WJCMS

Wasit Journal of Computer and Mathematics Science

Journal Homepage: https://wjcm.uowasit.edu.iq/index.php/WJCM

e-ISSN: 2788-5879 p-ISSN: 2788-5879

Fraud Detection and Identification in Credit Card Based on Machine Learning Techniques

Omega John Unogwu[1,*] and Youssef Filali[2]

[1]Department of Computer Science and Engineering, Universidad Azteca Chalco, Mexico
[2]EIGSI, La Rochelle-Casablanca, 17041-204010, France-Morocco

*Corresponding Author: Omega John Unogwu

DOI: https://doi.org/10.31185/wjcm.185
Received: May 2023; Accepted: Septemper 2023; Available online: Septemper 2023

ABSTRACT: Fraudulent internet transactions have caused considerable harm and losses for both people and organizations over time. The growth of cutting-edge technology and worldwide connectivity has exacerbated the rise in online fraud instances. To offset these losses, robust fraud detection systems must be developed. ML and statistical approaches are critical components in properly recognizing fraudulent transactions. However, implementing fraud detection models presents challenges such as limited data availability, data sensitivity, and imbalanced class distributions. The confidentiality of records adds complexity to drawing inferences and constructing improved models in this domain. This research explores multiple algorithms suitable for classifying transactions as either genuine or fraudulent using the Credit Card Fraud dataset. Given the extremely unbalanced nature of the dataset, the SMOTE approach was used for oversampling to alleviate the class distribution imbalance. In addition, feature selection was carried out, and the dataset was divided into training and test data. The experiments utilized NB, RF, and MLP algorithms, all of which demonstrated high accuracy in detecting credit card fraud. MLP method achieved 99.95% accuracy as compared to other methods.

Keywords: Multilayer Perceptron (MLP), Machine Learning (ML), Fraud, Random Forest (RF), Transaction, Naïve Bayes (NB).

1. INTRODUCTION

Financial institutions now provide the general public with practical business facilities, and internet banking is becoming more and more common. In today's cutthroat financial environment, e-payment systems have become indispensable since they make buying products and services easier. Credit and debit cards have transformed consumer convenience by enabling cashless transactions and providing insurance against things that have been damaged, lost, or stolen. An additional degree of security is added by requiring customers to confirm transactions with businesses before using their credit cards. However, credit card theft continues to be a problem, costing both financial institutions and consumers a lot of money. Fraudsters constantly develop new methods to take advantage of internet transactions, making it difficult for banks and other organizations to identify and stop fraudulent activity [1]. Because of this, these organizations are always looking for improved ways to win over customers' confidence and protect their operations. Although credit cards have many benefits, security and fraud are still issues. Unauthorized individuals who use credit cards fraudulently to obtain money are committing credit card fraud. Identity theft techniques or unprotected online platforms can be used to steal sensitive credit card information. Financial security is seriously threatened by fraudsters' ability to obtain consumers' credit and debit card data without their knowledge or agreement.

When unauthorized individuals use stolen credit card or account information to engage in illicit actions, credit card

fraud, a form of identity theft, happens. Fraud can happen when credit cards are misplaced, stolen, or counterfeited, as well as when card numbers are used for e-commerce purchases in card-not-present situations, especially with the development of online shopping [2]. Finding credit card theft has become a top priority as more and more people switch to digital payments. Businesses must adjust to the changing environment in this cashless age. Given that customers are increasingly prioritizing the simplicity and security of debit and credit card payments, conventional payment methods may no longer be sufficient. Companies will face a difficult environment in the future where they must adapt their payment systems to accommodate different payment methods and stay competitive. Uncomfortably, incidences of credit card fraud are on the rise. Additionally, there has been a noticeable increase in fraudulent activity involving new credit card accounts, necessitating the urgent need for efficient solutions. Payment card theft caused huge damages to the world economy of $24.26 billion, with the United States bearing the burden of 38.6% of the total card fraud losses in 2018 [3]. Automated fraud detection systems must be implemented in financial institutions due to the fragility of financial systems.

