Global Journal of Engineering and Technology Advances
eISSN: 2582-5003
Cross Ref DOI: 10.30574/gjeta
Journal homepage: https://gjeta.com/

(REVIEW ARTICLE)

Check for updates

Credit card fraud detection and classification by deep learning and machine learning

Kiran Bala *, Sakshi sharma, Meenakshi Garg and Deeksha Verma

Department of computer science and Engineering in Chandigarh Engineering College, Jhanjeri, Mohali, India.

Global Journal of Engineering and Technology Advances, 2022, 13(03), 022–027

Publication history: Received on 02 November 2022; revised on 12 December 2022; accepted on 15 December 2022

Article DOI: https://doi.org/10.30574/gjeta.2022.13.3.0202

Abstract

One of the most important contributors to the expansion and progression of a nation's economy is its banking and financial industry. In particular, over the recent past, there has been a significant increase in the utilization of credit and debit cards, whereby all customers trade transactions either digitally over the internet or physically at the stores. Here, the customers, banking institutions, and financial organizations are all being put in a difficult position by fraudulent actors. Because more recent technology is now readily available, internet banking has become an important avenue for commercial transactions. Fake banking activities and fraudulent transactions are serious problem that affects both the users' sense of safety and their trust in the system. In addition, fraudulent activities result in enormous losses because of the proliferation of sophisticated frauds such as virus infections, scams, and fake websites. These frauds are all examples of advanced fraud. This study makes three contributions toward the prevention of fraudulent activity involving credit card transactions.

Keywords: Machine learning; AI; Credit card; Deep learning

1. Introduction

1.1. Fraud Detection

The online banking fraud detection technique utilizes the broad-based Wisdom Web of Things (W2T) method [6] [7]. It also provides multi-feature data of digital banking consumers, it includes electronic fund transactions data, demographic data, credit card transactions data, and relevant types of data. These multi-aspects of data are transmitted through the Internet or World Wide Web (WWW) to the data center. These data center offers a platform to perform an e-banking fraud recognition process [10]. In online banking, customers, and computer systems are combined in one unit for recognizing their relationship and synchronization. In this W2T data cycle, fraud prediction is one of the most significant tasks. Many customers infrequently ensure their digital banking history recurrently and thus it is not capable of recognizing and reporting fraud transactions instantly after the occasion of fraud. This process makes the opportunity for loss revival very less [96]. Additionally, every alert produced from the detection system requires manual investigation and it is time-consuming. Normal, e-banking identification scheme requires very high detection rate, high accuracy, and less false positive rate.

Frauds can be classified into two methods, namely offline fraud, and online fraud.

- Offline fraud: Most offline fraud events present because of a wallet or purse and it includes several important documents. The documents, like identify card, credit cards, debit cards, driving license and so on and these documents include essential information, like transaction slips, name, bank account details, date of birth and so on.

* Corresponding author: Kiran Bala
Department of Computer Science & Engineering , CGC Jhanjeri, Mohali, Punjab , India .

Copyright © 2022 Author(s) retain the copyright of this article. This article is published under the terms of the Creative Commons Attribution Liscense 4.0.

- Online fraud: This method presents while, fraud presents their website as the real website for obtaining significant delicate data of consumer and executes legitimate transactionson particular client account.

In general cases, there are two approaches utilized to conflict the fraud, such as fraud avoidance and fraud recognition. Fraud preservation is used to sort out high-risk transactions and it is a starting process of security. Additionally, there are various authentication systems, such as expiry date, cardholder's address, signatures, and identification number are used for credit card fraud protection. Besides, the detection method is classified into two, like anomaly discovery and misuse recognition [2] [3]. The anomaly discovery normally used normal transactions to recognize frauds. Likewise, misuse detection used the labeled transaction to identify frauds [4]. The gradually progressing cashless economy directs business movement to electronic monetary transactions and it is utilized for the accurate reorganization of fraud [5]. Generally, fraud detection is the process of observing the transaction behaviour of the cardholder. Misuse detection types are utilized for identifying whether incoming transaction is fraud or not. Typically, misuse types have knowledge about present kinds of fraud for making methods through the learning of several fraud patterns. Similarly, anomaly detection type is used for creating a profile of normal transaction behavior of cardholders using their historical transaction data. The anomaly-based fraud detection identifies whether the new transaction is different from other general transaction behaviours [7].

