impact on daily death rate as correlation coefficient (r) decreased after lockdown. The T-test statistic T equal to 6.0078, which is not in the 95% region of acceptance, hence hypothesis is rejected.
In table 2 we have taken the data of Covid-19 cases came in India before second lockdown (01-02-2021 to 04-04-2021), during second lockdown (05-04-2021 to 15-06-2021) and after second lockdown (16-06-2021 to 16-08-2021). In which we have studies on new cases came
per day and death per day before lockdown, during lockdown and after lockdown and impact of lockdown in
respect of COVID-19 cases per day and death per day.

Table 2 shows that average daily new cases were round 29000 before lockdown and average daily death rate was around 172, and during lockdown number of cases per day increased up to 236726 and number of death cases also increased. Further, after lockdown number of cases still decreased and number of death cases also decreased but critically examining the facts it is brought that the correlation coefficient (r) increased after lockdown, which clearly shows that lockdown had no impact. t- test value before lockdown was 21.6146 and the same was decreased during lockdown and increased after lockdown.
Therefore, table2 indicates that lockdown could not make any effect in respect of numbers of cases per day and death per day but it had impact on daily death rate as correlation coefficient (r) decreased after lockdown. The test statistic T equals 21.6146, which is not in the 95% region of acceptance, hence hypothesis is rejected.

Table 1: Statistical Analysis of First Lockdown Period Data

Parameters	Before Lockdown	During Lockdown	After Lockdown
Pearson correlation coefficient (r) Between Daily New Cases and Daily Death	0.7027	0.9557	0.5258
P-value	6.128e-7	0	0.00001601
Covariance	12.2058	148912.7882	1939829.468
Sample size (n)	39 Days	68 Days	60 Days
Mean of Daily New Cases	13.7435897	2795.19	24145.68
Mean of Daily Deaths	0.28	80.01	512
t-test	6.0078	26.3879	4.7077

Table 2: Statistical Analysis Second Lockdown Period Data

Parameters	Before Lockdown	During Lockdown	After Lockdown
Pearson correlation coefficient (r) Between Daily New Cases and Daily Death	0.9405	0.5617	0.8234
P-value	0	2.844e-7	2.22e-16
Covariance	2882194.984	85976080.22	2446394.32
Sample size (n)	63	72	62
Mean of Daily New Cases	29036	236726	42219
Mean of Daily Deaths	172	3051	765
t-test	21.6146	5.6803	11.2405

