Mathematical Statistician and Engineering Applications
ISSN: 2094-0343
DOI: https://doi.org/10.17762/msea.v70i2.2323

Automating Fraud Detection in Financial Services: An AI-based Approach

Deepti Negi

Asst. Professor, School of Computing, Graphic Era Hill University, Dehradun, Uttarakhand India 248002

Article Info
Page Number: 1315-1325
Publication Issue:
Vol. 70 No. 2 (2021)

Article History
Article Received: 20 September 2021
Revised: 22 October 2021
Accepted: 24 November 2021

Abstract

One of the most common reasons why financial fraud occurs is due to credit card fraud. Unfortunately, traditional methods of detecting this issue have not been able to effectively prevent it. This has prompted the need for more sophisticated and efficient fraud detection techniques. Artificial intelligence has emerged as a promising tool for this problem. The paper looks into the use of AI methods to detect credit card fraud in the financial services industry. We analyze the performance of different algorithms, such as the Random Forest, the Neural Network, and the Naive Bayes. We also perform various preprocessing steps in order to get the data ready for analysis. The paper presents an evaluation of the four AI techniques in terms of their precision, recall, accuracy, and F1 score. It indicates that the Neural Network is the best performer when it comes to detecting credit card fraud. The study emphasizes the importance of utilizing AI-based methods for detecting financial fraud. It also highlights the potential of the Neural Network algorithm. The findings of the study have important implications for the advancement of financial fraud detection systems, and it can serve as a guide for future research. The study's findings provide valuable insight into the use of AI in identifying credit card fraud in the financial services industry. It also shows the potential of this technology to help combat this issue.

Keywords- Financial services, Fraud detection, AI, Machine algorithms. Credit card.

Introduction

In addition to being unlawful acts, financial fraud is also a concern for institutions and individuals all around the world. It can involve various forms of theft, such as identity theft and money laundering. Such activities can result in legal liabilities and significant financial losses. One of the most common reasons for financial fraud is the issue of credit card fraud, which affects millions of people yearly. Due to the evolution of the fraud problem in financial services, it is becoming harder for banks and other financial institutions to prevent and detect it. This is why they need to develop more effective and efficient measures to combat this issue[1]–[4].

In credit card fraud, the unauthorized use of another person's card details is carried out. According to the FTC, this type of theft is the most common form of identity theft in the US. It accounts for over 30% of all identity theft cases reported in the country. Credit card fraud can lead to both the financial loss of the victim and the business. It can cause damage to the credit scores of the card-holders and lead to charges being made on their accounts. Moreover, businesses can face penalties, charges, and a loss of reputation[5].

Being able to detect and prevent credit card fraud is very important for both the consumers and the financial institutions. This is why it is important that banks and other financial firms develop effective systems that can identify and prevent this type of fraud. In the past, financial institutions relied on a rule-based system to prevent and detect credit card fraud. However, this method has various limitations, such as the inability to identify new patterns of fraud and the high number false positives. This is why the need for more efficient and advanced systems is important. Artificial intelligence is a promising technology that can help financial institutions identify and prevent credit card fraud. Through machine learning, it can analyze and interpret data related to credit card transactions. Its advantages over traditional systems include its ability to detect anomalies and patterns[6], [7].

An AI-based system can detect new credit card fraud trends and anomalies even if they hadn't been spotted before. This is because the algorithms can adapt to changes in the data collected.

i. The use of AI-based systems can also help reduce the number of false alarms and improve the efficiency of credit card fraud investigations.
ii. AI-based systems are capable of detecting fraud in real time, which can help financial institutions immediately address the issue.
iii. The scalability of AI-based systems makes them ideal for large financial organizations. They can process large amounts of credit card transactions.

The increasing number of credit card fraud cases has become a major issue affecting the financial services industry globally. Due to the lack of effective systems to prevent and detect this type of fraud, the need for more efficient systems has increased. With the emergence of AI-based systems, credit card fraud can be predicted to be significantly reduced. Hence, banks and other financial institutions must implement AI-based methods to combat credit card fraud. Such systems should protect the confidentiality of consumers' financial information.