Supervised credit card fraud detection uses machine learning (ML) to address this sneaky menace. The goal is to separate fraudulent from valid transactions by building a strong ML model based on transactional credit card data, enabling prompt and precise decision-making. However, there are several obstacles on this trip. Rapid response, cost sensitivity, and rigorous feature pre-processing are all requirements for the system. The brain of this project turns out to be machine learning (ML), which uses historical data trends to create wise predictions. Financial institutions and businesses alike must be watchful and equipped with cutting-edge technology to safeguard their assets and consumers from the constantly expanding world of credit card fraud as the digital world intertwines with the fight against fraud. The adaptability of ML models in addressing numerous problems across many industries has been demonstrated [4]. In particular, DL algorithms have made substantial progress in a variety of applications, including video surveillance, banking, insurance, intrusion detection, and the prediction of heart disease. In this article, we investigate how machine learning, and more specifically, deep learning algorithms, can be used to identify credit card thefts in the banking sector. The SVM is a well-known supervised method for categorizing data among the ML techniques used to tackle these problems. Because SVM can handle both nonlinear and linear binary classification problems, it has achieved success in a variety of fields, including image identification and public safety. It is the best classifier because it successfully isolates input data using a hyperplane in the support vector. In the past, neural networks were the cutting-edge technique used to detect credit card theft. However, in recent years, attention has switched to DL strategies, which have produced encouraging results in a number of applications, including computer vision, natural language processing, and speech recognition [5].

Physical card theft and theft of sensitive card data, such as card numbers, CVV codes, and card kinds, are the two main categories of credit card fraud. Before the cardholder is aware of the fraud, thieves use stolen information to carry out unlawful transactions. Companies use a variety of ML approaches to separate fraudulent transactions from real ones in order to counteract these frauds. This research compares and analyses different machine learning (ML) algorithms, such as RF, NB, and MLP, to find the most effective approach for detecting credit card fraud. The goal of the research is to develop an efficient model that can recognize fraudulent transactions with accuracy, allowing businesses to guard against financial loss and shield their clients against fraud.

The rest of the article is organized as follows: We examine related works in the field in Section 2. Section 3 presents the various ML techniques. The results and analysis of the experiment are presented in Section 4. Finally, we summarize our findings in Section 5.

2. LITERATURE REVIEW

The goal of fraud detection is to identify valid and fraudulent credit card transactions, which poses a classification difficulty. Several research publications on this topic have been evaluated, and their important conclusions are described below.

Dhrawa et al. [6] proposed a novel method for detecting fraudulent online banking activities. The TRSGM algorithm is the product of their method, which integrates artificial intelligence, statistics, and data mining techniques. This system analyses transaction error scores by focusing on five major aspects. The system performed effectively for transactions that mimicked regular user behavior, as seen by low error scores. The error score was greater for transactions that deviated significantly from typical user behavior.

Taha et al. [7] undertook study to address the rising rate of credit card fraud caused by the rise of e-commerce and communication technology. They proposed an improved method based on Light Gradient Boosting Machine. For their examination, the researchers used two real-world public datasets containing both fraudulent and non-fraudulent transactions. In terms of accuracy, their proposed method surpassed rival systems, earning a remarkable accuracy rate of 98.40%, an AUC of 92.88%, a precision rate of 97.34%, and an F1-score of 56.95%. In terms of precision, they outperformed rival techniques, with an AUC of 92.88%.

Krivko et al. [8] proposed a unique fraud detection system based on rules and behavioral models that combines

supervised and unsupervised approaches. The number of true positive (TP) frauds detected and the number of positive prediction errors (PF) produced were used to evaluate this system.

Makki et al. [9] concentrated on combatting credit card fraud, which causes substantial financial losses. They discovered that present fraud detection systems can be costly, time-consuming, and labor-intensive. The imbalance in the dataset was noted by the researchers as a primary factor contributing to erroneous outcomes. Because of the presence of uneven data, these imbalanced classifications result in erroneous predictions and financial losses. To address this problem, the authors tested numerous algorithms and discovered that SVM, C5.0 decision tree, and ANN fared the best in terms of accuracy. They trained their models using a balanced dataset in order to increase fraud detection accuracy.

Bhusari et al. [10] introduced a secret Markov model-based method for detecting credit card fraud that is based on modeling user behavior. In this strategy, the decision-making process comprises determining the chance of a new transaction belonging to a given user. When compared to other existing systems, the researchers say that their technology is capable of effectively processing a large volume of transactions.