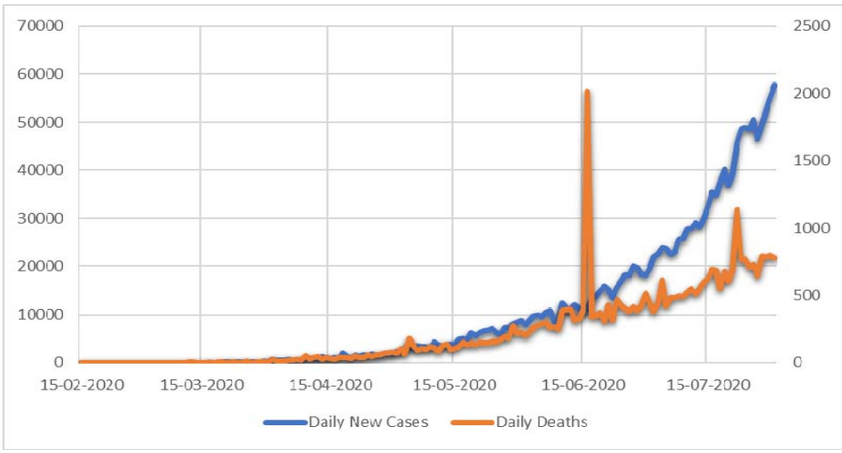


Figure 1 Daily Covid Cases and Daily Deaths in first lockdown

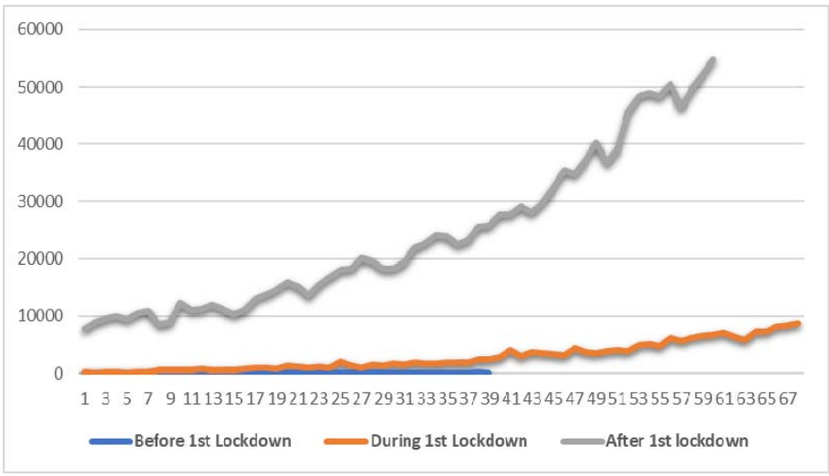


Figure 2 Daily Covid Cases Before, during and after First Lockdown

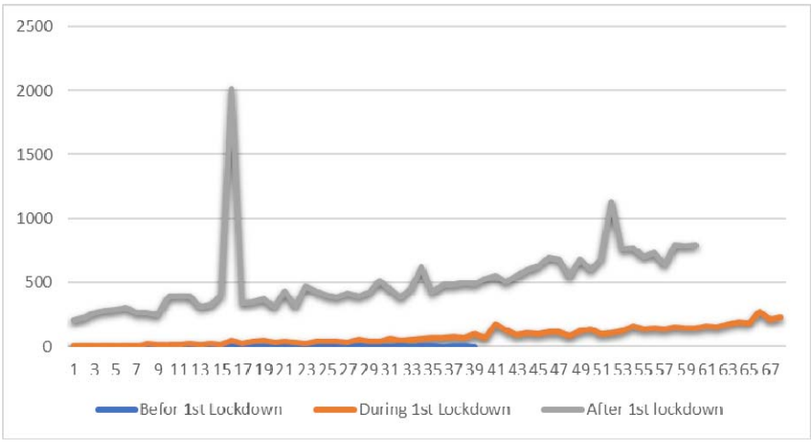


Figure 3 Daily Covid Deaths Before, during and after first Lockdown

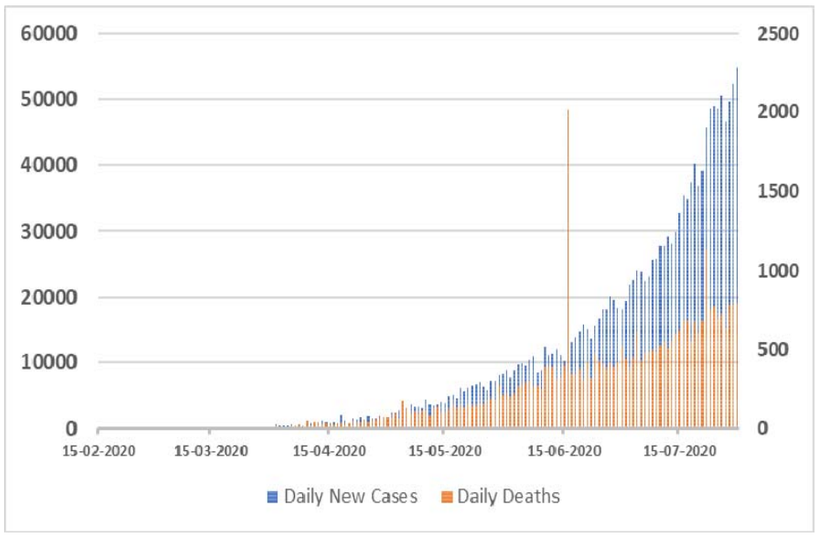


Figure 4 Daily new cases and daily deaths in Ist Lockdown period

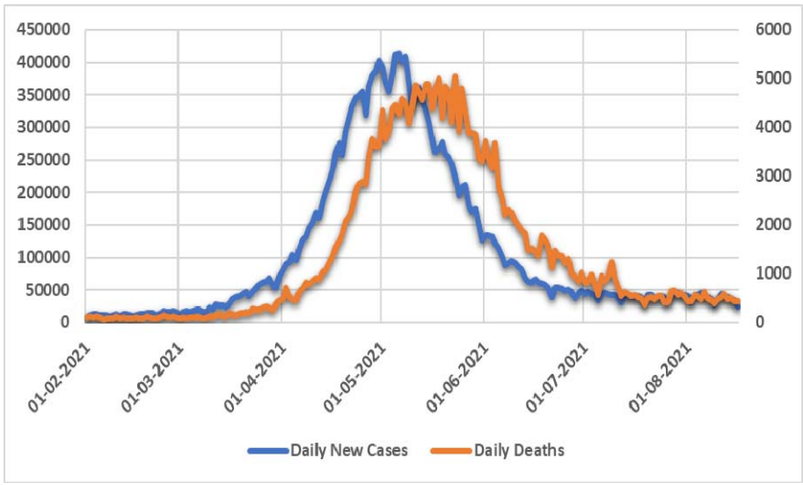


Figure 5 Daily Covid Cases and Daily Deaths in 2nd lockdown

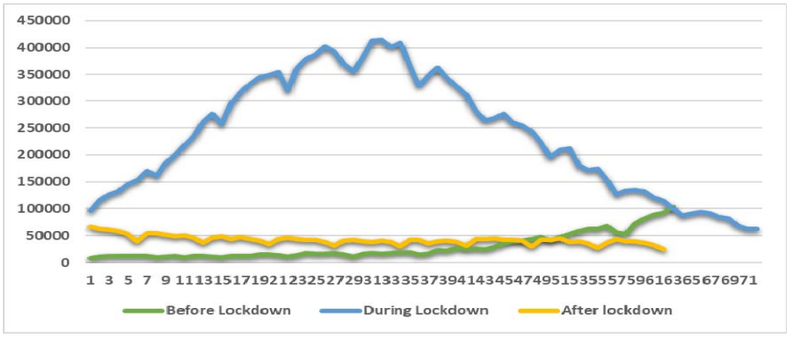


Figure 6 Daily Covid Cases Before, during and after 2nd Lockdown

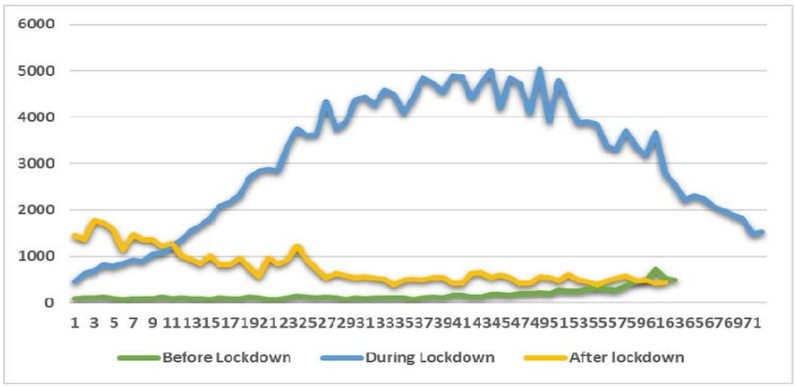


Figure 7 Daily Covid Deaths Before, during and after 2nd Lockdown

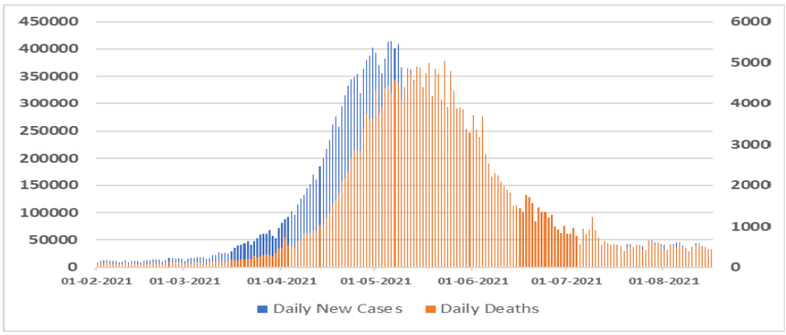


Figure 8 Daily new cases and daily deaths in 2nd Lockdown period

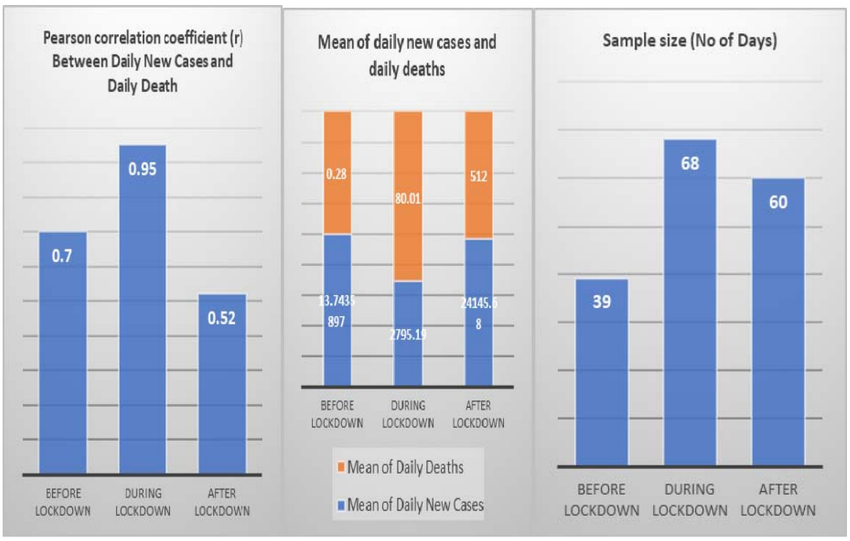


Figure 9 Descriptive statistical analysis of First Lockdown Period

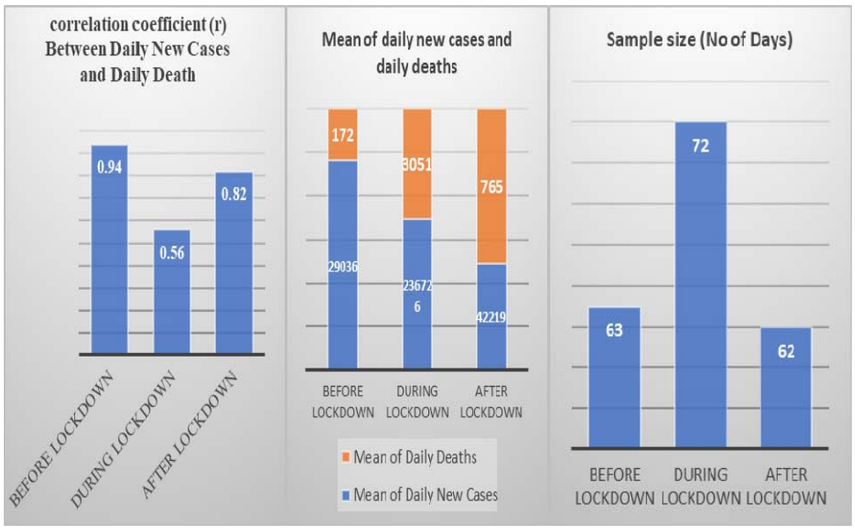


Figure 10 Descriptive statistical analysis of second Lockdown Period

3. CONCLUSION

This paper analyzes the existing data of the Indian epidemic situation. We have considered the data before, during, and after the first Indian lockdown period from 25-03-2020 to 31-05-2020 and the second lockdown period from 05-04-2021 to 15-06-2021 analyzed the data between daily new cases and daily deaths in that period. We found that the correlation coefficient (r) between daily new cases and daily death cases after lockdown has decreased in the first lockdown period, showing that the lockdown days affected the spreading of the virus. But the mean daily new cases and mean daily deaths were increased. The daily growth of COVID-19 cases and deaths has implied a global lockdown, quarantine, and limitations. While in the second lockdown period, the correlation coefficient between daily new cases and daily deaths decreased during the lockdown period, and daily deaths also decreased. This study provides a preliminary indication that a lockdown that can be used to control the COVID-19 pandemic by the Indian government was much more effective. As a result, governments' use of lockdown influences psychology, the environment, the economy, and the spreading COVID-19.

4. REFERENCES

[1] A. Noymer, M. Garenne, 'The 1918 influenza epidemic's effects on sex differentials in mortality in the United States. Popul Dev Rev.pp. 565-81,2000;26(3).

[2] J.S. Oxford, A.Sefton, R.Jackson, W.Innes, R.S.Daniels,NPAS Johnson . World War I may have allowed the emergence of "Spanish" influenza. Lancet Infect Dis, pp 111–4., 2002;2 (2):

[3] R. Collier, The Plague of the Spanish Lady. The Influenza Pandemic of 1918–19. Atheneum, pp-56-64 1974.

[4] H. Phillips,The recent wave of 'Spanish' flu historiography. Soc History Med, pp-789–808.2014, 27(4)

[5] Novel Coronavirus Visual Dashboard operated by the Johns Hopkins University Center for Systems Science and Engineering (JHU CSSE). Also, Supported by ESRI Living.

[6] Worldometer, Coronavirus,https://www.worldometers.info/coronavirus/country/turkey/, (2020).

[7] A. Atalan, Is the lockdown important to prevent the COVID-19 pandemic? Effects on psychology, environment and economy-perspective, Annals of Medicine and Surgery, pp 38-42, 2020, 56.

[8] Das and Kumari, Statistical Analysis of COVID cases in India, J. Phys.: Conf. Ser. 1797 012006, 2021.

[9] S. Ahmad Alajlan, N.K. Alhusseini, S. Mohammed Basheeruddin Asdaq et al., The impact of lockdown strategies on the basic reproductive number of coronavirus (COVID-19) cases in Saudi Arabia, Saudi Journal of Biological Sciences, pp 4926-4930, 2021. 28(9)

