Developing a Credit Card Fraud Detection Model using Machine Learning Approaches

Shahnawaz Khan[1]
Faculty of Engineering
Design and Information & Communications Technology
Bahrain Polytechnic
Isa Town, Bahrain

Abdullah Alourani[2]
Department of Computer Science and Information
College of Science in Zulfi
Majmaah University
Al-Majmaah, 11952
Saudi Arabia

Bharavi Mishra[3]
Department of Computer Science & Engineering
The LNM Institute of Information Technology
Jaipur, Rajasthan, India

Ashraf Ali[4]
Faculty of Computer Studies
Arab Open University, Kingdom of Bahrain

Mustafa Kamal[5]
College of Science & Theoretical Studies
Saudi Electronic University, Dammam, Saudi Arabia

***Abstract*—The growing application and usage of e-commerce applications have given an exponential rise to the number of online transactions. Though there are several methods for completing online transactions, however, credit cards are most commonly used. The increased number of transactions has given the opportunity to the fraudsters to mislead the customers and make them execute fraudulent transactions. Therefore, there is a need for such a method that can automatically classify detect fraudulent transactions. This research study aims to develop a credit-card fraud detection model that can effectively classify an online transaction as fraudulent or genuine. Three supervised machine learning approaches have been applied to develop a credit-card fraud classifier. These techniques include logistic regression, artificial intelligence and support vector machine. The classification accuracy achieved by all the classifiers is almost similar. This research has used the confusion matrix and area under the curve to demonstrate the score of the different performance measures and evaluate the overall performance of the classifiers. Several performance measures such as accuracy, precision, recall, F1-measure, Matthews correlation coefficient, receiver operating characteristic curve have been computed and analysed to evaluate the performance of the credit-card fraud detection classifiers. The analysis demonstrates that the support vector machine-based classifier outperforms the other classifiers.**

Keywords*—Credit card fraud detection; neural network; support vector machine; logistic regression; performance measures

I. Introduction

With the increased use of financial technology, the use of online transactions has increased manifolds in recent years. This expansion and use of electronic commerce have increased the trust of customers in online transactions. There are several kinds of financial fraud such as credit card fraud, securities fraud, insurance fraud, etc. that use online methods to accomplish the fraud. Most online transactions utilize credit cards. Therefore, the most common type of fraud among all the frauds types is credit card fraud [17]. Credit card frauds can be further categorized into offline fraud, application fraud, bankruptcy fraud, internal fraud, behavioural fraud, counterfeit fraud, cardholder-not-present fraud, etc. Online transactions are providing new opportunities for fraudsters. Frauds are activities by the fraudsters that are intended to yield the fraudster personal or financial gain [23]. These activities are often criminal or wrongful. Credit card frauds cause problems and losses to financial institutions as well as individuals. There are hundreds of transactions every second for any financial institution [2]. Manual fraud detection and prevention is not a feasible solution. There has been a tremendous amount of effort by the research community in developing efficient detection techniques for credit card frauds. So that the trust of the customers can increase in e-commerce and online transactions and the losses that occurred due to the frauds can be minimized.

Digital transactions can take place over the phone or on the internet. For executing a transaction, very basic information is required such as expiry date, card number, card verification number etc. Cardholders provide this information through phone or the internet. Fraudsters apply several techniques and attempt to steal the credit card information of the customers so that they can use it for doing fraudulent transactions. It is a very serious, and costly problem for financial service providers. Billions of dollars are subject to fraudulent transactions every year [22]. The fraudulent transaction is an issue of concern for all the credit card providers or by expansion for all the financial systems that provide the facilities for online transactions to their customers. It is usually the result of someone stealing the credit card information of the customers which also impact the brand value of the credit card service providers and the merchants.

The worldwide cost of fraudulent transactions is projected to be 38.5 billion U.S. dollars by 2027 and it was 32 billion dollars in 2021 [26]. The fraudulent transactions cause a huge loss for the merchants also because they usually have to bear all the related costs such as administrative charges, issuer fees etc. The number of digital transactions is huge, therefore, verifying each transaction for its genuineness is not an easy task for the financial service providers. Consequently, credit card providers often only investigate the cases when they are reported by the customers. The literature highlights several other issues in which the primary issue is the imbalance amount of the cases in the available historical data. The number of actual fraud cases in the data is usually very small in comparison to the number of genuine transactions. The imbalance of the training data creates the problem of biases in the classification accuracy of the classifier. The presence of the dominating class corners the other classes. Thus, the classifier keeps predicting the dominating class. Therefore, even if the classifier is predicting wrong, the accuracy of the classifier will not be impacted by large.

Therefore, there is a need for a system that can detect fraudulent transactions efficiently and raise an alarm as soon as the transaction is made, so that the credit card provider can take immediate action and reduce the risk of capital loss. Several researchers have utilized multiple machine learning and other computational methods to detect credit card frauds. This research aims to tackle the issue of data imbalance and develop a credit card fraud detection system. There are several techniques that aim to minimize the effect of the data imbalance. This research applies such techniques for data preparation and during the model evaluation phase. Because the researchers are in the view that only measuring the accuracy will not be a proper evaluation of the classifier.