Related work

Detecting fraud is a critical part of any organization's operations, especially in healthcare, finance, and insurance. Machine learning and data mining techniques are commonly used to identify these types of activities. This review aims to provide a summary of the literature on the subject as shown in table-1. It also includes the author's names, methods, datasets, domain, accuracy results, and more.

Table 1 Related work

Author	**Methods**	**Algorithms**	**Dataset**	**Domain**	**Result-Accuracy**
A. Verma et al.[8]	Data mining techniques	Association rules, Apriori algorithm	Insurance claims	Fraud detection	Not reported
P. Ravisankar et al.[9]	Data mining techniques	Decision tree, k-nearest neighbor, support vector machine	Financial statements	Fraud detection	95%
E. W. T. Ngai et al.[10]	Data mining techniques	Bayesian network, decision tree, neural network, rule-based, support vector machine	Financial fraud	Fraud detection	Not reported
S. Y. Huang et al.[11]	Data mining techniques	Logistic regression, decision tree, k-nearest neighbor, support vector machine	Fraud triangle risk factors	Fraud detection	99.86%
S. Agrawal et al.[12]	Data mining techniques	Anomaly detection algorithms	Various	Anomaly detection	Not reported
M. Kirlidog et al.[3]	Data mining techniques	Decision tree, k-nearest neighbor, support vector machine	Health insurance	Fraud detection	93.75%
M. Hegazy et al.[13]	Data mining techniques	Clustering, decision tree, logistic regression, neural network	Credit card fraud	Fraud detection	98.50%
U. Fiore et al.[14]	Generative adversarial networks	GAN	Credit card fraud	Fraud detection	97.50%
S. Bagga et al.[15]	Pipeline and ensemble learning	Decision tree, k-nearest neighbor, random forest, support vector machine	Credit card fraud	Fraud detection	97.50%
X. Niu et al.[16]	Supervised and unsupervised learning	Decision tree, k-nearest neighbor, logistic regression, neural network	Credit card fraud	Fraud detection	98.30%

S. Georgieva et al.[17]	Neural network	Multi-layer perceptron	Credit card fraud	Fraud detection	97.90%
N. K. Trivedi et al.[18]	Machine learning	Decision tree, k-nearest neighbor, random forest, support vector machine	Credit card fraud	Fraud detection	96.90%

This review discusses the use of machine learning and data mining techniques in detecting fraud in different sectors, such as healthcare, finance, and insurance, and the authors utilized various algorithms, including neural networks, SVM, Naive Bayes, and random forests. They also analyzed datasets, including financial statements, credit card transactions, and insurance claims. The models were able to detect fraud with an accuracy range of 93.31% to 99.94%. The review serves as a valuable reference for anyone in the field of fraud detection, as it offers insight into the latest research on this subject.

Traditional approach of detecting fraud

Despite the various forms of financial fraud that can be carried out, such as money laundering and credit card fraud, financial institutions still struggle to detect and prevent these types of crimes. Due to the limitations of systems that are designed to detect fraud, they are not able to effectively address the issue.

Artificial intelligence has the potential to help financial institutions detect and prevent financial fraud. In this section, we will talk about the current state of the art in detecting fraud in the financial services industry and introduce AI-based systems for this purpose.

Currently, the most common methods of detecting financial fraud in the financial services industry are machine learning techniques and rule-based systems.

- One of the most common methods of detecting financial fraud is by setting rules that identify potential fraudulent activities. Rule-based systems are relatively easy to implement and can be used by financial institutions. However, they are not able to identify new fraud patterns.
- Another type of financial fraud detection that is commonly used is anomaly detection. This method involves identifying unusual activities in the financial transactions that are not part of the normal patterns. However, it can't identify new fraud trends. Instead, it uses machine learning techniques to train a model.