[10] S. Alrashed, et al., Impact of lockdowns on the spread of COVID-19 in Saudi Arabia, Informatics in Medicine Unlocked, 20,100420, 2020.

[11] Sources of Covid-19, 2020 data: https://www.mohfw.gov.in/; https://www.who.int/

[12] Malik, Praveen, Singh, R., Gehlot, A., Akram, S. V., & Das, P. K. (2022). Village 4.0: Digitalization of Village with Smart Internet of Things Technologies. Computers & Industrial Engineering, 107938.

[13] Bagwari, S., Gehlot, A., Singh, R., Priyadarshi, N., & Khan, B. (2022). Low-Cost Sensor-Based and LoRaWAN Opportunities for Landslide Monitoring Systems on IoT Platform: A Review. IEEE Access, 10, 7107–7127. https://doi.org/10.1109/ACCESS.2021.3137841

[14] Malik, P, Gehlot, A., Singh, R., Gupta, L. R., & Thakur, A. K. (2022). A Review on ANN Based Model for Solar Radiation and Wind Speed Prediction with Real-Time Data. Archives of Computational Methods in Engineering. https://doi.org/10.1007/s11831-021-09687-3

[15] Singh, R., Gehlot, A., Rashid, M., Saxena, R., Akram, S. V, Alshamrani, S. S., & Alghamdi, A. S. (2021). Cloud server and internet of things assisted system for stress monitoring. Electronics (Switzerland), 10(24). https://doi.org/10.3390/electronics10243133

[16] Karunanidy, D., Ramalingam, S., Dumka, A., Singh, R., Rashid, M., Gehlot, A., ... & AlGhamdi, A. S. (2022). JMA: Nature-Inspired Java Macaque Algorithm for Optimization Problem. Mathematics, 10(5), 688.

[17] Sharma, V. K., Singh, R., Gehlot, A., Buddhi, D., Braccio, S., Priyadarshi, N., & Khan, B. (2022). Imperative Role of Photovoltaic and Concentrating Solar Power Technologies towards Renewable Energy Generation. International Journal of Photoenergy, 2022.

A Study of Various Stages In Human Computer Interaction

Vineet Kishore Srivastava[1], Rakesh Chandra Bhadula[2], Ajay Sharma[3], Sanjeev Kumar Shah[4], V.N. Kala [5]& Satya Jeet Singh [6]

[1]*Department of Mathematics, Uttaranchal, University, Dehradun*

2Department of Mathematics, Graphic Era Hill, University, Dehradun

[3]*School of Management, Graphic Era Hill, University, Dehradun*

[4]*Department of ECE, Uttaranchal, University, Dehradun*

[5]*Department of Applied Science, GBPC, Pauri Grahwal*

[6]*Department of Mathematics, Graphic Era, University, Dehradun*

Abstract.

Aim of this study is to recognize more deeply how computer and human interact together. This is to study the various of actions, which is performed to communicate with computers and humans. In each stage of actions, it is concluded that goal cannot be achieved without performing sequence of functions and finally comparison can be made to check the intentions and user goals. Designing of HCI is also studied in our study and it is analysed that five stage of designing play's important role to make Human computer interactions. After applying these five stages of designing human can easily communicate with computer.

Keywords. DataHuman computer interaction, Design, User Interface, Communication

1. INTRODUCTION

Human computer interaction came into existence in the 1980s and computers were designed in such a way that humans can use complex technology in a safe, comfortable, efficient and enjoyable manner. Many researchers from computer and psychological background who may have interest in this area are studying the behaviour as well as designers and people from industry and commercial background, HCI is a saleable product for them. Researchers are developing the HCI in such a way the human can interact with the computer in a safe and comfortable mode.