Credit card is also one of the most unauthorized kinds of fraud. A credit card is a synthetic card, which is provided to bank clients as one of the payment modes [89]. A credit card permits cardholders for purchasing products and goods from shopping websites or marketplace. This type of fraud involves a person using other person’s credit card for individual utilization whereas, the vendor of the credit card and card issuer is not conscious of the fact that their card was used by others. The huge utilization of credit cards and the deficient of efficient security methods lead to a billion-dollar loss to credit card fraud. Credit card firms are disinclined to declare information and it is complex for obtaining an accurate approximation of losses. The utilization of credit cards with a lack of tough security causes billion-dollar economic sufferers. The global economic losses that occur because of credit card fraud amount to 22.8 billion dollars in 2017 and constantly increase by 2022. In credit card fraud there are two classifications, namely behaviour fraud and application fraud [83]. Application fraud refers to a fraud during the time of new credit card application process, by providing fake identity information and the issuer also acknowledges. Besides, behaviour fraud happens after providing thecredit card, accurately and it indicates credit card transactions, which includes fraud behaviour [11]. Credit card fraud recognition is an imperative problem for financial organizations and credit card users. The fraud detection for a small amount helps to protect a huge amount of money and credit card fraud and detecting fraudulent activities is an important issue for research [1].

There are two stages, such as the Near Real-Time (NRT) stage and Real-Time (RT) Stage in the online automatic fraud detection system. The system decides rapidly whether to block a transaction using bared transaction data in the RT stage. Likewise, the system performs a slower ex-post assessment using a huge information context, which involves linked data in successive NRT stages. In the NRT stage, classification rules are applied to produce awareness of doubtful transactions. Besides, doubtful transactions are transformed into human investigators for the last evaluation. Moreover, the investigator task includes selection between investigate the transaction, carrying out a fast investigation on them, and hence, the alert is considered as fraud transaction or legal transaction. If any transaction cases are considered fraudulent, then the consequent credit card is blocked-up. In the NRT stage, managing regulations are mainly considered [6][7]. Apart from this, fraud detection in online shopping systems is a recently developing issue. The banking system, fraud investigators, and electronic payment systems, like PayPal, Gpay, etc have an effective fraud detection system for preventing fraud performances. Using this CyberSource information, it is needed to identify abnormalities across a prototype of fraud activities, which suffer by alternations comparative to past. The best fraud identification structure should capable to recognize fraud transactions precisely and make identification in real-time transactions. Fraud detection methods are of two types, namely misuse detection and anomaly detection. Anomaly detection system fetches trained normal transactions and it utilizes various approaches for identifying new frauds. On the other hand, misuse fraud recognition scheme utilizes labelled transaction as fraud transaction or legal transaction, which is trained by dataset history. Therefore, the misuse discovery method involves a supervised learning approach, whereas the anomaly identification structure involves an unsupervised learning method [85].

The increase in online transactions leads to the growth of credit card transactions. Online credit card transactions caused an increase in online fraud. To reduce this, various credit card fraud detection techniques are developed. Credit card fraud can be classified into three main types, Credit Card Fraud, Bankruptcy Fraud, and Credit Application Fraud.

Usage of credit card transactions can be categorized into two types. The first one is a card- present, where all the transactions are happened by inserting the card will fall under this type. Due to the recent chip security feature, if a

merchant does not swipe a card with the chip feature, then the merchant is exclusively responsible if that transaction is fraudulent. If they used the chip security feature, then the issuer is accountable.

The second type is card not present. All e-commerce transactions fall under this group and they are the utmost vulnerable to credit card fraud. If a fraudulent transaction is a card not present, then the merchant absorbs the costs. The reason for this is because the merchant cannot accomplish the necessary security procedures to the best of their capacity. Since they cannot employ the security policies issues by the processor or the issuer, the merchant assumes all liability.

This chapter discussed the various credit card fraud detection methods that are already

2. Related work

2.1. Credit card Fraud detection based on deep learning

This section explains the credit card fraud detection methods using deep learning methods. Deep learning is an artificial learning method that learned and made decisions on its own. The credit card fraud detection methods using the deep learning methods are as follows, Zhou, X.-H.*et al.*[128] modelled a credit card fraud detection using a Generative adversarial network. This method detected the fraudsters using the inference method that was based on the discriminator model and generative model. The inference method differentiated the samples fitting effectively. The probabilistic relationship was learned in the inference model with the help of a deep denoising autoencoder. This method had a low rate of misclassification and high accuracy. The negative samples were separated from the positive samples by developing a minimax game using adversarial training. This method failed to provide a solution to the complex problems in classification and the efficiency was decreased for an increasing number of parameters.