Randhawa et al. [11] investigated credit card fraud detection using ML algorithms. They tested the effectiveness of their model using publicly available credit card datasets. The trial findings showed that their proposed algorithm functioned effectively in assessing credit card creditworthiness.

Li et al. [12] proposed Kernel-based Super-vised Hashing, a fraud detection system that uses the principle of approximate nearest neighbor for huge datasets with high-dimensional data. This novel application of KSH for prediction outperforms existing methods.

Mittal et al. [13] used common supervised and unsupervised machine learning methods to detect credit card transaction fraud on a highly imbalanced dataset. Their results showed that unsupervised ML algorithms were effective in detecting fraud in the banking system and improving information classification.

Kumar M. et al. [14] used RF approaches to construct a fraud detection model for credit card transactions, a supervised ML approach that employs decision trees for categorization. A confusion matrix was used to test the model's performance, and it attained an accuracy of 90% [10].

Sailusha et al. [15] analyzed the outcomes of two ML systems used to detect credit card transaction fraud. The Adaboost algorithm stands out as being extremely effective at detecting credit card transaction fraud.

Rao et al. [16] presented their findings on credit card fraud detection using several machine learning algorithms. They acquired and examined data from many journals in this topic.

Akila et al. [17] proposed an ensemble-based misrepresentation location algorithm to control unbalanced datasets and prevent noise in transactions. The bagging model efficiently addressed the dataset's anomalies and non-vital features. Bagging and risk-based base learners were used, with Nave Bayes resolving noise-related difficulties in transactions. Using the NRBE methodology, the suggested model outperformed other methods, generating significant gains in Balanced Classification Rate and Balanced Error Rate, with 50% enhanced recall and reduced cost for fraud detection (2x to 2.5x).

Hussein et al. [18] proposed a novel strategy for detecting credit card fraud that combines numerous classification algorithms based on the fuzzy method. Their research comparing this combination technique to seven other algorithms revealed that its use could greatly minimize the occurrence of fraud in financial transactions.

Sohony et al. [19] proposed an ensemble learning technique for detecting credit card fraud that takes into account the appropriate fraud-to-normal transaction ratio. They discovered that RF had higher accuracy, but neural networks were effective at detecting fraud. The researchers used a combination of RF and neural networks to conduct experiments on big real-world credit card transaction datasets.

Kumar P. et al. [20] investigated many ways for detecting MasterCard fraud using machine learning algorithms. They used relevant criteria to assess the performance of various strategies. Extensive research in this arena indicated a need for more efficient systems that function effectively in a variety of conditions.

Bazczyski et al. [21] investigated a novel dataset relating to auto loan applications and applied the Dominance-Based Rough Set Balanced Rule Ensemble approach, which had not previously been investigated for financial fraud prediction. When they compared their methodology to existing tools for forecasting financial fraud, they discovered many advantages in their proposed strategy.

Zainab et al. [22] examined ML models, with a specific focus on boosting algorithms, to test their accuracy in predicting fraudulent transactions. The study discovered that boosting algorithms outperformed other standard models, producing more precise and trustworthy fraud prediction results.

Zamini et al. [23] created an unsupervised model for clustering that employs autoencoders. When tested on the European dataset, their model beat other contemporary systems with three hidden layers and k-means clustering.

Sarma et al. [24] proposed a unique technique for detecting bank fraud by identifying patterns linked with fraudulent occurrences using a community detection algorithm. They created a web-based service that acted as a central hub between banks and clients, allowing for more rapid fraud detection.

Husejinovic et al. [25] examined the predictive performance of NB, C4.5, and ML algorithms in routine and fraudulent transactions. The ML algorithms performed admirably in differentiating binary class 0 in the dataset, with PRC rates ranging from 0.999 to 1.000. The C4.5 algorithm performed the best at predicting fraud transactions, with a success percentage of 92.74%.

Nayak et al. [26] introduced a machine learning approach for efficient fraud detection in online transactions, highlighting the significance of discriminating between real and fraudulent transactions. They used multiple ML techniques and graphed the results to effectively highlight the interdependencies within the data.

Jiang et al. [27] suggested a multi-stage process for detecting credit card fraud. They first gather cardholder transactions and then aggregate the data based on behavioral patterns. Following that, the dataset is labeled, and a model is trained and tested. In the event of anomalous behavior, a feedback mechanism is utilized to inform the system about the abnormality.