Various researchers have applied several machine learning and hybrid methods for detecting fraudulent transactions and have developed classifiers that can detect fraudulent transactions. Several researchers have used standalone methods [7], [8], [13], [17], [27] while many researchers have also applied hybrid approaches [3], [22] for detecting credit card frauds. The issue of the presence of the highly imbalanced amount of data samples is the primary challenge in developing an effective credit-card fraud detection model. Several approaches have been applied such as feature selection, feature engineering, sequence classification, supervised and unsupervised machine learning methods, data pre-processing to balance the data classes.

Credit card fraud detection methods discussed in this research focus on identifying if a transaction on a credit card is fraudulent or not by applying several machine learning techniques such as logistic regression, artificial neural networks, and support vector machines. Credit card fraud detection systems use historical transaction data to train. The decision of these systems relies on the spending behavioural patterns learned during the training process from historical transaction data. The system aims to develop an efficient credit-card fraud detection model that can effectively classify the transactions into genuine transactions or fraudulent transactions efficiently. Several performance measures such as accuracy, precision, recall, F1-measure, Matthews correlation coefficient, receiver operating characteristic curve have been calculated to evaluate the performance of different classification models. Among the implemented models, the support vector machine model performs better.

This paper has been organized into five sections. The second section of the paper examines the literature review and presents the background work briefly. Section three discusses machine learning techniques, data and pre-processing. Section four presents the results obtained from different classifiers implemented by this research study and also exhibits the various performance measures and evaluates the performance of the credit card fraud detection models. The last section presents the conclusion.

II. Related Work

In binary classification, the basic concept is to find the threshold value that enables the classifier to assign a particular label to the case or instance being predicted. There have been conducted several research studies on detecting credit card frauds based on the spending behaviour of the customer. However, sometimes the customer spending behaviour changes during certain conditions such as holidays or other special occasions. This might create an issue for supervised machine learning systems. Research by [3] proposes a hybrid approach for managing customer abnormal spending behaviour. The proposed approach combines supervised and unsupervised machine learning approaches and presents effective results in case of abnormalities of the spending behaviours.

A research study [7] applies a generative adversarial network (GAN) to detect credit card frauds. As the credit card transaction data is usually highly imbalanced, the proposed framework by the research study [7] has a higher false-positive rate if the sensitivity is improved. Research [22] applies and compares several machine learning techniques such as Naïve Bayes, random forest, logistic regression, decision tree, AdaBoost, multiple layer perceptron etc. The research demonstrates that the AdaBoost with majority voting produces the best results among all the alternatives. There is an interesting fact to consider for this research [22], that the transactions data used for developing the classifier has only 0.0355% fraudulent transactions. The data used is highly imbalanced and no measures have been taken to counter the imbalances of the data. In the research study presented by [17], several machine learning methods have been applied for credit card fraud detection. The study illustrates that SVM, ANN, C5.0 decision tree, and LR performs better among the tested criteria. However, the number of false positives is high among all the implemented methods.

A random forest algorithm is an effective method for developing supervised classification models. A research study by [27] implements random forest supervised machine learning techniques to detect the behavioural patterns for genuine transactions and fraudulent transactions. A similar research study by [13] proposes a random forest algorithm-based machine learning model for detecting credit card frauds. The model presented by the research [13] exhibits good accuracy. Though, the performance measure is based on the statistics obtained from the confusion matrix only. As has been discussed above by several researchers the credit card fraud detection problem poses the challenge of the imbalance data classes. Therefore, there should have been some other performance measure such as the receiver operating characteristic (ROC) curve that could have been employed for a better performance measure. In a similar domain, the research study by [18] focuses on minimizing the number of incorrect fraud classifications. Actually, that's the primary target of any researcher working in this domain. The research study by [18] employs multiple algorithms for anomaly detection and implements algorithms such as isolation forest and local outlier factor algorithms. The results presented by the study are sensitive to the quantity of the data and face the challenge of imbalanced data for the classification classes.

A research study by [20] presents an interesting perspective on credit card fraud detection and infer that there is no constant pattern for fraud. Therefore, supervised machine learning techniques are not efficient in credit card fraud detection. It [20] proposes an unsupervised machine learning approach using a restricted Boltzmann machine (RBM) and deep Auto-encoder. However, the results achieved are less promising than some of the supervised machine learning approaches. A Bayesian network classifier-based approach presented by [6] in a research study uses a hyper-heuristic evolutionary algorithm to detect the patterns. The presented solution like the other approaches discussed in this section targets the class imbalance and misclassification issues of the credit card fraud detection problem. Research by [2] introduces a real-time fraud detection system using machine learning and big data. This solution primarily focuses on the detection speed of the transactions, use of big data and scalability.