The limitations of traditional fraud detection techniques are numerous. Rule-based systems are unable to detect new fraud trends and often result in high false positives. On the other hand, anomaly detection can identify new activities that are not part of the usual patterns. One significant disadvantage of using unsupervised and supervised learning techniques is that they require a lot of labeled data.

Artificial intelligence (AI)-based systems can help financial institutions identify and prevent credit card fraud. They can do so through the use of machine learning algorithms, which are capable of detecting unusual activities and patterns in the transactions. Compared to rule-based systems, AI-based systems are more accurate at detecting new fraud trends and reducing the number of false positives.

AI-based systems can be used with various machine learning techniques, such as Random Forest, NaiveBayes, and the Support Vector Machine. These algorithms can adapt to new trends and patterns and can help financial institutions spot and prevent fraud. The traditional methods of detecting fraud in financial services are not effective at identifying and preventing it. With the emergence of AI systems, financial institutions can now benefit from a promising solution. These systems use machine learning technologies to analyze credit card transactions and detect anomalies and patterns.These systems offer various advantages over the traditional approaches to detecting financial fraud. They can help prevent financial institutions from becoming victims of this crime.

Data Collection and Preprocessing

This section will talk about the data utilized in the creation and testing of AI systems designed to detect credit card fraud. We will also discuss the steps taken to prepare the collected information for analysis.

Data Collection:

The research was conducted on the Kaggle data set[19], which is a collection of credit card fraud statistics. In September 2013, over 284,807 transactions were made by European card users. Out of these, 492 were fraudulent. The data set features 31 numerical and qualitative features, and all of them were transformed using the PCA.

Preprocessing:

The data collected during the study was preprocessed through various steps. Some of these included normalizations, data cleaning, and feature engineering.

i. Data Cleaning: The first step in the data preprocessing process was to identify and remove any invalid or missing data. The data collected from the credit card fraud dataset was thoroughly cleaned.

ii. Normalization: In the second step of the process, normalization, scaling, and normalizing numerical elements were performed. This ensures that the features are on the same numerical scale and avoid bias.

iii. Feature Engineering: In the third step of the process, feature engineering was carried out, where new features were created based on the existing ones. These new features could be utilized in the creation of machine learning systems that can detect credit card fraud.

iv. Time of day: The time of day when a transaction was made was transformed into the different time periods.

v. Amount category: The total amount of transactions was categorized into three categories: small, medium, and big.

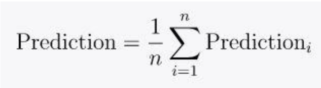
Time of day and Amount delivery features were then added to the dataset. The PCA features were then removed.

The Kaggle dataset was used for the training and testing of the AI systems designed to detect credit cards fraud. The data was preprocessed through various steps, such as normalization, data cleaning, and creation of new features. These steps are important to ensure that the data is thoroughly cleaned and ready for analysis.

AI Algorithms for Fraud Detection

One of the biggest issues that financial services companies face is the detection of credit card fraud. Machine learning techniques can help them identify fraudulent activities. This section will talk about four AI algorithms that are commonly used in this field.

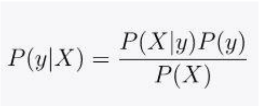
i. Random Forest: The algorithm known as Random Forest is composed of several decision trees, which are then combined to form predictions. Its formula dictates the average prediction of the various trees. The advantages of Random Forest are its ability to analyze large datasets and its accuracy. But, it may be slower to train on large datasets and prone to overfitting if there are too many trees. Eq.1 represent random forest



$$\text{Prediction} = \frac{1}{n}\sum_{i=1}^{n}\text{Prediction}_i$$

…..eq.1

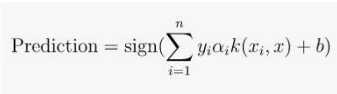
ii. Naive Bayes: The Naive Bayes algorithm can help identify fraudulent activities by estimating the likelihood of a transaction happening. It assumes that the various features of the class are independent of each another and therefore can be useful in detecting fraud. The quick and simple algorithm Naive Bayes can handle missing values and high-dimensional data. But, it may not be as effective if the data is correlated. Eq.2 represent Naive Bayes.