Ananthu et al. [28] concentrated on detecting fraudulent transactions through the analysis of previous transaction records. They wanted to combine big data analytics with machine learning techniques to detect fraud in massive real-time datasets quickly and efficiently. In addition, the article gave a quick comparison of several ML algorithms applied to huge data.

3. METHODOLOGY

3.1 DATASET DESCRIPTION

The research utilized the Credit Card Fraud Detection dataset, available for download from Kaggle [29]. This dataset comprises two days' worth of transactions conducted by European cardholders in September 2013. The dataset consists of 31 numerical features, some of which contain sensitive financial information. To preserve data anonymity, the PCA transformation was applied to these specific input variables. Three aspects, however, remained unaltered. "Time" represents the time difference between the dataset's initial and subsequent transactions. "Amount" is the transaction amount entered by credit card users. The label is provided by the "Class" feature, which has two values: 1 for fraudulent transactions and 0 for authentic transactions. Only 492 transactions out of a total of 284,807 were categorized as fraudulent, demonstrating a severely skewed dataset in which only 0.173% of transactions are tagged as fraudulent. Given the large imbalance in class distribution, data pretreatment is critical to maintain model quality and precision throughout future analysis and ML.

3.2 PREPROCESSING

Feature selection is a crucial technique that identifies the most relevant variables in a dataset, aiding in reducing overfitting, improving accuracy, and reducing training time. To achieve this, the Feature Selector tool [30] by Will Koehrsen was employed in the experiment, helping determine the most important features. Features contributing less than 95% cumulative importance were removed, leaving 27 features for further experimentation. Given the highly imbalanced data, ML algorithms struggle to learn efficiently. To overcome this, class distribution balancing strategies such as undersampling the dominant class, oversampling the minority class, or a mix of the two were used. The Synthetic Minority Oversampling (SMOTE) [31] was a popular oversampling approach that improved on the random oversampling methodology. As machine-learning algorithms typically expect inputs to be on a similar scale, scaling was performed to bring all features to a consistent level of magnitudes, especially for variables with highly varying values like "time" and "amount."

The experiment was carried out using a Windows 10 OS with the Spyder scientific Python programming environment, which is part of the Anaconda platform. Essential libraries used in the experiment include numpy, pandas, sklearn, and matplotlib.

3.2.1. MODELS

Naïve Bayes:

NB is a Bayesian probabilistic machine learning algorithm. It is feature independent, which makes it computationally efficient and well-suited for high-dimensional data. Despite its simplicity, NB has demonstrated efficacy in a variety of applications, including text classification and spam filtering. Because of its speed, efficiency, and ability to process high-dimensional data, NB is utilized in credit card fraud detection. It is especially well-suited for processing high-volume transactions in real time and can adapt to skewed datasets. Despite its simplicity, NB produces interpretable findings that help analysts understand its decision-making process. Fraudulent transactions are identified using the Bernoulli distribution.

Random Forest

RF is an ensemble learning strategy that combines many decision trees to improve classification and regression issues.

By pooling predictions from individual trees, it reduces overfitting and improves accuracy. RF is a robust, scalable, and widely utilized technology in a variety of fields including as classification, regression, and anomaly detection. Because of its high accuracy, robustness against overfitting, and capacity to handle huge and complicated datasets, RF is a powerful ML technique used in credit card fraud detection. It increases generalization and identifies crucial factors for fraud classification by combining numerous decision trees. The ensemble technique improves the model's detection of fraudulent transactions while being computationally efficient for real-time applications.

Multilayer Perceptron

The MLP is a feedforward artificial neural network with three layers: input, hidden, and output. Each node in the network employs an activation function that computes the weighted sum of its inputs and incorporates a bias. This configuration helps us to determine which neurons should be deleted in order to simplify the network's external connections. Using the ReLU activation function, we used an MLP with four hidden layers of 50, 30, 30, and 50 units, respectively. According to study, deeper networks produce better results than shallower ones. We utilized Adam, a stochastic gradient-based optimizer, to optimize the network's weights. The optimum hyper-parameters were chosen after thorough investigation, and increasing the network's complexity did not greatly enhance results while significantly increasing processing time. Because of its ability to learn complicated patterns from data, MLP(MLP) is a neural network architecture often used in credit card fraud detection. MLP can capture deep correlations and nonlinearities in data with several hidden layers, improving fraud detection accuracy. The capacity of the algorithm to modify and update weights during training ensures that it remains current with evolving fraud strategies. However, as compared to other algorithms, its training process may necessitate more data and processing.