The research study by [4] considers the spatial and temporal features among others and presents a 3D convolutional neural network for credit card fraud detection. The present approach [2] implements the model on the real-world data collected from multiple locations. The research by [15] proposes a hidden Markov model-based approach for automated feature engineering to improve the performance of the classifier and to model temporal feature correlation. A similar research study [28] that focus on the features of the transactional data, develops a deep learning-based solution that uses homogeneity-oriented behaviour analysis for feature engineering. A research study [1] proposes an optimized light gradient boosting based machine learning technique for predicting credit card frauds. This research [1] relies on parameter optimization for improving the performance of the classifier.

An interesting approach to solving the credit card detection problem is sequence classification or prediction problem. Research [8] formulates credit card fraud detection as a sequence classification problem. It applies long short-term memory neural network to identify the fraudulent transaction. The research concludes that articulating the fraud detection task as a sequence-learning problem leads to an increased number of false positives. As a matter of fact, online transactions should not be considered a sequential classification problem, because the amount, time, and point of the online transactions usually change randomly. It will require a highly disciplined spending behaviour to express online transactions as a sequence learning problem.

III. Methodology

A. Dataset and Pre-processing

One of the primary issues for data preparation for credit-card fraud transaction data is the labelling of the data. Often the fraudulent label of the transactional data can only be decided posterior the transaction has been executed and reported by the customer. The dataset used in this research study consists of transactions made by European cardholders in September 2013 [14]. The dataset contains 284,807 transactions made during two days. The fraudulent transactions made during this time were 492 which is just 0.172% of all transactions made during this time. As it is evident that the data is imbalanced, therefore, this research uses the resampling technique and makes an effort to oversample the fraudulent transactions and to remove the genuine transactions. The dataset was transformed using principal component analysis to maintain the confidentiality of the transactions [14] and the principal components are used as features for training the classifiers. The dataset contains 30 input features such as transaction time, transaction amount and 28 principal components. The output classes have two labels 1 and 0. The fraudulent transactions are assigned label 1 and the genuine transactions are labelled as 0.

B. Modelling for Credit Card Fraud Detection

Several approaches and algorithms have been implemented for credit card fraud detection. Some of these solutions have been discussed in the related work section. Several features and affairs have been taken into consideration for credit card fraud detection classifiers. One of the common issues that are discussed throughout the literature in credit card fraud detection is the presence of class data imbalance. However, none of the algorithms or approaches discussed precisely tackle the class imbalance issue. Therefore, this research implements a two-step process for handling the issue. The first phase is data pre-processing. In the data pre-processing phase, the study aims to reduce the class data imbalance by increasing the number of cases for the minority class, and by reducing the number of cases for dominating class. This section discusses the approaches used in this study.

1) Logistic regression: Logistic regression is a probabilistic modelling process that produces the probability of the discrete output variables based on the input variables. Often logistic regression is applied for binary classification when the input variable is single or multiple. However, logistic regression can be applied to classify more than two output classes, which is known as multinomial logistic regression. Furthermore, it can be used for ordering the level of the output variable classes which is known as ordered logistic regression. However, logistic regression is often used for binary classification problems. Credit card fraud detection is a binary classification problem in which the output of the transaction is either fraud or a genuine transaction given the input features for the transactions. Therefore, logistic regression can be used as a credit card fraud detection technique.

Credit card detection can be performed by computing the probability of the given transaction using the given features and comparing it with a threshold value such as 0.5. If the computed probability is more than 0.5 then it will be classified as fraud if less than the threshold then it will be classified as a genuine transaction. Let us assume that the probability of the fraudulent transaction based on the transaction features x is P(y = 1|x) or simply P(x). To compute the probability-estimate log-odds can be computed. Log-odds are directly proportional to the probability of the transaction label. Higher the odds, the higher the probability of the given label for the transaction.

It can be defined as: $\frac{P(x)}{1-P\ (x)}$. (1)

For modelling and simplifying the computational process, natural logarithm was applied as follows:

$$logit(x) = \log\left(\frac{P(x)}{1-P(x)}\right) \quad (2)$$

Let's consider, $\log\left(\frac{P(x)}{1-P(x)}\right) = w'x + b$ (3)

Here, w' is the transpose of the weight vector and b is the offset variable. The above equation (Eq. 3) can be further simplified by applying exponential on both sides:

$$P\ (x) = \frac{e^{w'x+b}}{1+ e^{w'x+b}} \quad (4)$$

Therefore, the probability of the fraudulent $P\ (x)$ can be estimated using the above equation (Eq. 4) in which x are the features of the transaction. The aim is to optimize the values of w and b based on the transactional data. It can be learnt by converting the above problem into maximum likelihood estimation problem and optimizing it for w and b using the transactional data.

The log-likelihood from the equation (eq 4) can be derived as following (Eq 5):

$$log\ (w,b) = \sum_1^n [y_i \log P(x_i) + (1 - y_i) \log(1 - P(x_i))] \ (5)$$

The optimized values of w and b are estimated by maximizing the log-likelihood (Eq. 5) or by converting the above problem into minimization problem after multiplying with a negative sign.

2) Artificial neural network: Artificial neural network (ANN) is one of the most powerful machine-learning techniques. ANN aims to simulate the behaviour of biological organisms. In the human nervous system, neurons are connected to other neurons through connections which are known as axons and dendrites. The strength of the connections is subject to change in accordance with the external stimuli which are referred to weights in ANN. The computational units in ANN are termed as neurons. Though, the ANN simulation of the biological organism is very basic still, the complexity and computation capabilities of the artificial neural network is very powerful. Artificial neural networks have been applied to solve complex computational problems for example in machine translation [9], [10], [11], [24], [25], image processing [12], time series forecasting [21], classification etc. There are several neural network architectures that are employed in machine learning for various different tasks. The following diagram (Fig. 1) presents a general feed-forward neural network architecture. This study employs a feed-forward neural network. Input layer is the first layer and works as input to the neural network. The input layer of the neural network used in this study contains 25 neurons corresponding to the features of a transaction. Two middle layers, popularly known as hidden layers, have been used each of which contains 10 neurons. The output layer has 2 neurons corresponding to each class. The network uses the backpropagation algorithm for learning. The layers are fully connected layers. The activation function used is rectified linear units (ReLU).

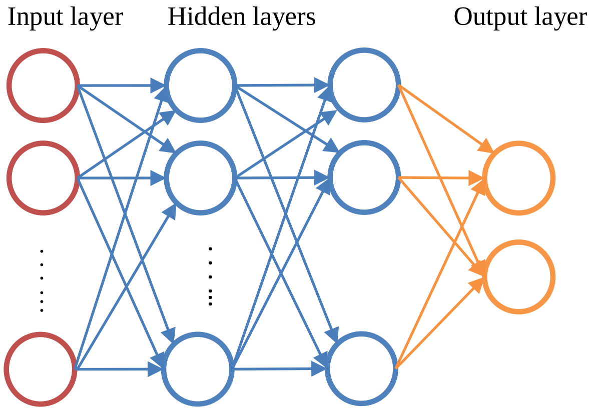

Fig. 1. Feed-forward Neural Network General Architecture.

The ReLU activation function is a very simple but effective activation function. It returns a zero if the input received by the activation function is negative, otherwise, no change is applied on the input and the input values is returned as it is if the input is positive. It can be simply stated as:

$$f(x) = \begin{cases} 0, & x \leq 0 \\ x, & x > 0 \end{cases} \quad (6)$$

3) Support vector machines: Support vector machines like artificial neural networks have been among the most popular machine learning algorithms. SVMs are commonly applied in solving supervised machine learning problems such as regression, classification and outlier detection. Though, the number of samples for this research study is ample, but, SVMs can also be applied in a scenario where the number of dimensions is more than the sample size. SVMs, for classification, functions by finding the best hyperplanes that separate the data points in accordance with the classes. The hyperplanes set apart the data points of one class from the data points of other classes. The optimization is applied for finding the best hyperplane that can find the maximum margin among the data points of one class from the data points of the other class. The support vectors are the closes data points to the hyperplane. SVMs can be implemented using different types of kernel functions. Kernel methods are a set of algorithms that are used in machine learning techniques for pattern analysis and detection. Kernel functions, transforms the data into higher dimensions, expecting to find clearer decision boundaries for data separation. Kernel functions aids in efficiently transforming high dimensional data for creating optimal boundaries for decision making. The kernel function used in this research is the quadratic kernel function.

A quadratic kernel is a non-stationary and special form of the polynomial kernel. The general form of the polynomial function looks as follows:

$$K(f, g) = (f^T g + c)^d \tag{7}$$

Here, f and g are the computed vectors of features from the input data samples, c is a free parameter and has a value of c ≥ 0, and d is the degree of polynomial. When the degree d = 2. Then, the kernel function is called quadratic kernel function and can be presented as follows:

$$K(f, g) = (f^T g + c)^2 \tag{8}$$

IV. Results and Discussion

This research study implements several machine learning methods for credit card fraud detection. The selected methods are logistic regression, artificial neural network and support vector machines. Three models were trained using the above-mentioned machine learning technique on the selected training data. The training process has applied five-cross validation. Test data samples were randomly selected before applying resampling techniques to evaluate the performance of the system on real-world data. The accuracy achieved for the logistic regression method is 99.92% and the prediction speed is around 300 thousand predictions per second. The accuracy for the neural network-based model is 99.92% while the prediction speed is 650 thousand predictions per second. Support vector machines model has the prediction speed of about 350 thousand predictions per second and the accuracy of 99.94%. The prediction speed of the artificial neural network-based model is the fastest among the tested models, but the application of the model on big data has yet to be tested [2]. As the accuracy achieved by all the developed models is similar, therefore, some other evaluation metrics must be used to measure the performance of the developed models. The confusion matrix for classification models demonstrates the true positive (TP), true negatives (TN), false positives (FP), and false negatives (FN). It illustrates how many of the instances have been classified to their actual class and how many have been misclassified. The following Tables I to III illustrate the confusion matrix and several performance metrics for the three classification models:

1) Accuracy: Accuracy is the primary performance evaluation metric and measures the ratio of correct prediction over the total number of predictions by the classifier. It can be presented as:

$$Accuracy = \frac{True\ Positive\ +\ True\ Negative}{Total\ Predictions}$$

The models (logistic regression, artificial neural network and support vector machines) achieved an accuracy of 99.91%, 99.91% and 99.94% respectively.

2) Precision: Precision is also known as the positive predictive value and is the ratio of the true positive predictions over the total positive predictions:

$$Precision = \frac{True\ Positive}{True\ Positive\ +\ False\ Positive}$$

The models (logistic regression, artificial neural network and support vector machines) achieved a precision of 87.32%, 76.91% and 87.67% respectively.

3) Recall: Recall measures the true positive rate of the classifier and is also known as sensitivity in binary classification. It is calculated as the fraction of the true positive predictions over all the positive cases that were retrieved for the testing:

$$Recall = \frac{True\ Positive}{True\ Positive\ +\ False\ Negitive}$$

The sensitivity or recall measured for logistic regression classifier is 61.59%, recall for the artificial neural network is 75.81% and for support vector machines is 78.05%.

TABLE I. CONFUSION MATRIX AND PERFORMANCE METRICS FOR LOGISTIC REGRESSION

		Logistic Regression		
		Predicted Class		
		Genuine	***Fraud***	
True Class	**Genuine**	284271	44	99.98%
	Fraud	189	303	61.59%
			87.32%	99.92%

TABLE II. CONFUSION MATRIX AND PERFORMANCE METRICS FOR ANN

		Artificial Neural Network		
		Predicted Class		
		Genuine	***Fraud***	
True Class	**Genuine**	284203	112	99.96%
	Fraud	119	373	75.81%
			76.91%	99.92%

TABLE III. CONFUSION MATRIX AND PERFORMANCE METRICS FOR SVM

		Support Vector Machine		
		Predicted Class		
		Genuine	***Fraud***	
True Class	**Genuine**	284261	54	99.98%
	Fraud	108	384	78.05%
			87.67%	99.94%

4) Specificity: To measure the true negative prediction rate of the classifier, specificity is calculated as the proportion of the true negative predictions over the total negative cases that were retrieved for the testing:

$$Specificity = \frac{True\ Negitive}{True\ Negitive\ +\ False\ Positive}$$

The negative prediction rate or the specificity of the logistic regression is 99.98%, specificity for the artificial neural network is 99.96%, and the specificity for the support vector machines-based classifier was measured as 99.98%.

5) F1-Score: F1-score or F-measure considers the importance of true positive and true negative. It is the harmonic mean of the two performance measures calculated earlier which are precision and recall:

$$F1 - score = 2 * \frac{Precision * Recall}{Precision\ +\ Recall}$$

F1-score measured for logistic regression classifier is 72.23%, 76.36% for artificial neural network classifier, and 82.58% for support vector machines.

6) Matthews Correlation Coefficient (MCC): Matthews correlation coefficient [19] calculates the correlation between the actual classes and predicted classes of the cases. Matthews correlation coefficient (MCC) provides a more accurate evaluation of the overall performance of the binary classifier than other performance measures such as precision, recall, F1-score, and accuracy [5]. MCC is the ration of the covariance of the actual classes of the cases and the predicted labels over the product of the standard deviations of the true classes (σ_t) and the predicted classes (σ_p). MCC is measured as following:

$$MCC = \frac{Cov(t,p)}{\sigma_t.\sigma_p}$$
$$= \frac{TP.TN - FP.FN}{\sqrt{(TP + FP)(TN + FN)(TP + FN)(TN + FP)}}$$

MCC has been calculated for all the three classifiers. The value of the Matthews correlation coefficient remains between -1 and +1. Higher the value of MCC, better the model is. The MCC value for logistic regression is 0.733 or 73.3%, 76.32% for the artificial neural network classifier and 82.69% for the support vector machine classifier.

7) Receiver operating characteristic curve: Another important criterion to consider is the highly imbalanced amount of data points in the training data. Highly imbalanced data introduces several issues in developing machine learning models. One of such issues is biasness. In the case of highly imbalanced data, the prediction accuracy is usually biased. In the case of imbalanced data, the accuracy calculated based on the confusion matrix might be misleading because it will not address the issue of biased classification. Therefore, some other evaluation measure should also be considered while evaluating the performance of the classification model. This research has considered the Receiver Operating Characteristic (ROC) curve as an additional performance measure for the classifiers. ROC curve was initially developed and applied during the world war II for detecting the enemy objects [16]. ROC is fundamentally a graphical representation or a plot that illustrates the accuracy of the classification capability of a binary classifier (Lusted, 1971). ROC curve is a widely used performance measure to evaluate the performance of the binary classifiers. ROC curve plots the sensitivity of the classifier against the false positive rate. False positive rate can be obtained by subtracting the specificity of the classification model from one. The graph is drawn on a 1x1 space which means that the scale on each of the x and y-axis is in the range of 0 to 1. The line connecting the coordinates (0, 0) and (1, 1) will represent a random classifier. An ideal classifier would score a point on the upper left corner (0, 1) which represents the case of zero false positives and zero false negatives.

The following Fig. 2 demonstrates the ROC curve for logistic regression model. The logistic regression classification model yields a point (0.38, 1) and the area under curve is 0.97.

The following Fig. 3 demonstrates the ROC curve for artificial neural network classifier. The artificial neural network model yields the threshold point (0.24, 1) and the area under the curve for artificial neural network model is 0.90. The following Fig. 4 illustrates the ROC curve for support vector machine model. The support vector machine classification model yields a point (0.22, 1) and the area under curve is 0.94. It can be seen that the point yield by the support vector machine vector model is the closest to the point of the best classification model.

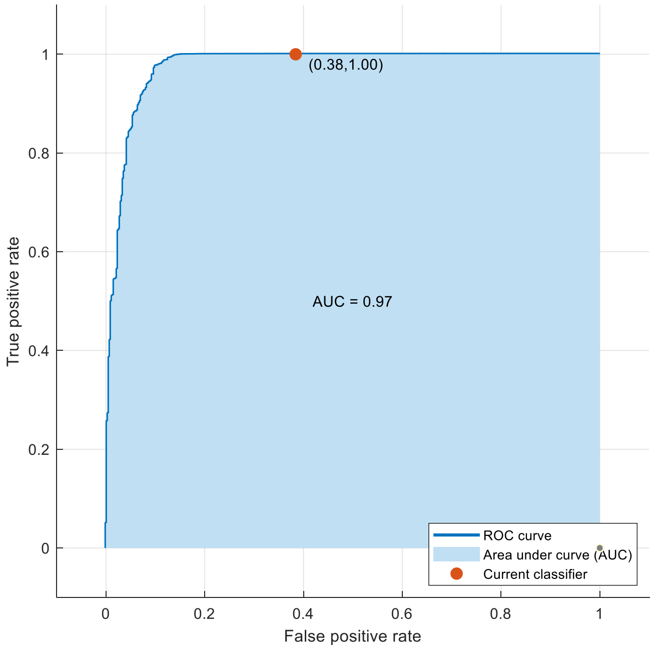

Fig. 2. ROC Curve for Logistic Regression Model.

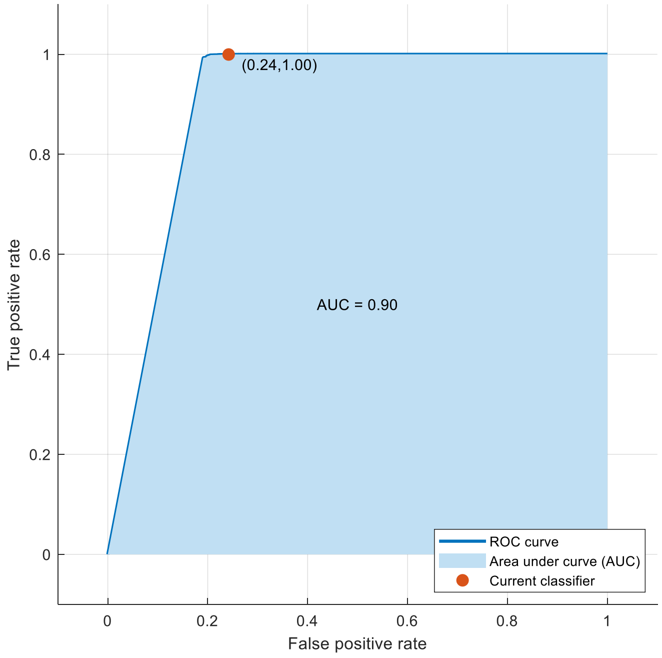

Fig. 3. ROC Curve for Artificial Neural Network Model.

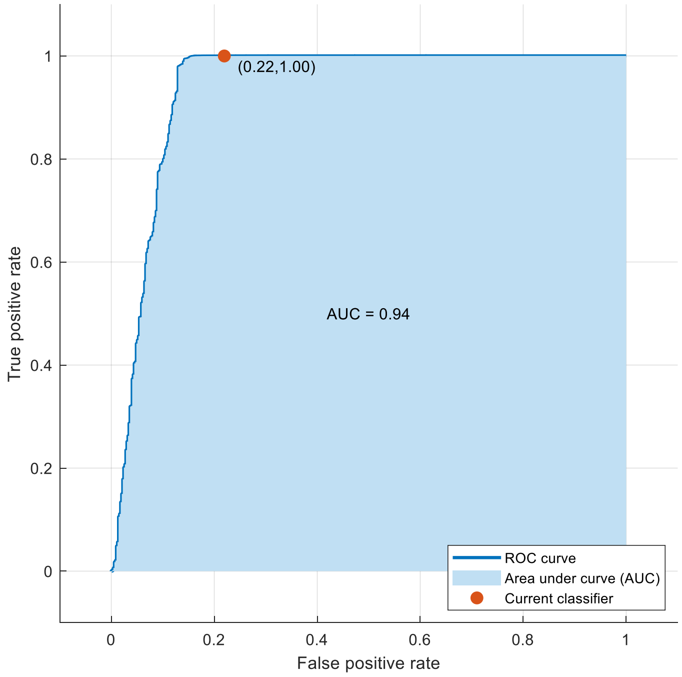

Fig. 4. ROC Curve for Support Vector Machine Model.

V. Conclusion

Credit card fraud is an issue of concern among financial institutions and causes huge financial losses for service providers. Fraudulent transactions have cost over 32 billion United States dollars worldwide in 2021. This amount is projected to increase by over 38 billion dollars in the next 5 years by 2027. Several computational approaches have been employed to develop an effective model for credit card fraud detection. Researchers have employed supervised and unsupervised machine learning approaches. However, the supervised machine learning approaches have produced better results. There are several issues while developing the credit-card fraud detection model. Availability of the highly imbalanced class data is the issue of major concern. The presence of the dominating class corners the other classes. Thus, the classifier keeps predicting the dominating class. Therefore, even if the classifier is predicting wrong, the accuracy of the classifier will not be impacted by large. This research study has applied a resampling technique to counter the effect of imbalanced class data. However, due to the nature of the problem, it is neither feasible nor practical to completely ignore and eliminate the gap of imbalanced data classes. This research study has implemented three machine learning techniques which are logistic regression (LR), artificial neural network (ANN), and support vector machines (SVM).

The models have been evaluated thoroughly using different performance evaluation measures and matrices. Though based on the accuracy computed from the confusion matrix, all the model scores same. But, further analysis using different performance measures demonstrates that the support vector machines classification model outperforms the other models. The prediction accuracy and specificity are almost the same for all the classification models, while the precision is almost 12% lower for the ANN model than the other two models. While the SVM model has slightly higher precision than the LR model. Recall of the SVM model is almost 21% higher than the LR model and almost 3% higher than the ANN model. Similarly, the MCC value and F1-score for the SVM model are over 12% higher than the LR model and 7% higher than the ANN model. Receiver operating curve yields a point (0.38, 1) for the LR model, (0.24, 1) for the ANN model, and (0.22, 1) for the SVM model. The best-case scenario for the classifier on the ROC curve is to yield a point on the upper left corner (0, 1) which represents the case of zero false positives and zero false negatives. Among, the three tested models, SVM is the closest classification model to the best point (0, 1). Therefore, it can be concluded based on the various performance measures that the SVM model outperforms the other models for credit card detection.

Acknowledgment

The authors would like to thank the Deanship of Scientific Research at Majmaah University for supporting this work under project no: R-2022-50.

References

[1] Altyeb, A. A., & Malebary, S. J. (2020). An Intelligent Approach to Credit Card Fraud Detection Using an Optimized Light Gradient Boosting Machine. IEEE Access, 1–1. doi:10.1109/access.2020.2971354.

[2] Carcillo, F., Dal Pozzolo, A., Le Borgne, Y.-A., Caelen, O., Mazzer, Y., & Bontempi, G. (2018). SCARFF: A scalable framework for streaming credit card fraud detection with spark. Information Fusion, 41, 182–194. doi:10.1016/j.inffus.2017.09.005.

[3] Carcillo, F., Le Borgne, Y.-A., Caelen, O., Kessaci, Y., Oblé, F., & Bontempi, G. (2019). Combining Unsupervised and Supervised Learning in Credit Card Fraud Detection. Information Sciences. doi:10.1016/j.ins.2019.05.042.

[4] Cheng, D., Xiang, S., Shang, C., Zhang, Y., Yang, F., & Zhang, L. (2020). Spatio-Temporal Attention-Based Neural Network for Credit Card Fraud Detection. Proceedings of the AAAI Conference on Artificial Intelligence, 34(01), 362–369. doi:10.1609/aaai.v34i01.5371.

[5] Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC genomics, 21(1), 1-13.

[6] De Sá, A. G., Pereira, A. C., & Pappa, G. L. (2018). A customized classification algorithm for credit card fraud detection. Engineering Applications of Artificial Intelligence, 72, 21-29.

[7] Fiore, U., De Santis, A., Perla, F., Zanetti, P., & Palmieri, F. (2017). Using generative adversarial networks for improving classification effectiveness in credit card fraud detection. Information Sciences. doi:10.1016/j.ins.2017.12.030.