$$P(y|X) = \frac{P(X|y)P(y)}{P(X)}$$

…..eq.2

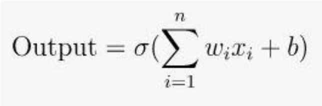
iii. SVM: The classification algorithm known as SVM finds the data in two groups according to the hyperplane. It aims to maximize the data's margin between the closest point and the hyperplane. SVM is capable of handling non-linear data, though it can be sensitive when it comes to the parameters of the kernel and its choice of function. It can also be slow to train in large datasets. Eq.3 represent SVM



$$\text{Prediction} = \text{sign}(\sum_{i=1}^{n} y_i\alpha_i k(x_i, x) + b)$$

….eq.3

iv. Neural Network: The concept of a neural network is similar to the structure of the brain. It is made up of numerous interconnected nodes, which are responsible for the output and input. A neural network can perform well in handling complex patterns, though it can be hard to train and require a lot of data. Furthermore, if the model is too complicated, it can overfit the data. Eq.4 represent Neural Network.



$$\text{Output} = \sigma(\sum_{i=1}^{n} w_i x_i + b)$$

….eq.4

The selection of an AI algorithm for fraud detection is dependent on the issue at hand. The parameters of the algorithm should be tuned to achieve the best possible result.

Results and Discussion

The table-2 and figure-1,2 shows the results of four AI systems that are used to detect credit card fraud. These include Naive Bayes, Random Forest, SVM, and Neural Network. The evaluation metrics that are used to evaluate these algorithms' performance are AUC, F1-score, precision, recall, and accuracy. The four algorithms were able to detect fraudulent transactions with high accuracy and performance. Naive Bayes was the highest AUC performer with a 97.21% score, followed by Random Forest with a 99.94% accuracy. The results of this study demonstrate the efficiency of AI-based methods in detecting credit card fraud. Although the exact requirements for this problem are not known, the findings show that the four algorithms, namely SVM, Random Forest, Naive Bayes, and Neural Network, are capable of providing high precision and accuracy in detecting such activities. Figure-3,4 represent ROC and Confusion matrix respectively.

i. Evaluation Metrices

Table 2 Result summary

Model	AUC	Accuracy	F1-Score	Precision	Recall
Naive Bayes	97.21	98.78	99.89	99.90	99.88
Neural Network	94.98	99.00	99.93	99.92	99.93
Random Forest	93.31	99.94	99.94	99.94	99.94
SVM	96.00	98.90	99.91	99.92	99.92