The dataset was divided into a training set (80%) and a test set (20%) for model training and evaluation. The model was updated over several epochs, using a specified tolerance for optimization (TOL). Convergence was assessed by monitoring the loss or score, and if it did not improve by at least TOL for three consecutive iterations, training was terminated as convergence was reached. This approach ensures the model stops training once it reaches a point of diminishing returns in terms of improvement.

4. RESULT ANALYSIS

In the pursuit of identifying the most suitable algorithm for fraud transaction detection, various evaluation criteria have been employed, with accuracy, recall, and precision being the most commonly used metrics. These metrics are derived from a Confusion Matrix [Fig. 1], enabling the assessment of model performance [Table 1]. Both original and oversampled data were used for testing the models, and the results emphasized the significance of data sampling. As the test set constitutes 20% of the entire dataset with a total of 56,962 samples, 98 transactions were flagged as fraud. The analysis of the outcomes reveals that accuracy is notably high, although it does not guarantee perfect results. The importance of interpreting accuracy in conjunction with other metrics is highlighted. MLP method got F1-Score and Accuracy as 0.9998, 99.95% respectively.

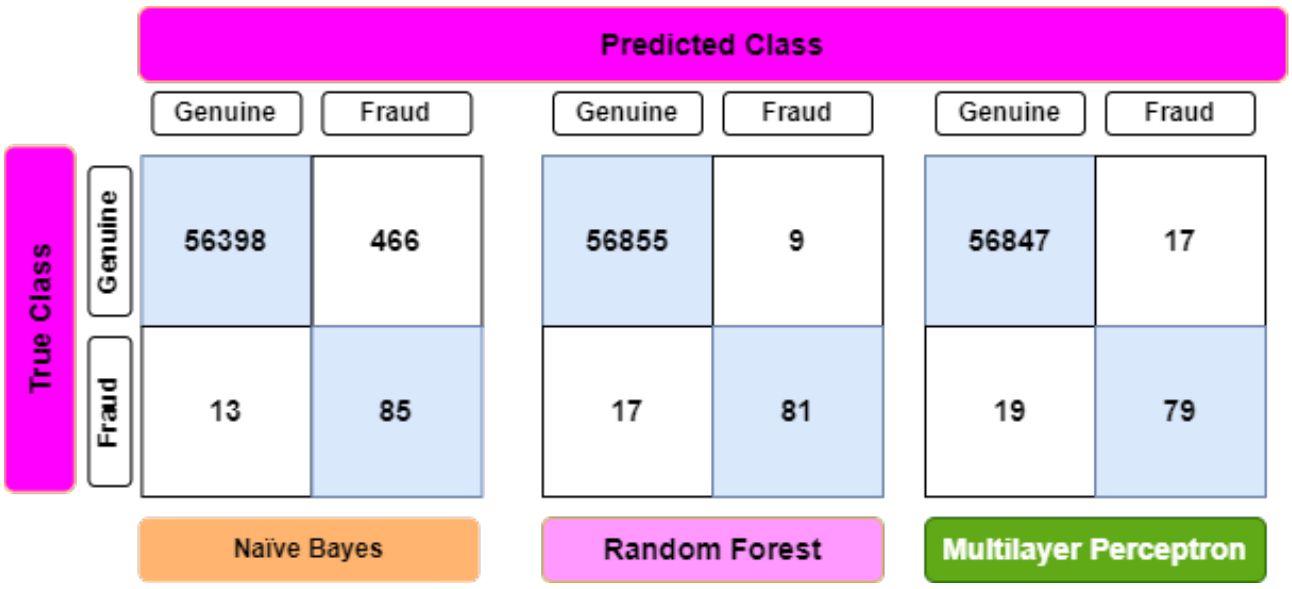

FIGURE 1. Confusion Matrix of all ML models

In further detail, the aforementioned metrics may be described as follows:

1. $Precision = \frac{True\ Positive\ (TP)}{True\ Positive + False\ Positive\ (FP)}$
2. $Sensitivity\ or\ Recall = \frac{True\ Positive}{True\ Positive + False\ Negative\ (FN)}$
3. $F1 - Score = \frac{2\ x(Precision\ x\ Recall)}{(Precision + Recall)}$
4. $Specificity = \frac{True\ Negative\ (TN)}{True\ Negative + False\ Positive}$
5. $Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$

Table 1. Performance Metrics for all ML models mple

Models	Recall	Precision	F1-Score	Specificity	Accuracy (%)
NB	0.9998	0.9918	0.9958	0.1543	99.16%
Random Forest	0.9997	0.9997	0.9997	0.8229	99.94%
Multilayer Perceptron	0.9997	0.9998	0.9998	0.9000	99.95%

5. CONCLUSION

In recent years, the rise in financial fraud, credit card fraud, and money laundering has garnered significant concern. Detecting and preventing such frauds has become a top priority for companies to avoid substantial losses. This paper aimed to compare various ML algorithms for fraudulent transaction detection. The MLP method achieved remarkable performance with an F1-Score of 0.9998 and an accuracy of 99.95%, showcasing its exceptional ability to accurately detect and classify credit card fraud transactions. These impressive results emphasize the effectiveness of the MLP algorithm in identifying fraudulent activities and its potential for bolstering fraud detection systems. Further research could explore alternative ML algorithms like genetic algorithms and various stacked classifiers, accompanied by extensive feature selection to enhance results.

FUNDING

None

ACKNOWLEDGEMENT

None

CONFLICTS OF INTEREST

The author declares no conflict of interest.

REFERENCES

[1] Y. Abakarim, M. Lahby, and A. Attioui, "An efficient real-time model for credit card fraud detection based on deep learning," *Proceedings of the 12th International Conference on Intelligent Systems: Theories and Applications*, pp. 1–7, 2018.
[2] S. B. E. Raj and A. A. Portia, "Analysis on credit card fraud detection methods," *2011 International Conference on Computer, Communication and Electrical Technology (ICCCET)*, pp. 152–156, 2011.
[3] V. N. Dornadula and S. Geetha, "Credit card fraud detection using machine learning algorithms," *Procedia Computer Science*, vol. 165, pp. 631–641, 2019.
[4] J. O. Awoyemi, A. O. Adetunmbi, and S. A. Oluwadare, "Credit card fraud detection using machine learning techniques: A comparative analysis," *2017 International Conference on Computing Networking and Informatics (ICCNI)*, pp. 1–9, 2017.
[5] H. Najadat, O. Altiti, A. A. Aqouleh, and M. Younes, "Credit card fraud detection based on machine and deep learning," *2020 11th International Conference on Information and Communication Systems (ICICS)*, pp. 204–208, 2020.
[6] J. N. Dharwa and A. R. Patel, "A data mining with hybrid approach based transaction risk score generation model (TRSGM) for fraud detection of online financial transaction," *International Journal of Computer Applications*, vol. 16, no. 1, pp. 18–25, 2011.
[7] A. A. Taha and S. J. Malebary, "An intelligent approach to credit card fraud detection using an optimized light gradient boosting machine," *IEEE Access*, vol. 8, pp. 25579–25587, 2020.
[8] M. Krivko, "A hybrid model for plastic card fraud detection systems," *Expert Systems with Applications*, vol. 37, no. 8, pp. 6070–6076, 2010.
[9] S. Makki, Z. Assaghir, Y. Taher, R. Haque, M. S. Hacid, and H. Zeineddine, "An experimental study with imbalanced classification approaches for credit card fraud detection," *IEEE Access*, vol. 7, pp. 93010–93022, 2019.
[10] V. Bhusari and S. Patil, "Application of hidden Markov model in credit card fraud detection," *International Journal of Distributed and Parallel Systems*, vol. 2, no. 6, pp. 203–203, 2011.
[11] K. Randhawa, C. K. Loo, M. Seera, C. P. Lim, and A. K. Nandi, "Credit card fraud detection using AdaBoost and majority voting," *IEEE Access*, vol. 6, pp. 14277–14284, 2018.