[8] Jurgovsky, J., Granitzer, M., Ziegler, K., Calabretto, S., Portier, P.-E., He-Guelton, L., & Caelen, O. (2018). Sequence classification for credit-card fraud detection. Expert Systems with Applications, 100, 234–245. doi:10.1016/j.eswa.2018.01.037.

[9] Khan, S. N., & Usman, I. (2019). Amodel for english to urdu and hindi machine translation system using translation rules and artificial neural network. Int. Arab J. Inf. Technol., 16(1), 125-131.

[10] Khan, S., & Mishra, R. B. (2011). Translation rules and ANN based model for English to Urdu machine translation. INFOCOMP Journal of Computer Science, 10(3), 36-47.

[11] Khan, S., & Mishra, R. B. (2012). A neural network based approach for English to Hindi machine translation.

[12] Khan, S., Thirunavukkarasu, K., Hammad, R., Bali, V., & Qader, M. R. (2021). Convolutional neural network based SARS-CoV-2 patients detection model using CT images. International Journal of Intelligent Engineering Informatics, 9(2), 211-228.

[13] Kumar, M. S., Soundarya, V., Kavitha, S., Keerthika, E. S., & Aswini, E. (2019). Credit Card Fraud Detection Using Random Forest Algorithm. 2019 3rd International Conference on Computing and Communications Technologies (ICCCT). doi:10.1109/iccct2.2019 .8824930.

[14] Le Borgne, Yann-A., and Gianluca Bontempi. " Machine Learning for Credit Card Fraud Detection - Practical Handbook " Universit'e Libre de Bruxelles, 2021, https://www.researchgate.net/publication/351283764_ Machine_Learning_for_Credit_Card_Fraud_Detection_-_Practical_ Handbook.

[15] Lucas, Y., Portier, P. E., Laporte, L., He-Guelton, L., Caelen, O., Granitzer, M., & Calabretto, S. (2020). Towards automated feature engineering for credit card fraud detection using multi-perspective HMMs. Future Generation Computer Systems, 102, 393-402.

[16] Lusted, L. B. (1971). Signal detectability and medical decision-making. Science, 171(3977), 1217-1219.

[17] Makki, S., Assaghir, Z., Taher, Y., Haque, R., Hacid, M.-S., & Zeineddine, H. (2019). An Experimental Study with Imbalanced Classification Approaches for Credit Card Fraud Detection. IEEE Access, 1–1. doi:10.1109/access.2019.2927266.

[18] Maniraj, S. P., Saini, A., Ahmed, S., & Sarkar, S. (2019). Credit card fraud detection using machine learning and data science. International Journal of Engineering Research and, 8(09).

[19] Matthews, B. W. (1975). Comparison of the predicted and observed secondary structure of T4 phage lysozyme. Biochimica et Biophysica Acta (BBA) - Protein Structure, 405(2), 442–451. https://doi.org/https:// doi.org/10.1016/0005-2795(75)90109-9.

[20] Pumsirirat, A., & Yan, L. (2018). Credit card fraud detection using deep learning based on auto-encoder and restricted boltzmann machine. International Journal of advanced computer science and applications, 9(1), 18-25.

[21] Qader, M. R., Khan, S., Kamal, M., Usman, M., & Haseeb, M. (2021). Forecasting CO2 Emissions Due To Electricity Power Generation In Bahrain.

[22] Randhawa, K., Loo, C. K., Seera, M., Lim, C. P., & Nandi, A. K. (2018). Credit Card Fraud Detection Using AdaBoost and Majority Voting. IEEE Access, 6, 14277–14284. doi:10.1109/access.2018. 2806420.

[23] Sahin, Y., Bulkan, S., & Duman, E. (2013). A cost-sensitive decision tree approach for fraud detection. Expert Systems with Applications, 40(15), 5916-5923.

[24] Shahnawaz, & Mishra, R. B. (2015). An English to Urdu translation model based on CBR, ANN and translation rules. International Journal of Advanced Intelligence Paradigms, 7(1), 1-23.

[25] Shahnawaz, M. R. (2011). ANN and rule based model for English to Urdu-Hindi machine translation system. In Proceedings of National Conference on Artificial Intelligence and agents: Theory& Application, AIAIATA (pp. 115-121).

[26] Szmigiera, M. (2021). Value of fraudulent card transactions worldwide 2021-2027, Statista.com, accessed on 15-Jan-2022. https://www.statista .com/statistics/1264329/.

[27] Xuan, S., Liu, G., Li, Z., Zheng, L., Wang, S., & Jiang, C. (2018). Random forest for credit card fraud detection. 2018 IEEE 15th International Conference on Networking, Sensing and Control (ICNSC). doi:10.1109/icnsc.2018.8361343.

[28] Zhang, X., Han, Y., Xu, W., & Wang, Q. (2019). HOBA: A Novel Feature Engineering Methodology for Credit Card Fraud Detection with a Deep Learning Architecture. Information Sciences. doi:10.1016/j.ins.2019.05.023.