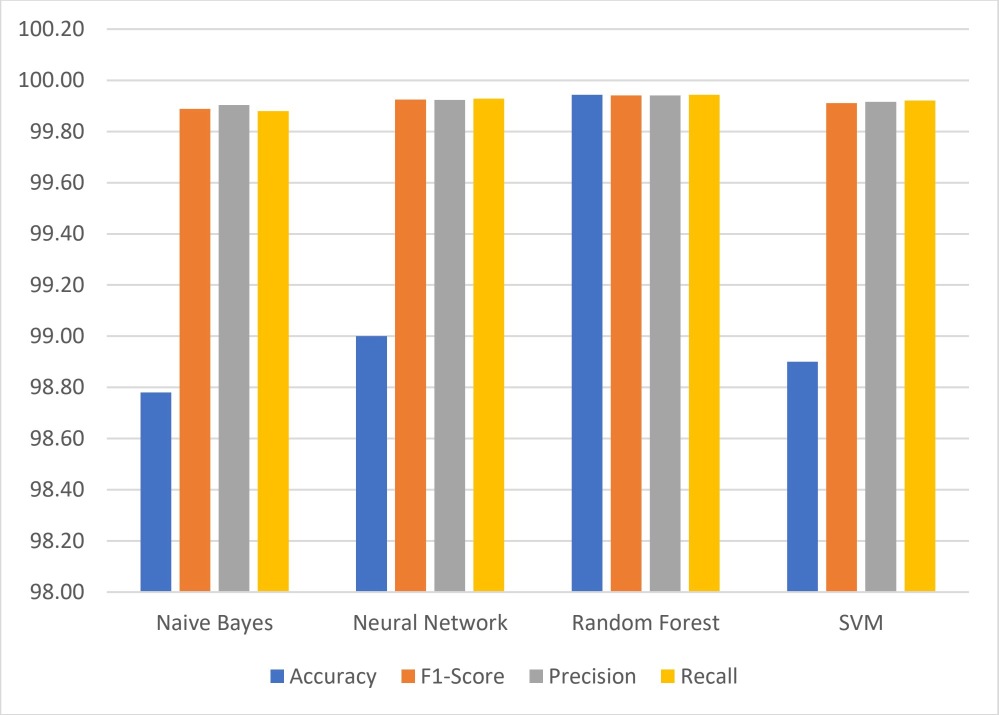


Figure 1 Various Evaluation Metrices

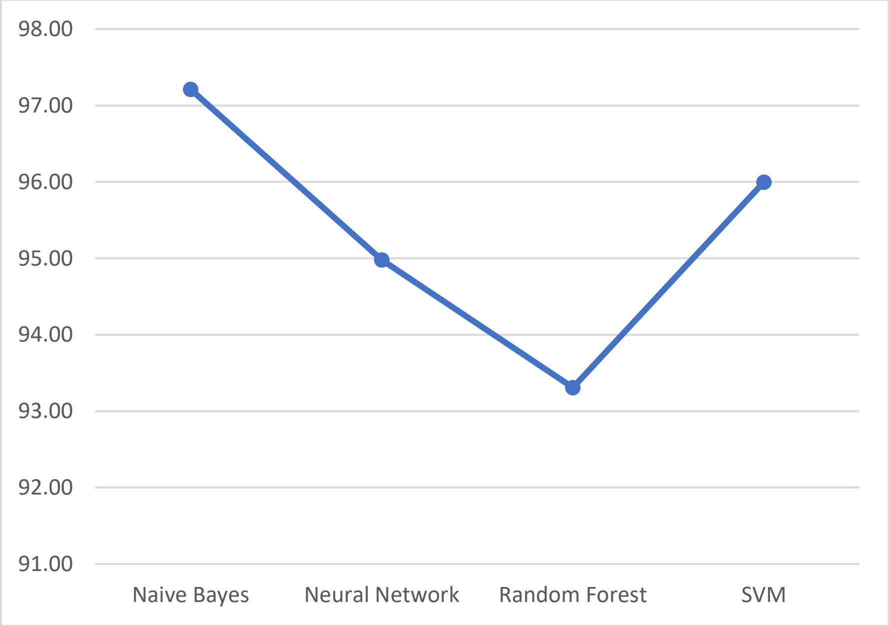


Figure 2 AUC

ii. ROC

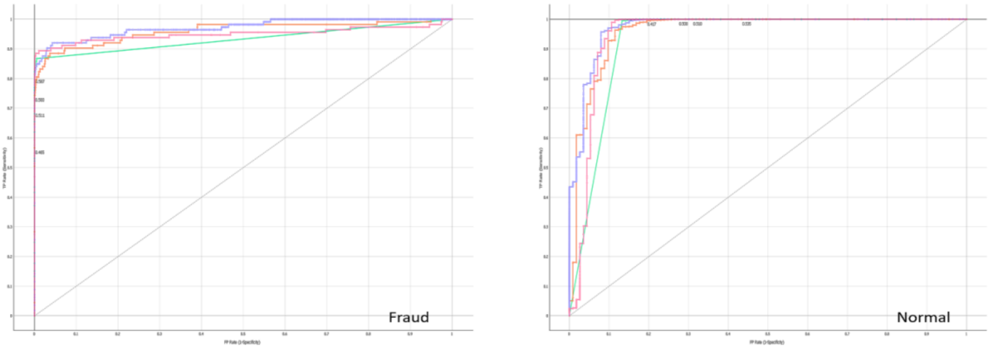
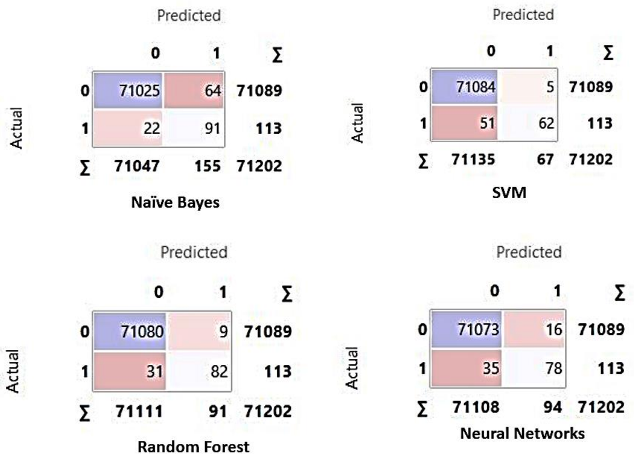


Figure 3 ROC - Fraud and Normal

iii. Confusion Matrix


Predicted
Actual

Naïve Bayes

	0	1	Σ
0	71025	64	71089
1	22	91	113
Σ	71047	155	71202

SVM

	0	1	Σ
0	71084	5	71089
1	51	62	113
Σ	71135	67	71202

Random Forest

	0	1	Σ
0	71080	9	71089
1	31	82	113
Σ	71111	91	71202

Neural Networks

	0	1	Σ
0	71073	16	71089
1	35	78	113
Σ	71108	94	71202

Figure 4 Confusion Matrix

Conclusion and future scope

The paper explores the use of AI in detecting credit card fraud. We evaluated four popular methods, namely the SVM, Random Forest, Naive Bayes, and Neural Network. We found that all of them can perform well in detecting fraudulent activities with high precision and accuracy, though each has its own limitations and advantages. Artificial intelligence (AI) can help banks and credit card companies detect and prevent fraud. It can do so by automatically learning and adapting to new patterns of activity, which helps them respond faster to potential fraud. In addition, these systems can reduce the time it takes to respond to fraud reports by up to 90%. Although our findings show that four AI algorithms can effectively detect credit card fraud, more research is needed to learn and analyze other systems that can be used in this field. Our study only looked at this aspect, which means that other areas of financial services, such as insurance and loans, could also benefit from AI. Although our study only examined one dataset, future research will examine the performance of AI systems on multiple datasets to improve the generalizability and robustness of the results. In addition, investigations on the use of algorithms that can explain their decisions will give regulators and the public more transparency regarding how these systems make decisions.