Card et al. (1983) studied the GOMS (Goal Operators Method System) model for HCI, in this model one has to identify the goal then the user has to operate the metal activity to complete the goal. Gregor and Newell (2001) presented the HCI model to study organizational strategies for some old adults. Michel et al. (2014) studied a new method for skin deduction in colour image. In this method they used texture based discriminative skin presence feature, which is very helpful for Human computer interaction. Kawulok (2012) presented discriminative skin presence features, which is based on discriminative textural features.

Card et al. (2018) investigated the complexity between the human and computer interaction and described the kinds of actions those arises when more than two artifacts interact. They described communication between computers and humans with specific models and characteristics. Understanding the communication between human and computer involves connectivity, assessment and applications, this makes easy and convenient computer user interface. The term human computer interaction was invented a few years ago, but industry has recognised its need and importance in beginning of 19th century and working to improve this area. Basic purpose of this chapter to explore the critical actions in human computer interaction. Now a days computer interaction knowledge is vital because in every field of human life, computer communication is needed (Sawer &Ting , 2014).

Human computer interaction can be described in two ways, first one is direct communication and second one is indirect communications, direct communication includes various types of inputs like touch, sound, pointer etc. which received by input devices, control unit process this communication. While indirect communication involving the background process in human computer interaction. (Sunder et al., 2016).Tan &Nijholt (2010), suggested that the computer technology is ranging from small computer desktops to large one, all these involve input system, processing control system and output system. They described that the group of human beings who are confronted with such a problem can be encountered with computer technology.

Klumpp et al. (2019) conferred different ways to stop an artificial partition in HCI by social sustainability and logistics creativity. In logistics and supply chain process HCI plays an important role to resolve the problems. Zhang, P., and Li, N. (2004) studied MIS (Management Information Systems) as a subfield of Human Computer Interaction and revealed the relations among human interaction with information technology. He recognized the importance of HCL especially in contest of business administration, managerial and organization.

2. HUMAN COMPUTER INTERACTION WITH STAGES OF ACTIONS

***Stages of Actions*:** In Human computer interaction any task can be obtained by creating goals and goal can be achieve by applying sequence of various actions. These stages can be divided in the following seven categories.

(i) Forming the Goals (ii) Intentions (iii) Specify the actions (iv) Execution (v) Perceiving the states of the world (vi) Interpreting the state of the world (vii) Comparison to intentions and user's goals.

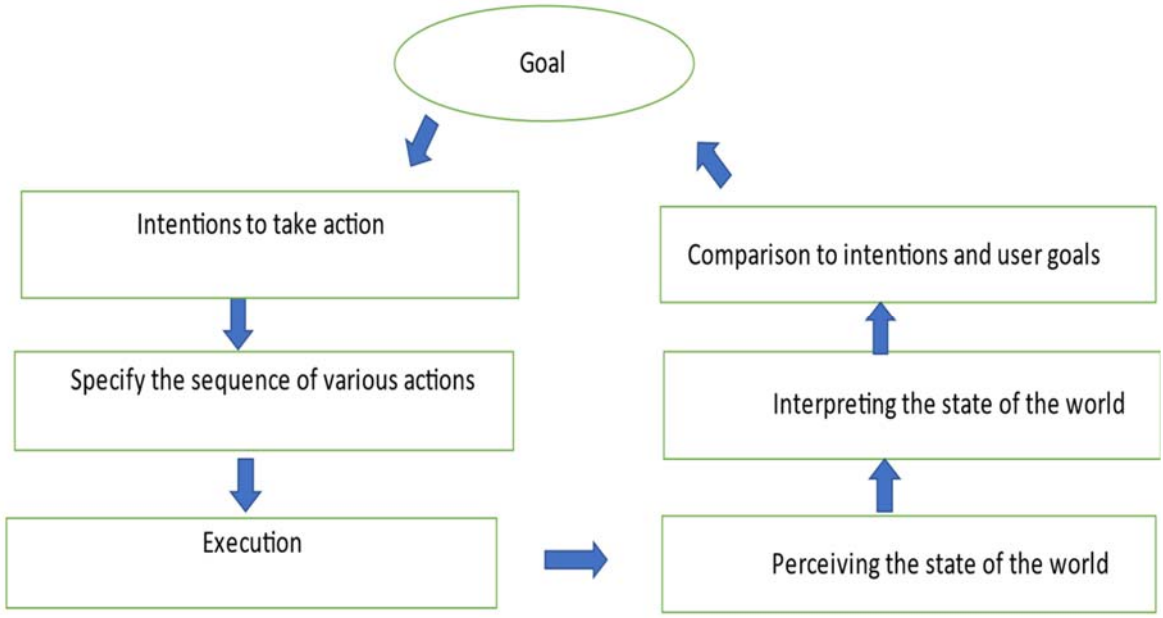


Figure 1: Human Computer Interaction with Stages of Actions

3. HUMAN COMPUTER INTERACTION AND DESIGNING

This focus upon evaluation displays a problem into modelling. Designing components plays an important role to achieve the goal. Human computer interaction can be made in such a way that every human can interact with computer. Landauer (1995) conducted a study through econometric analysis of computer industry, he suggested that design flaws have to be identified in beginning designing user interface and interaction because this makes work easy and cost effective.

Basically, there are five stages of designing (i) Visibility (ii) Affordance (iii) Legible mapping (iv) Feedback (v) Constraints.

(i) ***Visibility:*** In Human computer interaction visibility means the mapping between indeed actions and actual operations. These sequences of operations are used to communicate with computer.

(ii) ***Affordance:*** Affordance means actual properties of the thing, which can be determine just how they could possibly be used in the Human Computer

interaction. Thus, affordance providing leads to operate the computer i.e., machine

(iii) ***Mapping*:** It is the relationship between the control and the results. These results can be obtained by doing interaction between the computer.

(iv) ***Feedback:*** - This is important process of every interaction; it sends back to user. What communication has user done with the computer.

(v) ***Constraints***: This part limit the user action to the system. This part reduced the error by performing the communication between human and computer.

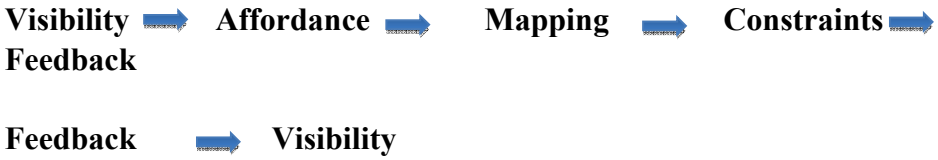


Figure 2: Human Computer Interaction and Designing

4. CONCLUSION:

In this chapter main focus remained to establish series of actions in the communication between human and computer. Seven stages of action are studied and it is concluded that perfect completion of actions in each stage are as important as to achieve the goal. Five stages of designing i.e., visibility, affordance, mapping, feedback and constraints are as much important as the feedback process, in such process interaction can be send back to user, what communication has human done with computer. In computer-based learning i.e., human computer interaction, computer experts can design a system that can be user friendly. It is essential that communication should be executed by the user according to his or her requirement.

5. REFERENCES

[1] Card S.K, Moran T.P. Newell A. The Psychology of Human Computer Interaction. Eribaum. Hillsdale. NJ, 1983.

[2] Card, S.K. The psychology of human-computer interaction. CRC Press, 2018

[3] Gregor P and Newell, A. F.. Deigning for dynamic diversity-making accessible interfaces for older people. In wUAUC01-2001 ECNSF Workshop on Universal Accessibility Ubiquitous Computing: Providing for the Elderly, pp.90-92, 22-25 May 2001, Portugal.

[4] Michal K., Jolanta K. and Jkub N. Spatial Based Skin Deduction Using Discriminating Skin Presence Features. Pattern Recognition Letter, Vol 41, pp. 3-13, 2014.

[5] Kawulok, M., Kawulok, J., Smolka. Discriminative Textural Features for Image and Video Colorization. IEICE tran. Inf. Syst. 95- D (7), pp. 1722-1730, 2012.

[6] Sawers, A., & Ting, L.H.. Perspectives on human-human sensorimotor interactions forthe design of rehabilitation robots.Journal of Neuro engineering and Rehabilitation,11(1), pp. 1 – 13, 2014.

[7] Sundar, S.S., Bellur, S., Oh, J., Jia, H., & Kim, H.S. Theoretical importance of contingency in human-computer interaction: Effects of message interactivity on user engagement. Communication Researc,43 (5), pp. 595 – 625, 2016.

[8] Tan, D., &Nijholt, A. Brain-computer interfaces and human-computer interaction. In Brain-Compute Interfaces, pp. 3 – 19, 2010.

[9] Klumpp, M., Hesenius, M., Meyer, O., Ruiner, C., & Gruhn, V. Production logistics and human-computer interaction—State-of-the-art, challenges and requirements for the future. The International Journal of Advanced Manufacturing Technology, 105(9), pp. 3691–3709, 2019.

[10] Zhang, P., & Li, N. An assessment of human–computer interaction research in management information systems: topics and methods. Computers in Human Behaviour, 20(2), pp.125–147, 2004.

[11] Landauer T. The Trouble with Computer. MIT Press Cambridge, MA, 1995.

[12] Malik, Praveen, Singh, R., Gehlot, A., Akram, S. V., & Das, P. K. (2022). Village 4.0: Digitalization of Village with Smart Internet of Things Technologies. Computers & Industrial Engineering, 107938.

[13] Bagwari, S., Gehlot, A., Singh, R., Priyadarshi, N., & Khan, B. (2022). Low-Cost Sensor-Based and LoRaWAN Opportunities for Landslide Monitoring Systems on IoT Platform: A Review. IEEE Access, 10, 7107–7127. https://doi.org/10.1109/ACCESS.2021.3137841

[14] Malik, P, Gehlot, A., Singh, R., Gupta, L. R., & Thakur, A. K. (2022). A Review on ANN Based Model for Solar Radiation and Wind Speed Prediction with Real-Time Data. Archives of Computational Methods in Engineering. https://doi.org/10.1007/s11831-021-09687-3

[15] Singh, R., Gehlot, A., Rashid, M., Saxena, R., Akram, S. V, Alshamrani, S. S., & Alghamdi, A. S. (2021). Cloud server and internet of things assisted system for stress monitoring. Electronics (Switzerland), 10(24). https://doi.org/10.3390/electronics10243133

[16] Karunanidy, D., Ramalingam, S., Dumka, A., Singh, R., Rashid, M., Gehlot, A., ... & AlGhamdi, A. S. (2022). JMA: Nature-Inspired Java Macaque Algorithm for Optimization Problem. Mathematics, 10(5), 688.

[17] Sharma, V. K., Singh, R., Gehlot, A., Buddhi, D., Braccio, S., Priyadarshi, N., & Khan, B. (2022). Imperative Role of Photovoltaic and Concentrating Solar Power Technologies towards Renewable Energy Generation. International Journal of Photoenergy, 2022.

A Machine Learning-Based Method to Predict of Drug–Target Association Based On Multiple Feature Information Selection and Reduction Techniques

[1]Deepak Srivastava, [2] Dr.Pramod Kumar,[3] Dr. Sunil Ghildiyal

[1]Assistant Professor, Department of Computer & Information Sciences, Himalayan School of Science & Technology, Swami Rama Himalayan University, Dehradun, India,

[2]Professor, Krishna Engineering College, Ghaziabad, U. P., India

[3]Associate Professor, UIT, Uttaranchal University, Dehradun, India

ABSTRACT

Identifying drug–target (protein) interactions is critical for research and development of innovative drugs, providing a significant benefit to pharmaceutical businesses and patients. However, predicting Drug Target Indications by clinical trial procedures is typically costly and time consuming. As a result, many machine learning-based algorithms have been created for this goal, yet significant unknown interactions remain. Additionally, feature selection and reduction concerns are a key barrier in drug-target datasets, since they might affect classifier performance if not handled well. This study offered a unique approach for predicting drug–target interactions. To begin, the amino acid composition (AAC), dipeptide composition (DC), and tripeptide composition (TC); and drug SMILES substructure fingerings are used to extract the protein sequence's feature vectors. PCA (Principal Component Analysis) is used to eliminate superfluous and redundant characteristics in order to get the most optimum features. Finally, balanced and optimum features are supplied to SVM with RBF kernel function in order to detect Drug Target indications, and the proposed approach's prediction capacity is evaluated using the 10-fold CV validation test method. The prediction findings suggest that the proposed model outperforms other current approaches in predicting Drug Target interaction.

KEYWORDS: Drug Repurposing, Support Vector Machine (Kernel), Principal Component Analysis, Feature selection and reduction, Classification

I. INTRODUCTION

Predicting novel drug–target association is a critical stage in the pipelines of drug discovery and design [1–3]. Drug repurposing is a rising area in pharmaceutical science, with an emphasis on uncovering previously undiscovered interactions between current drugs and novel target proteins. The advancement of the entire genes and the expansion of the molecular biology project provide valuable information for predicting novel therapeutic targets. Numerous attempts have been made in recent years to identify new treatments, but relatively few have been approved by the Food and Drug Administration (FDA) and reached patients, while a large number of pharmaceuticals have been rejected in clinical trials due to unacceptable toxicity. DTI wet-lab investigations are often time-consuming, labor-intensive, and expensive; as a result, such failures are difficult to accept and result in significant financial loss. As a result, researchers are particularly motivated to develop machine learning (ML)-based algorithms for detecting Drug Target Association [3], which may successfully narrow the search space of drug–target possibilities to be evaluated in wet-lab trials, therefore reducing work and expense. Recently, machine learning-based computational approaches have grown increasingly advantageous due to the vast amount of heterogeneous pharmacological and protein data.

chemogenomic approaches often make advantage of the genomic and chemical information associated with target proteins and medicines. As a result, chemogenomics techniques are becoming increasingly common for identifying Drug Target associations. The chemogenomic model's prediction challenge might be addressed utilising powerful machine learning methods [12].

Numerous machine learning classifiers, including deep learning, SVM, fuzzy logic, and closest neighbour, have been successfully used to various sorts of prediction tasks. Whereas feature-based algorithms use input vectors of drug chemistry and protein sequence features and display the class label as a binary value (1 or 0).

1.1 Feature Selection and Reduction

Feature Selection in Computer Aided Diagnosis (CAD) is a difficult module to learn when it comes to classifying Drug Target indications. This is mainly due to the increased number of features to be analyzed with high desirable accuracy. When the feature sets are huge, or the input dataset is voluminous, the classification becomes a highly time-consuming task. Feature selection generally deals with selecting the most appropriate useful features and minimizes the redundancy in improving the performance of classification subsystem. Thus, the optimal feature selection will effectively increase the accuracy, reduce the time complexity, and improve the performance of any CAD system. [11].

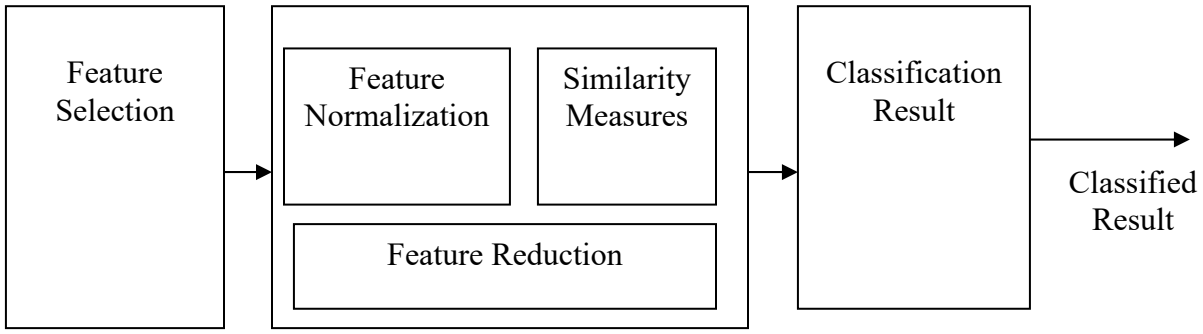


Figure 1: Depict Feature Selection and reduction

In this paper, a combined feature selection algorithm that uses multi Zero Mean and Unit Variance Normalization with Correlation Distance Method is presented and analyzed. The single-objective feature selection algorithms provided only a single bag of optimal solution. This approach circumvented the limitations of typical single goal algorithms by generating a collection of optimum solutions that trade off distinct objectives. The multi objective approach ensured that the minimal features with high impact on classification were selected and it achieved improved accuracy with lesser.

The multi-objective feature selection algorithm's system architecture is divided into three phases: feature selection, feature reduction, and classification. The selection phase's input was produced by features collected from drugs and targets. Feature representation and neighborhood formation were carried out during the selection phase. [20] During the feature reduction step, redundant features and those with a negligible effect on categorization were filtered. The classification step was used to classify the massive amount of input data into distinct categories. Cross validation, data trimming, nearest-neighbor computation, and normal distribution model were all performed throughout the classification phase. To classify input data into training and testing sets, the cross validation approach was applied. To compute the class probability, data reduction was performed. To compare the real input features to the training set of features, the nearest class algorithm was utilized. Finally, the categorized results were determined using a normal distribution model.

II. BACKGROUND/ LITERATURE SURVEY

The PCA-based technique outlined in [11] was used to perform a survey on drug repurposing, which has gotten a lot of interest in recent years. It is capable of providing effective solution for applications with certain limits after several years of study.

Orawan et al.[17] built the system by collecting Fuzzy Co-occurrence Matrix and fractal dimension characteristics, and then used PCA to reduce the system's dimensionality. Multi-class SVM is used to

classify cancer patients with a 91.7 percent correct classification rate, 93 percent sensitivity, and 91 percent specificity.

Zhu et al. [14] used principle component analysis (PCA) and the K-means clustering technique to construct an improved logistic regression model for diabetes prediction. PCA is used in this proposal to translate diabetic data to a lower dimension. Integration of PCA enhanced the accuracy of K-means clustering and logistic regression, as demonstrated by simulation results.

III. DATA COLLECTION AND METHODOLOGY

We created a prediction model that employs Support Vector Machine with RBF kernel nonlinear classification techniques to determine the potential of particular drug-target associations. However, prior research indicates that the dataset is classifiable using linear models. Feature selection and reduction is done into two parts that is feature normalization and feature similarity. External validation and tenfold cross-validation were performed to determine the accuracy of each prediction model. Below steps are involved to determine Drug Target association. Methodology is explained in below steps: [4]

Step 1: Data Collection

I acquired data from Drug Bank in order to utilize it for Drug Repurposing. A Drug Bank that has a huge number of drugs and target information. It is composed of a diverse set of licenced small molecule drugs, biotech pharmaceuticals, and experimental drugs that are linked to non-redundant protein sequences.

Step 2: Data Pre-processing:

Pre-processing is used to eliminate undesired noise and increase contrast between regions of varying brightness. To eliminate undesirable items, pre-processing is used to separate the arithmetic data from the non-numerical data.

Step 3: Compute Descriptors

To begin, the amino acid composition (AAC), dipeptide composition (DC), and tripeptide composition (TC) of the protein sequence are retrieved, as well as the drug SMILES substructure fingerings. The word "descriptors" refers to the terms used to describe the chemical, topological, and geometrical properties of drugs and targets. We gathered 591 medication descriptions by binding Drugs and targets.

Step 4: Check Similarity

In this work, immuno-oncology proteins (UniProt ids - Q6UWE3, P42677, P63173, and Q9Y243) were compared to known protein structures from a drug repository to uncover new biomarkers. Generally, the immune system kills cancer cells. We compared the immune-oncology compound's similarity scores.[15]

Step 5: Feature selection and representation

The retrieved collection of characteristics served as the initialization phase's input. The characteristics were represented as a two-dimensional array, each of which was seeded with a random particle. The grid was constructed using the neighbours of the first random particle that generated the leader. A random array with a size equal to the number of features was formed and randomly filled with values ranging from 0 to 1 using a uniform random function defined by the position of each particle. If the associated index I in the array was greater than the threshold value of 0.45, a feature was picked. [17]

Step 6: Feature Reduction and Extraction

We focus our efforts on the two phases of feature normalization and similarity in order to produce dependable and flexible recognition of features. Accuracy, sensitivity, specificity, precision and F1-Score were utilized as performance indicators for the evaluation of three distinct feature descriptor databases used for PCA-based recognition. Zero Mean and Unit Variance Normalizing feature normalization and similarity Correlation Distance approaches are used in this study to evaluate feature descriptor performance based on principal component analysis (PCA).

- **Zero Mean and Unit Variance Normalization**

Zero Mean and Unit Variance Normalization [9] [10] normalize all of the elements ai (i=1, 2,...,d) of a, it translate and scale the axes so that all the feature vector have zero mean and unit variance. Following expression will produce the normalized feature vector a’.

$$a'_i = \frac{a_i - \mu}{\gamma} \quad (1)$$

Where, μ and γ are the vector mean and the vector standard deviation of that feature respectively.

- **Correlation Distance methods**

Similarity measure used to match the similar subjects (persons) as well as being able to discriminate dissimilar one. Let x, y be the feature vectors of length n. then we can calculate the following distances between these feature vectors

$$\partial(a, b) = 1 - \sum_{i=1}^{n} \frac{(a_i - \bar{a})(b_i - \bar{b})}{\sqrt{\sum_{i=1}^{n}(a_i - \bar{a})^2 \sum_{i=1}^{n}(b_i - \bar{b})^2}} \quad (2)$$

Where,

$$\bar{a} = \sum_{i=1}^{n} a_i \quad \text{And } \bar{b} = \sum_{i=1}^{n} b_i \quad (3)$$

Apply PCA based algorithm on the dataset to select the best feature set PCA-200, PCA-500 and PCA1000.

Step 7: Classification Phase:

Performs training and testing only on the best features set selected by PCA based Algorithm and create a model using support vector machine with RBF kernel classifier. [19]

Step 8: Performance

Evaluate the performance of this model based on some parameters like accuracy, sensitivity, specificity, precision, F1-score and analyse the prediction using mean AUROC.

IV. RESULT AND DISCUSSION

This section contains result and discussion about prediction of Drug target association for breast cancer. For implementing the proposed technique, we have used python. The proposed system has been tested on the data sets. These three dataset repository uses 591 associations for the purpose of classification. Confusion matrix for the model is represented in Table 1.

Table 1: Confusion matrix for Model

Confusion Matrix		200 Descriptor		500 Descriptor		1000 Descriptor	
		Predicted					
		P	N	P	N	P	N
Actual	P	443	17	447	15	453	11
	N	19	112	13	116	9	118

Simulation results comparing expected interactions between a drug and its target. The suggested system's accuracy may be considerably enhanced by employing the rule base. These metrics also validate the specificity and sensitivity of the proposed system. The suggested system is assessed using the following metrics. Table 2 displays the result for model.

Table 2: Table 5.2 Comparative analysis of proposed system using nonlinear SVM with RBF function

Parameter (%)	**200**	**500**	**1000**
Accuracy	93.9	95.3	96.6
Sensitivity	95.8	97.2	98.1
Specificity	86.8	88.5	91.5
Precision	96.3	96.7	97.6
F1 – Score	96.0	96.9	97.8
AUROC	96.9	97.9	98.3

Evaluation findings for classification-based prediction models are depicted in Table 2. In this part, we represented the results of our calculations.

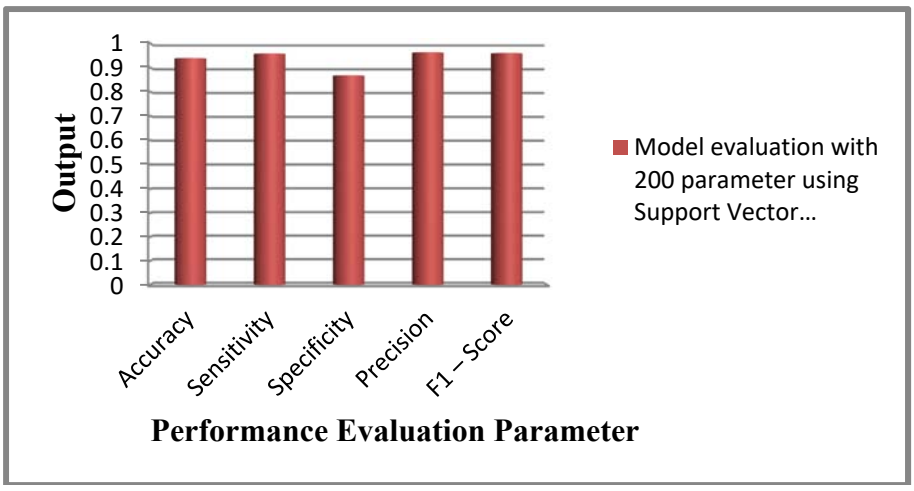


Figure 2: Comparative representation for proposed system with 200 feature attribute

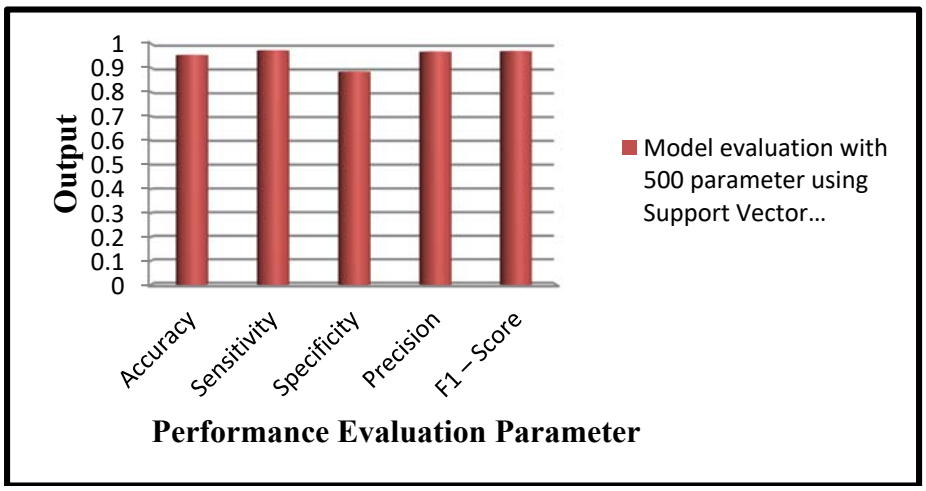


Figure 3: Comparative representation for proposed system with 500 feature attribute

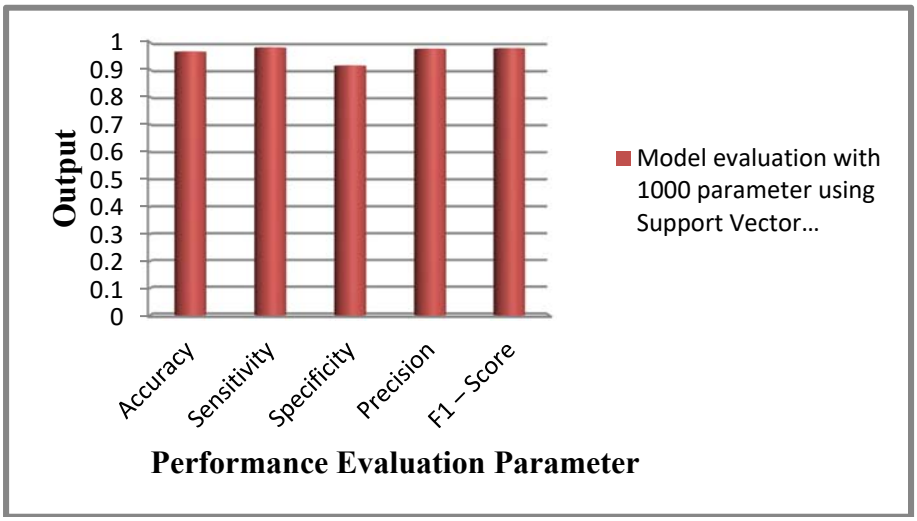


Figure 4: Comparative representation for proposed system with 1000 feature attribute

AUROC measures the model's capacity to distinguish between "cases" (positive instances) and "non-cases" in terms of performance (negative examples.) Assuming that 90 percent of the time, a model accurately assigns a higher absolute risk to a randomly picked patient with an incident than to another randomly selected patient without an event, this indicates that the model has strong discriminating capacity. Comparative analysis of proposed system with different number of feature selection using Support Vector Machine using RBF kernel function below:

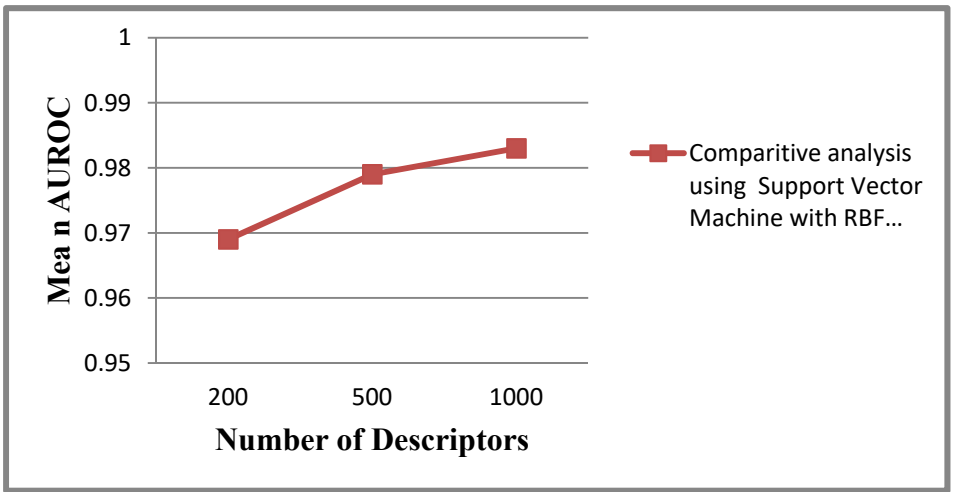


Figure 5: Average Mean of AUROC values for different number of feature descriptors using SVM with RBF kernel function

V. SUMMARY AND CONCLUSION

In this paper, a combined feature selection algorithm was presented to select the best subset of features from bag of features. Best features were chosen using feature selection with feature similarity. The model was used to forecast connections between immuno-oncology agents and disease. The data set consisted of 591 Drug target associations. Zero Mean and Unit Variance Normalization and Correlation Distance methods extracted the features from the dataset. The performance analysis of the present feature selection model was compared with Support Vector Machine with RBF kernel function. The present model performed better in terms of performance parameters against all the aforementioned algorithms with 200, 500 and 1000 features. For the dataset considered, the present algorithm was effective due to selection of lesser number of features in sequential processing of data. We performed cross validation on each model to determine its performance The model's accuracy and mean AUROC were greater than 95%, while increasing the number of descriptor features. The findings indicated that immuno-oncology compounds may be useful as therapeutic candidates for a variety of disorders cancer treatment. The proposed prediction models can aid in drug development by identifying the potential for immuno-oncology compounds to be repurposed for cancer treatment.

VI. REFERENCES

1. Ali Ezzat, Peilin Zhao, Min Wu, Xiao-Li Li, and Chee-KeongKwoh, "Drug-Target Interaction Prediction with Graph Regularized Matrix Factorization", in IEEE/ACM Transactions on Computational Biology and Bioinformatics (2015).
2. Antonio Lavecchia, “Machine-learning approaches in drug discovery: methods and applications,” in Elsevier, Drug Discovery Today Volume 00, Number 00 November 2014.

3. Ashis Kumer Biswas, Nasimul Noman and Abdur Rahman Sikder, “Research article Machine learning approach to predict protein phosphorylation sites by incorporating evolutionary information,” Eleventh BMC Bioinformatics 2010, 11:273.
4. Brian Delavan , Ruth Roberts, Ruili Huang, WenjunBao, Weida Tong and Zhichao Liu, "Computational drug repositioning for rare diseases in the era of precision medicine," in Elsevier, Drug Discovery Today, Volume 00, Number 00, October 2017.
5. Christopher C. Yang, Mengnan Zhao, " Mining heterogeneous network for drug repositioning using phenotypic information extracted from social media and pharmaceutical databases ",in Elsevier, Artificial Intelligence In Medicine 96 (2019) 80–92 .
6. Fabian Pedregosa , Gael Varoquaux , AlexandreGramfort, Vincent Michel and Bertrand Thirion, "Scikit-learn: Machine Learning in Python," Journal of Machine Learning Research 12 (2011) 2825-2830.
7. HuiminLuo, Jianxin Wang, Min Li, Junwei Luo, Peng Ni, Kaijie Zhao, Fang-Xiang Wu, and Yi Pan, "Computational drug repositioning with random walk on a heterogeneous network," IEEE/ACM Transactions On Computational Biologyand Bioinformatics, 2017.
8. Jianlin Cheng, Allison N. Tegge and Pierre Baldi, “Machine Learning Methods for Protein Structure Prediction”, IEEE Reviews In Biomedical Engineering, Vol. 1, 2008.
9. Jiaying You, Robert D. McLeod, Pingzhao Hu, "Predicting drug-target interaction network using deep learning model", in Elsevier, Computational Biology and Chemistry 80 (2019) 90–101.
10. Liang Yu, Ruidan Su, Bingbo Wang, Long Zhang, YapengZou, Jing Zhang, Lin Gao, "Prediction of novel drugs for hepatocellular carcinoma based on multi-source random walk”, IEEE/ACM Transactions on Computational Biology and Bioinformatics,2015.
11. Zhu, J., Xie, Q., & Zheng, K. (2015), “ An improved early detection method of type-2 diabetes mellitus using multiple classifier system”. Information Sciences, 292, 1-14.
12. Lu Zhang, Jianjun Tan, Dan Han and Hao Zhu, "From machine learning to deep learning: progress in machine intelligence for rational drug discovery," in Elsevier, Drug Discovery Today, Volume 22, Number 11, November 2017.
13. Majid Rastegar-Mojarad, Ravi kuma,r Komandur Elayavilli, Liwei Wang, Liwei Wang, Rashmi Prasad, Hongfang Liu, "Prioritizing Adverse Drug Reaction and Drug Repositioning Candidates Generated by Literature- Based Discovery", Proceedings of the 7th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics October 2016 Pages 289–296.
14. Fangjun Kuang, Weihong Xu Siyang Zhang (2014). “A novel hybrid KPCA and SVM with GA model for intrusion detection”. Applied Soft Computing, Volume 18, May 2014, Pages 178-184.
15. Ahmadi, N., Nilashi, M., Samad, S., Rashid, T. A., & Ahmadi, H. (2019). An intelligent method for iris recognition using supervised machine learning techniques. *Optics & Laser Technology*, *120*, 105701.
16. Pemovska T, Johnson E, Kontro M, et al. Axitinib effectively inhibits BCR-ABL1 (T315I) with a distinct bre conformation. Nature. 2015;519:102–105.
17. Orawan, C., Panwadee, S., & Bandit, S. (2016). “Application of artificial neural networks on growth prediction of Staphylococcus aureus in milk”. International Food Research Journal, 23(1), 415.
18. R. Burbidge, M. Trotter, B. Buxton, S. Holden, “Drug design by machine learning: support vector machines for pharmaceutical data analysis”, in Elsevier, Computers and Chemistry 26 (2001) 5–14.
19. Thein, H. & Tun, K. M. M. (2015). “An approach for breast cancer diagnosis classification using neural network”. *Advanced Computing*, *6*(1), 1.