[12] Z. Li, G. Liu, S. Wang, S. Xuan, and C. Jiang, "Credit card fraud detection via kernel-based supervised hashing," *2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation*, pp. 1249–1254, 2018.
[13] S. Mittal and S. Tyagi, "Performance evaluation of machine learning algorithms for credit card fraud detection," in *9th International Conference on Cloud Computing*, pp. 320–324, 2019.
[14] M. S. Kumar, V. Soundarya, S. Kavitha, E. S. Keerthika, and E. Aswini, "Credit card fraud detection using random forest algorithm," *2019 3rd International Conference on Computing and Communications Technologies (ICCCT)*, pp. 149–153, 2019.
[15] R. Sailusha, V. Gnaneswar, R. Ramesh, and G. R. Rao, "Credit card fraud detection using machine learning," *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, pp. 1264–1270, 2020.
[16] R. Rao, H. R. Rao, N. V. Kumar, and M. Kolla, "A study on machine learning approaches to detect credit card fraud," *AIP Conference Proceedings*, vol. 2358, pp. 100008–100008, 2021.
[17] S. Akila and U. S. Reddy, "Credit card fraud detection using non-overlapped risk-based bagging ensemble (NRBE)," *2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, pp. 1–4, 2017.
[18] A. S. Hussein, R. S. Khairy, S. M. M. Najeeb, and H. T. Alrikabi, "Credit card fraud detection using fuzzy rough nearest neighbor and sequential minimal optimization with logistic regression," *International Journal of Interactive Mobile Technologies*, vol. 15, no. 5, pp. 24–42, 2021.
[19] I. Sohony, R. Pratap, and U. Nambiar, "Ensemble learning for credit card fraud detection," *Proceedings of the ACM India Joint International Conference on Data Science and Management of Data*, pp. 289–294, 2018.
[20] P. Kumar and F, "Credit card fraud identification using machine learning approaches," *2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT)*, pp. 1–4, 2019.
[21] J. Błaszczyński, A. T. D. A. Filho, A. Matuszyk, M. Szeląg, and R. Słowiński, "Auto loan fraud detection using dominance-based rough set approach versus machine learning methods," *Expert Systems with Applications*, vol. 163, pp. 113740–113740, 2021.
[22] K. Zainab, N. Dhanda, and Q. Abbas, "Analysis of various boosting algorithms used for detection of fraudulent credit card transactions," *Information and Communication Technology for Competitive Strategies*, pp. 1083–1091.
[23] M. Zamini and G. Montazer, "Credit card fraud detection using autoencoder-based clustering," *2018 9th International Symposium on Telecommunications (IST)*, pp. 486–491, 2018.
[24] D. Sarma, W. Alam, I. Saha, M. N. Alam, M. J. Alam, and S. Hossain, "Bank fraud detection using community detection algorithm," *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, pp. 642–646, 2020.
[25] A. S. Husejinovic, R. S. Khairy, S. M. M. Najeeb, and H. T. Alrikabi, "Credit card fraud detection using naive Bayesian and C4.5 decision tree classifiers," *Periodicals of Engineering and Natural Sciences*, vol. 8, no. 1, pp. 1–5, 2020.
[26] H. D. Nayak, L. Anvitha, A. Shetty, D. J. Souza, and M. P. Abraham, "Fraud detection in online transactions using machine learning approaches-a review," *Advances in Artificial Intelligence and Data Engineering*, pp. 589–599, 2021.
[27] C. Jiang, J. Song, G. Liu, L. Zheng, and W. Luan, "Credit card fraud detection: A novel approach using aggregation strategy and feedback mechanism," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 3637–3647, 2018.
[28] S. Ananthu, N. Sethumadhavan, and H. N. Ag, "Credit card fraud detection using Apache Spark analysis," *5th International Conference on Trends in Electronics and Informatics (ICOEI)*, pp. 998–1002, 2021.
[29] C. Kaggle *Credit Card Fraud Detection*, 2019.
[30] Github *Feature selector*, 2019.
[31] A. Fernández, S. Garcia, F. Herrera, and N. V. Chawla, "SMOTE for learning from imbalanced data: Progress and challenges, marking the 15-year anniversary," *Journal of Artificial Intelligence Research*, vol. 61, pp. 863–905, 2018.