In[8], developed a credit card detection method using data mining. This method detected the score of Fraud Suspicion by developing a risk scoring system. The risk scoring system was a machine learning model that approved the order automatically for the threshold value higher than the score and the order was manually approved for the threshold value lower than the score. This method detected the risk score and improved the manual revision process effectively. However, this method had problems in presenting the solutions in e-tail for fraud detection.

In[9], designed a semantic fusion method for fraud detection. This method developed a semantic fusion method by integrating the artificial bee colony algorithm (ABC) and the k-means algorithm. In the ABC, the inability in handling the real cluster was eliminated by combining the global search with the neighborhood search. The relevant features were determined using the rule engine and the unified frame was developed by combining the ABC optimizer and k-mean classifier with the optimized classifier. The behaviour of the customer like the frequency usage, geographical locations, and book balance was considered for determining the fraudulent. This method improved the accuracy in classification and convergence speed but the computational complexity was high.

In[10], developed a hybrid method by integrating the Genetic Algorithm and neural network (GANN) for fraud detection in credit cards. The training of the neural network was done for determining the parameters, like type of the network, weight, number of nodes, and layers using the BPN. This method used the fact that the success rate was high for the talented person. This method provided accurate detection of the fraud in the credit card. However, this method failed to determine the transactions in the credit card in advance.

In[11], designed a Convolutional Neural Networks (CNN) for the detection of fraud in the credit card. In this method, the intrinsic patterns were learned from the labelled data for determining the behaviour of fraud. Then, the feature matrix was obtained from the available transaction data. After that, the latent patterns set were obtained for the samples using CNN. This method predicted the credit card fraud accurately but the complexity in the computation is a major concern.

In[12], developed a homogeneity-oriented behaviour analysis (HOBA) for the detection of fraud in the credit card. The variables of the features used for the detection of credit card fraud were generated using HOBA. The behaviour in the transaction was analysed using deep learning methods. The performance was determined by considering the false-positive rates. This method provided good performance by reducing the losses in the fraud and reducing the regulatory costs. However, this method failed to reduce the cost of computation which is increased due to the increase in the variable set.

In[13], designed a credit card fraud detection method by the analysis of the behaviour of the user. This method provided two-level tracking of the credit card. Theprecision of the method was improved using the ABC algorithm

and k-means algorithm. Initially, normal behaviour was determined for the reliability of the fraud. Then, the fraudulent data were classified using the rule-based engine by finding the deviations from the normal data. The first-level classification was performed using the k-means algorithm whereas the second-level classification was done using the ABC algorithm. The k-means algorithm failed to evaluate the actual clusters which were overcome by the ABC algorithm. Finally, the KNN algorithm was used for the transaction matching and the closeness distance was evaluated using the incoming transaction. However, this method required the extra rules for enhancing the accuracy in the rule engine.

In[14], modelled a machine learning method for fraud detection. This method determined a large number of fraudulent transactions. The available transaction amount was determined by adjusting the cost matrix. This method provided good performance in the detection of fraud. However, this method failed to include the historical bankcard transactions for the accurate detection of fraudsters and also failed to consider the small frauds.

In[15], designed a Bayes minimum risk fraud detection model. The Bayes minimum risk used the quantifying tradeoffs for making decisions in the credit card fraud detection. Although this method minimized the cost, the savings needs to be further increased for a cost-sensitive system.

In[16], modelled an Optimized Light Gradient Boosting Machine (OLightGBM) for the detection of fraud in the credit card. In this approach, the LightGBM was incorporated with the hyper-parameter optimization algorithm based on Bayesian for parameter tuning. This method discriminated against both the fraudulent and legitimate transactions in the credit card. Initially, the data was pre-processed followed by the selection of features. The features were selected for reducing the dimensionality using the information gain method. The similarities were determined in the information gain method and the weights were assigned. The features with the greatest weight were considered for the evaluation of the best features. Then, the OLightGBM was used for the processing and handling of a large amount of data. The OLightGBM bundled the features into a bundle and created a histogram with the same feature depending on the feature bundles through the feature-scanning algorithm. Although this method enhanced the performance of the prediction method, it failed to reduce the complexity of computation.

In[17], modelled a Federated learning for Fraud Detection (FFD) for the detection of fraud. The updates were locally computed by aggregating the shared Fraud Detection System (FDS). In this method, the private data from the bank was not required to be sent to the data centre with the federated fraud detection method. The unavailability and the sensitivity of the dataset were influenced by the decentralized data. The information was learned using the parameters of the global shared method and by accessing the individual bank update. This method detected fraud in the credit card by overcoming the skew distribution problem. However, this method failed to aggregate the updates of the model in an efficient manner.

In[18], modelled a fraud detection method using the Restricted Boltzmann Machine and Deep Learning based on Auto-Encoder. For cleansing the data, the data attributes were classified and transformed into Principal Component Analysis (PCA) through XLSTAT. This method used the auto-encoder for setting the output and the input equivalent using the backpropagation algorithm. This method detected the maximum frauds in the system with good accuracy. The input was encoded and decoded to the output through the hyperbolic tangent function. The backpropagation was realized using the parameter gradients with the help of an auto-encoder. This method provided accurate detection but the computational complexity was high.

In[19], designed a game theory-based approach for the detection fraudulent transactions. The normalized scores were assigned into the individual rule for the selection to quantify the influence of the rule in the pool performance. The performance was summarized by collaborating the Shapley Value and the Coalitional Game Theory for generating the power-index. The score was determined for predicting the maintenance of the rule in the assessment process of the periodic rule and the compact rule-set was used for selecting the rules that were top-ranked. The individual rule's historical performance was done by selecting the rules for maintaining the operation in the NRT pool. This method achieved moderate precision and high recall rate but failed to assign the rules with the normalized score for summarizing the performance.

In[20],developed detection method using the Dataset Shift Quantification. The covariate shift was quantified in the dataset by classifying the transactions of every day. For the efficient classification, there was a presence of covariate shift and the days were different. The distance matrix was evaluated between the days using the agglomerative clustering algorithm by matching the calendar with the shift pattern for particular periods of time. Then, the detection process was improved by considering the knowledge of the dataset shift as the significant feature. This method improved the performance but the recall and precision measure were improved only for small percentage.

In[21],developed a fraud detection method in credit cards based on machine learning. This method notified the end-user through GUI with the API module for the detection of the fraudulent transaction. After the detection of the suspicious transaction, the decision was moved into the next step for the investigation of fraud. This method detected the skewness in the data distribution effectively but failed to detect the location of the occurrence of fraud.

In[22],designed a fusion model for the identification of fraud in the credit card. Initially, the under-sampling method was used for pre-processing the data. In this method, the XGBoost and Lasso-Logistic were combined and optimized for predicting the fraud in the credit card. The drawback of the XGBoost method was the retaining of the variables that were unnecessary whereas the Lasso-Logistic was slower. These drawbacks of both the XGBoost and Lasso-Logistic method were rectified by fusing both the models. This method provided good extrapolation thus providing better classification accuracy for the prediction.However, the fusion of the methods leads to an increase in computational complexity.

3. Conclusion

In this day and age, the idea that is referred to as "Digitalization" has a significant amount of weight on the minds of today's youth. This understanding of digitalization is playing an extremely important role in all aspects of the banking industry, as well as the financial sector, the insurance industry, and other areas of the economy. In general, moving towards digitalization is important for the Indian banking sector because it plays a most significant role in financial inclusion, which is mostly disturbed for the purpose of offering the best services to customers with the possibility of gaining more in the future. When it comes to transferring money from one bank account to another, online banking has largely established itself as the industry standard. Online banking is gradually gaining popularity, which increases the number of online transactions with improved facilities in a variety of domains such as the payment of insurance premiums, online reservation for public transportation (buses and trains), payment of utility bills (electricity, house, and water taxes), online shopping, and so on. The effectiveness of using the internet for banking is continually improving. On the other hand, this development has also one major drawback such as an increase in fraudulent activities. [Citation needed] [Citation needed] [Citation needed] [Citation needed] [Online banking, also known as e-banking or internet banking, is a relatively recent innovation that has experienced rapid expansion in recent years. Internet banking has become so widespread in recent years that even the average person needs to have access to it. People now have the ability to interact with their banking accounts through the use of an innovative banking service known as electronic banking, which is accessible via the internet. Electronic banking enables customers to take advantage of a wide range of banking services, including those offered by Automatic Teller Machines (ATMs), direct deposits, electronic transfers of funds (EFT), and automatic bill payments (ABP), among others.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest.

References

[1] Mbama C I and Ezepue P O, “Digital banking, customer experience and bank financial performance”, International Journal of Bank Marketing, April 2018.

[2] AleksandarLukic, “Benefits and Security Threats in Electronic Banking International”, Journal of Managerial Studies and Research, vol.3, no.6, pp.44-47, 2015.

[3] Revathi P, “Digital Banking Challenges and Opportunities in India”, EPRA International Journal of Economic and Business Review, vol.7, no.12, pp.20-3, 2019.

[4] Nayak R, “A Conceptual Study on Digitalization of Banking-Issues and Challenges in Rural India”, International Journal of Management, IT and Engineering, vol.8, no.6, pp.186-91, 2018.

[5] Dagada R, “Digital banking security, risk and credibility concerns in South Africa”, In proceedings of The Second International Conference on Cyber Security, Cyber Peacefare and Digital Forensic, 2013.

[6] Achituve I, Kraus S, Goldberger J, “Interpretable Online Banking Fraud Detection Based on Hierarchical Attention Mechanism”, In proceedings of 2019 IEEE 29th International Workshop on Machine Learning for Signal Processing (MLSP), pp.1-6, October 2019.

[7] Wei W, Li J, Cao L, Ou Y, Chen J, "Effective detection of sophisticated online banking fraud on extremely imbalanced data", World Wide Web, vol.16, no.4, pp.449-75, July 2013

[8] Singh P and Singh M, "Fraud detection by monitoring customer behavior and activities", International Journal of Computer Applications, vol.111, pp.11, January 2015.

[9] Abdelhamid D, Khaoula S, Atika O, "Automatic bank fraud detection using support vector machines", In proceedings of The International Conference on Computing Technology and Information Management (ICCTIM), pp.10, January 2014.

[10] Taha A and Malebary S J, "An Intelligent Approach to Credit Card Fraud Detection Using an Optimized Light Gradient Boosting Machine", IEEE Access, vol.8, pp.25579-87, February 2020.

[11] Gianini G, Fossi L G, Mio C, Caelen O, Brunie L, Damiani E, "Managing a pool of rules for credit card fraud detection by a Game Theory based approach", Future Generation Computer Systems, vol.102, pp.549-61, January 2020.

[12] Zhu H, Liu G, Zhou M, Xie Y, Abusorrah A, Kang Q, "Optimizing Weighted Extreme Learning Machines for Imbalanced Classification and Application to Credit Card Fraud Detection", Neurocomputing, May 2020.

[13] Pumsirirat A and Yan L, "Credit card fraud detection using deep learning based on auto-encoder and restricted boltzmann machine", International Journal of advanced computer science and applications, vol.9, no.1, pp.18-25, January 2018.

[14] Omariba Z B, Masese N B, Wanyembi G, "Security and privacy of electronic banking", International Journal of Computer Science Issues (IJCSI), vol.9, no.4, pp.432, July 2012.

[15] Darwish S M, "A bio-inspired credit card fraud detection model based on user behavior analysis suitable for business management in electronic banking", Journal of Ambient Intelligence and Humanized Computing, vol.10, pp.1-5, February 2020.

[16] Belás J, Korauš M, Kombo F, Korauš A, "Electronic banking security and customer satisfaction in commercial banks", Journal of security and sustainability issues, 2016.

[17] Gąsiorowski J, "Managing security in electronic banking–legal and organisational aspects", In Forum Scientiae Oeconomia, vol.4, no.1, pp.123-136, 2016.

[18] Yazdanifard R, WanYusoff W F, Behora A C, Sade A B, "Electronic banking fraud: The need to enhance security and customer trust in online banking", Advances in Information Sciences and Service Sciences, vol.3, no.10, pp.505-9, 2011.

[19] Claessens J, Dem V, De Cock D, Preneel B, Vandewalle J, "On the security of today's online electronic banking systems", Computers & Security, vol.21, no.3, pp.253-65, June 2002.

[20] Mohammadi S and Abedi S, "ECC- biometric signature: A new approach in electronic banking security", In proceedings of 2008 International Symposium on Electronic Commerce and Security, pp.763-766, August 2008.

[21] Thamizhchelvy K and Geetha G, "E-banking security: Mitigating online threats using message authentication image (MAI) algorithm", In proceedings of 2012 International Conference on Computing Sciences, pp.276-280, September 2012.

[22] Quah J T and Sriganesh M, "Real-time credit card fraud detection using computational intelligence", Expert systems with applications, vol.35, no.4, pp.1721-32, November 2008.