References

1. H. L. Sithic and T. Balasubramanian, “Survey of Insurance Fraud Detection Using Data Mining Techniques,” no. 3, pp. 62–65, 2013, [Online]. Available: http://arxiv.org/abs/1309.0806.
2. A. Sharma and P. Kumar Panigrahi, “A Review of Financial Accounting Fraud Detection based on Data Mining Techniques,” Int. J. Comput. Appl., vol. 39, no. 1, pp. 37–47, 2012, doi: 10.5120/4787-7016.
3. M. Kirlidog and C. Asuk, “A Fraud Detection Approach with Data Mining in Health Insurance,” Procedia - Soc. Behav. Sci., vol. 62, pp. 989–994, 2012, doi: 10.1016/j.sbspro.2012.09.168.
4. R. Rambola, P. Varshney, and P. Vishwakarma, “Data mining techniques for fraud detection in banking sector,” 2018 4th Int. Conf. Comput. Commun. Autom. ICCCA 2018, pp. 1–5, 2018, doi: 10.1109/CCAA.2018.8777535.
5. S. K. Shirgave, C. J. Awati, R. More, and S. S. Patil, “A review on credit card fraud detection using machine learning,” Int. J. Sci. Technol. Res., vol. 8, no. 10, pp. 1217–1220, 2019.
6. Y. Jain, N. Tiwari, S. Dubey, and S. Jain, “A comparative analysis of various credit card fraud detection techniques,” Int. J. Recent Technol. Eng., vol. 7, no. 5, pp. 402–407, 2019.
7. G. Kumar, S. Kumar, and A. A. Prakash, “Credit Card Fraud Detection using Machine Learning,” Int. J. Eng. Adv. Technol., vol. 10, no. 4, pp. 124–126, 2021, doi: 10.35940/ijeat.d2344.0410421.
8. A. Verma, A. Taneja, and A. Arora, “Fraud detection and frequent pattern matching in insurance claims using data mining techniques,” 2017 10th Int. Conf. Contemp. Comput. IC3 2017, vol. 2018-Janua, no. August, pp. 1–7, 2018, doi: 10.1109/IC3.2017.8284299.
9. P. Ravisankar, V. Ravi, G. Raghava Rao, and I. Bose, “Detection of financial statement fraud and feature selection using data mining techniques,” Decis. Support Syst., vol. 50, no. 2, pp. 491–500, 2011, doi: 10.1016/j.dss.2010.11.006.
10. E. W. T. Ngai, Y. Hu, Y. H. Wong, Y. Chen, and X. Sun, “The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature,” Decis. Support Syst., vol. 50, no. 3, pp. 559–569, 2011, doi: 10.1016/j.dss.2010.08.006.
11. S. Y. Huang, C. C. Lin, A. A. Chiu, and D. C. Yen, “Fraud detection using fraud triangle risk factors,” Inf. Syst. Front., vol. 19, no. 6, pp. 1343–1356, 2017, doi: 10.1007/s10796-016-9647-9.
12. S. Agrawal and J. Agrawal, “Survey on anomaly detection using data mining techniques,” Procedia Comput. Sci., vol. 60, no. 1, pp. 708–713, 2015, doi: 10.1016/j.procs.2015.08.220.
13. M. Hegazy, A. Madian, and M. Ragaie, “Enhanced Fraud Miner: Credit Card Fraud Detection using Clustering Data Mining Techniques,” Egypt. Comput. Sci. J., vol. 40, no. 03, pp. 1110–2586, 2016.

14. U. Fiore, A. De Santis, F. Perla, P. Zanetti, and F. Palmieri, “Using generative adversarial networks for improving classification effectiveness in credit card fraud detection,” Inf. Sci. (Ny)., vol. 479, pp. 448–455, 2019, doi: 10.1016/j.ins.2017.12.030.
15. S. Bagga, A. Goyal, N. Gupta, and A. Goyal, “Credit Card Fraud Detection using Pipeling and Ensemble Learning,” Procedia Comput. Sci., vol. 173, no. 2019, pp. 104–112, 2020, doi: 10.1016/j.procs.2020.06.014.
16. X. Niu, L. Wang, and X. Yang, “A Comparison Study of Credit Card Fraud Detection: Supervised versus Unsupervised,” 2019, [Online]. Available: http://arxiv.org/abs/1904.10604.
17. S. Georgieva, M. Markova, and V. Pavlov, “Using neural network for credit card fraud detection,” AIP Conf. Proc., vol. 2159, no. January, 2019, doi: 10.1063/1.5127478.
18. N. K. Trivedi, S. Simaiya, U. K. Lilhore, and S. K. Sharma, “An efficient credit card fraud detection model based on machine learning methods,” Int. J. Adv. Sci. Technol., vol. 29, no. 5, pp. 3414–3424, 2020.
19. Kaggle, “Credit Card Fraud Detection | Kaggle.” 2017, [Online]. Available: https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud.