© April 2026 | IJIRT | Volume 12 Issue 11 | ISSN: 2349-6002

A Review of Machine Learning Techniques for Credit Card Fraud Detection

Yug Saini[1], Prabhat Prajapati[2], Pratik Bhosale[3], Harsh Ramlan[4], Sumit[5]
[1,2,3,4] *Department of Computer Science & Engineering, Parul Institute of Engineering & Technology, Parul University, Vadodara, Gujarat, India.*
[5] *Assistant Professor, Department of Computer Science & Engineering, Parul Institute of Engineering & Technology, Parul University, Vadodara, Gujarat, India.*

***Abstract*—The recent surge in the adoption of digital payment systems has resulted in a substantial increase in the risk of credit card fraud faced by financial institutions. As the volume of transactions increases, the traditional rule-based monitoring systems have proven to be inadequate in identifying sophisticated patterns of fraud. Machine learning approaches have emerged as a practical solution for modelling complex transaction behaviour. The most important problem in dealing with fraud detection is the extreme imbalance in the proportion of normal to fraudulent transactions, which is less than 0.2%. As a result, the accuracy of the model is high, but the detection accuracy of fraudulent transactions is low. This paper critically examines recent studies on machine learning-based approaches for credit card fraud detection. Emphasis is given to the methods used for addressing the problems of class imbalance. Some of the widely used techniques such as Logistic Regression, Random Forest, and XGBoost have been analysed for their performance, interpretability, and computability. In addition, the performance of balancing techniques such as the Synthetic Minority Oversampling Technique (SMOTE) is also assessed based on the improvements reported in terms of Recall, F1-Score, and Precision-Recall AUC metrics for a set of benchmark problems. Apart from the performance metrics, the review also touches upon the practical aspects of the deployment of the models in a real-world scenario. The overall findings that can be obtained from the evidence indicate that the ensemble-based methods, in combination with the resampling techniques, can offer a better balance between the detection of fraud and false positives. The paper will conclude with the challenges that need to be addressed in the implementation of AI-based fraud detection methods.**

***Index Terms*—Credit Card Fraud Detection, Machine Learning, Class Imbalance, SMOTE, XGBoost, Ensemble Learning, Financial Security**

I. INTRODUCTION

The increase growth in the adoption of digital payment technologies, including online banking, e-commerce, Mobile banking systems have transformed the global ecosystem of finance [1], [3]. Which have enhanced the transaction speed, convenience, accessibility, which have also introduced the security vulnerabilities. Among these, the credit card fraud has comes out as the most prevalent as well as the costliest, causing financial losses to the banks as well as the customers [1]. The detection of the credit card fraud is highly challenging as a huge number of transactions are taking place every second. Modern fraud techniques involve coordinated attacks and manipulations that are hard to detect using normal rules. Hence, an automated system that could learn from the data and prevent fraud detection has become a necessity in financial organizations.

The Machine Learning technique has received significant attention in the field of fraud detection because it is capable of modeling complex and non-linear relationships between the data [1], [5]. Previously, traditional classifiers such as Logistic Regression were used, and now the focus is on learning-based classifiers. Despite all these advancements, there exists a limitation in most of the credit card transactions dataset, which is that the fraudulent transactions are just a fraction of the total data, i.e., less than 0.2% of all transactions [2], [9]. This imbalance in the data causes the machine model to favor the majority class, resulting in depreciatingly high accuracy but failing to effectively identify the fraud cases. Missing fraud transactions, also referred to as false negatives, resulted in human experiences of financial loss and lack of trust in financial systems. And therefore, recent studies have attempted to address the issue of learning such strategies which is fully aware of imbalances. This review paper focuses on consolidating the existing researches on machine learning based on credit card fraud detection, finding the commonly used techniques for handling class imbalance. This paper aims to provide a structured understanding of the practices, limitations, and opportunities for future improvement in fraud detection.

II. BACKGROUND AND PROBLEM DEFINITION

Credit card fraud detection is a binary classification problem where each transaction is labelled as legitimate or fraudulent [1], [3]. Though this problem appears to be straightforward but transactions data make this problem very complex to handle. The major challenge is the highly imbalance nature of credit card datasets [2], [9]. In real world scenario, legitimate transaction is vast in number as compared to fraud ones which create skewed class distribution that creates a barrier for effective learning. Machine learning models which are trained on that data often provides the outcome that are poorly sensitive towards fraudulent transaction. Another challenge which is the hidden nature of transaction data. In order to protect the user's privacy many publicly available datasets apply transformations like Principal Component Analysis (PCA) which hides the semantic meaning of the feature [10]. While it is necessary for data security, it limits interpretability and makes complications for feature engineering.

Additionally, fraud patterns are dynamic in nature; It is because attackers try to detect mechanism continuously over time and so the patterns change continuously. Models which are trained on past data requires to be upgraded otherwise there will be a performance degradation. This phenomenon is called concept drift, which further complicated the long-term deployment of fraud detection systems [6], [12]. Some careful considerations are also required for this type of models, just checking accuracy is not enough to detect fraudulent transactions as the fraud transaction can still achieve high accuracy. And to overcome this many researchers uses measurers like Recall, Precision, ROC-AUC Curve, F1 Score which give better assessment of how well fraud is detected.

III. REVIEW OF CLASS IMBALANCE HANDLING TECHNIQUES

Class imbalance is the most significant challenge in credit card fraud detection [2], [9]. In real world transaction, fraudulent transactions are available only a small fraction of total observations, often less than 0.2%. This imbalance causes machine learning classifiers for biasness their prediction towards the majority class, which results in poor fraudulent transactions. The literature widely categorizes the imbalance handling techniques into data level, algorithm level and hybrid approaches [2], [11]. Data level methods make changes in the training data distribution whereas algorithm level methods adjust learning mechanism to panelise misclassification of minority classes. Amongst these, the data level approaches are most common adopted because of the compatibility with standard classifiers.

Random under sampling reduces the number of majority class samples to balance the dataset [2]; however, it makes difficult to discard potential valuable information. Random oversampling replicates minority class samples, but it leads to overfitting [2]. To address this the technique of synthetic oversampling has been proposed. The Synthetic Minority Oversampling Technique (SMOTE) is one of the most widely used technique to handle class imbalance in fraud detection [2], [10]. Instead of making multiple duplicate minority samples it generates synthetic instance between existing minority class data points. This expands the minority class decision and improves model generalization. Multiple studies report improvement in recall When SMOTE is applied prior to model training [4], [10]. Hybrid strategies that combine oversampling and under sampling has also been discovered To balance the performance trade-off between recall and precision. Recent literature shows that appropriate sampling techniques are precondition for achieving reliable fraud detection performance.

Figure 1 illustrates a representative workflow commonly used in the literature, highlights the integration of data processing, class balancing, model training and fraud detection decision making.

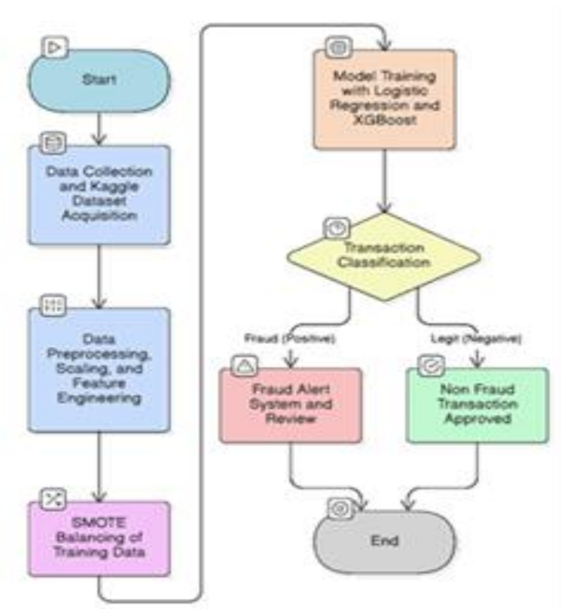

Fig. 1. General workflow of a machine learning-based credit card fraud detection system, including preprocessing, class balancing, model training, and prediction stages.

V. COMPARATIVE REVIEW OF MACHINE LEARNING MODELS

There has been a wide range of machine learning algorithm that has been applied to credit card fraud detection, with each providing something different in terms of interpretability, efficiency, and predictive performance. Studies evaluate various classifiers to understand the behaviour of highly imbalance transactional data. This section reviews the most commonly used models in the literature, with particular emphasis on their strengths and limitations in fraud detection scenarios.

4.1. Logistic Regression

Logistic Regression is frequently used algorithm as a baseline classifier in credit card fraud detection research because of its simplicity [1], [8]. This estimates the probability of transaction being fraudulent combination of inputs. It allows analysts and financial institutions to understand how the predictions are generated.

Despite these, Logistic Regression has limitations when applied to fraud detection. Its decision boundary restricts the ability to model complex, non-linear fraud patterns present in real-world transaction data. Under highly imbalanced datasets, Logistics Regression achieves overall accuracy while exhibiting poor recall for fraudulent transactions and therefore many fraudulent cases remain undetected.
Several studies report that applying class balancing techniques such as the Synthetic Minority Oversampling Technique (SMOTE) improves the ability to identify fraudulent transactions. However, even if balanced data, Logistic Regression generally underperform as compared to more advance algorithms.

4.2. Random Forest

Random Forest is an ensemble learning algorithm that combine the multiple decision trees to improve prediction [4], [6]. By aggregating the outputs of individual trees, Random Forest reduces variance and enhances robustness. This makes it well suited for modelling non-linear relationships and feature interactions that are commonly observed in transactional behaviour. This indicates that Random Forest achieves more balanced trade-off between precision and recall as compared to that of linear classifiers, particularly when trained on balanced datasets. They are more effective at identifying fraudulent transactions while maintaining reasonable false positive rates.

However, Random Forest models can be computationally intensive, especially when applied to large dataset transactions. Also, it makes the model less interpretable, which can create issues in a regulated environment.

4.3. XGBoost

Extreme Gradient Boosting (XGBoost) has emerged as one of the most effective machine learning algorithms for credit card fraud detection [4], [5], [12]. XGBoost constructs a decision tree sequence, with each new tree focuses on correcting the errors made by previous tree. Various studies show that combining XGBoost with data balancing techniques such as SMOTE which consistently produces superior results across multiple evaluation metrics [4], [9], including recall, F1-score, ROC-AUC, and Precision-Recall AUC. XGBoost is effective in capturing fraud patterns as compared to simpler models. In addition, XGBoost offers mechanisms for feature importance analysis, which improves model's efficiency. This allows practitioners to understand which transaction contribute most significantly to fraud detection decisions, supporting regulations and increase trust in the system.

Overall, this literature suggests that ensemble-based model i.e. XGBoost algorithm provides the most reliable performance for credit card fraud detection under highly imbalance.

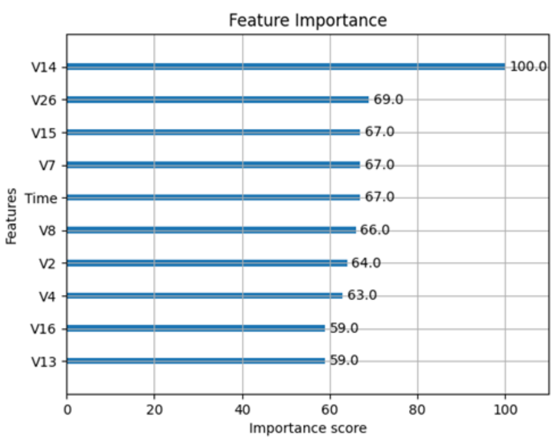

Fig. 2. Feature importance distribution generated by ensemble models, highlighting influential transaction attributes.

V. EVALUATION METRICS USED IN FRAUD DETECTION LITERATURE

As we found that credit-card transaction data is highly imbalanced in nature, due to which appropriate evaluation metrics is important to detecting fraud performances. The literature emphasizes that traditional way of finding accuracy is unreliable way to approach it, as this model achieve extreme accuracy by predicting all transactions as legitimate [1], [2]. Precision and recall are the most widely used metrics [1], [5]. Precision measures the selected part of transactions predicted as fraudulent that are actually fraud transaction, thereby reflecting all the inconvenience that are cause to consumers by the false alarms. Recall, which is also known as sensitivity, is important in fraud detection, as it measures the proportion of actual fraudulent transactions that are correctly identified. High recall is important in fraud detection, as false negatives directly correspond to undetected financial loss.

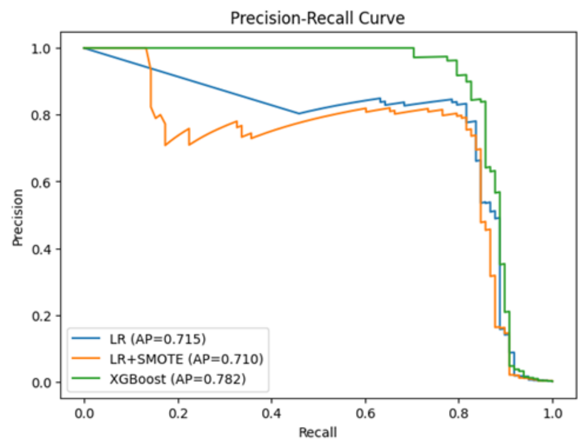

Fig. 3. Precision–Recall curve illustrating the trade-off between precision and recall for fraud detection models.

The F1-score is defined as the harmonic mean of precision and recall [5], it is frequently used to keep the balance between precision and recall. In the studies we found that models that only have recall often suffers from excessive false positives, while models focused only on precision often miss-out the fraud cases. F1 provides a single metric which effectively capture these limitations.

Threshold-independent metrics such as Receiver Operating Characteristic Area Under the Curve (ROC-AUC-curve) is widely used evaluation model for its capability across different decision threshold [1], [5].

However, in recent research we found that Precision recall (PR-AUC) provides more informative assessment under extreme class imbalance scenarios, it focuses mainly on minority class performance [9].

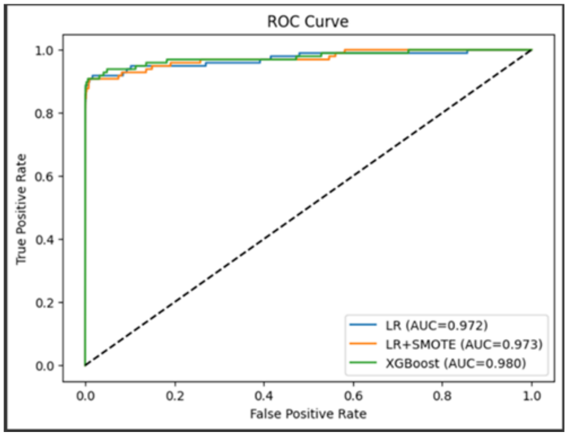

Fig. 4. ROC curve demonstrating the classification capability of ensemble-based fraud detection models under imbalanced data conditions.

Overall, this literature indicates that robust fraud detection system must be evaluated by different combination of metrices rather than single metrices. Use of multi-metric system ensure that improvement of performance and meaningful output.

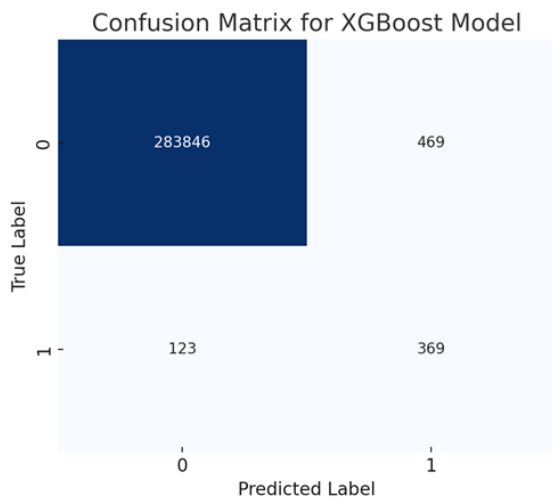

Fig. 5. Confusion matrix representation showing classification outcomes and the importance of minimizing false negatives.

VI. DISCUSSION AND PRACTICAL IMPLICATIONS

Collectively, the review of the above studies aims to show that the balance of detection ability and feasibility is a key requirement for effective credit card fraud detection. It is shown that the ensemble-based models i.e., the XGBoost algorithm, when combined with balancing techniques such as SMOTE, performs better than the traditional models in the detection of fraudulent transactions from the datasets. One of the key factors that is considered important in the detection of credit card fraud is the balance between the precision and the recall of the models. High recall is a key requirement for reducing the financial losses due to the detection of false transactions. This may result in the decline of the actual transactions, which may result in the dissatisfaction of the customers and may also result in a high workload for the investigators. It is also shown that the balance is better achieved using the boosting-based models rather than the linear models.

We also found out that interpretability is also an essential requirement for deployment. Financial institutions are run with very stringent regulations that require transparency and accountability in the decision-making system. Feature importance analysis allows the stakeholder to understand which transactions are responsible for the maximum fraud predictions, thus ensuring trust. Finally, the scalability of the system and the need for a real-time system increase the significance of model selection. It is a requirement of fraud detection systems to process a large number of transactions in the least time possible. This literature shows that the ensemble model can achieve high performance with computational feasibility.

Thus, all these findings suggest that fraud detection systems need not be designed with maximum performance, but rather with balanced performance.

VII. FUTURE RESEARCH DIRECTIONS

Although considerable progress has been achieved in the operation of machine literacy in fraud discovery, there are still numerous exploration challenges to be addressed. The operation of successional model like Longa- Short Term Memory (LSTM) model and Motor model can also be used in the discovery of fraud in deals. These models can dissect the deals in sequences rather than collectively. This can increase the rate of discovery of fraud in deals. The operation of Unsupervised literacy and Semi-supervised learning algorithms, videlicet Autoencoders and insulation timber, can also be used in the discovery of fraud in deals. The reason for this is that labelling of fraud is a time- consuming exertion. The operation of sequestration- conserving machine literacy has also entered considerable attention in recent times. The operation of Federated literacy can also be used in the discovery of fraud in deals. The operation of Federated literacy can allow numerous institutions to work together in the discovery of fraud in deals without participating sensitive client information.

Also, unborn systems can integrate resolvable artificial intelligence (XAI) ways similar as SHAP and LIME that helps in furnishing transactional position explanations. This can further enhance stoner's trust and assists in fraud judges in decision material.

VIII. CONCLUSION

In this paper, a detailed review and comparison of various approaches to detecting credit card fraud using machine learning algorithms will be provided. The review will focus on the fact that imbalanced data poses a significant problem in detecting credit card fraud, which needs to be handled properly by applying the right processing techniques. It can be inferred from the literature review that applying XGBoost along with SMOTE to handle imbalanced data will produce better results compared to other evaluation metrics like F1-score, ROC AUC, and Precision Recall AUC. The false positive rate can be handled properly. In addition to the effectiveness of the model, the literature review highlights the importance of scalability and interpretability in a practical setting, which can be inferred to be a part of the thought process and will help in gaining the trust of users. In conclusion, it can be said that a machine learning-based system for detecting credit card fraud must be technically viable and take into consideration various other factors that are practical and ethical in nature.

REFERENCES

[1] Ali, M. Qamar, and S. Zafar, "Financial fraud detection based on machine learning: A systematic literature review," Int. J. Account. Inf. Syst., vol. 44, 2022.

[2] R. B. Sulaiman, N. A. Rahman, and M. K. Hassan, "Review of machine learning approach on credit card fraud detection," Humanit. Soc. Sci. Commun., vol. 11, no. 1, 2024.

[3] L. Hernandez Aros, J. F. Delerue, and P. B. Duarte, "Financial fraud detection through the application of machine learning techniques: A literature review," Int. J. Account. Inf. Syst., vol. 40, 2021.

[4] A. R. Khalid, S. U. Rehman, and M. A. Khan, "Enhancing credit card fraud detection: An ensemble machine learning approach," IEEE Access, vol. 12, pp. 34121–34134, 2024.

[5] F. K. Alarfaj, A. Almutairi, and S. Alzahrani, "Credit card fraud detection using state-of-the-art machine learning and deep learning algorithms," IEEE Trans. Comput. Social Syst., vol. 9, no. 4, pp. 1156–1166, 2022.

[6] S. K. Hashemi, M. R. Keyvanpour, and F. Moradi, "Fraud detection in banking data by machine learning techniques," in Proc. IEEE Int. Conf. Big Data, 2022, pp. 1874–1881.

[7] U. Detthamrong, T. Srisuruk, and P. Sornlertlamvanich, "Enhancing fraud detection using advanced machine learning techniques," Int. J. Econ. Financial Issues, vol. 14, no. 2, pp. 112–121, 2024.

[8] R. Achary and C. J. Shelke, "Fraud detection in banking transactions using machine learning," in Proc. IEEE Int. Conf. Intell. Syst., 2023, pp. 98–104.

[9] N. S. Alfaiz and S. M. Fat, "Enhanced credit card fraud detection model using machine learning," Electronics, vol. 11, no. 19, pp. 1–18, 2022.

[10] E. Ileberi, Y. Sun, and Z. Wang, "A machine learning-based credit card fraud detection using the GA algorithm for feature selection," J. Big Data, vol. 9, no. 1, pp. 1–24, 2022.

[11] P. K. Kamuangu, "A review on financial fraud detection using artificial intelligence and machine learning," J. Econ. Financial Account. Stud., vol. 6, no. 1, pp. 45–59, 2024.

[12] X. Feng and S. K. Kim, "Novel machine learning-based credit card fraud detection systems," Appl. Sci., vol. 14, no. 3, pp. 1–17, 2024.

[13] D. Mienye and Y. Sun, "A machine learning method with hybrid feature selection for

improved credit card fraud detection," Electronics, vol. 12, no. 5, pp. 1–20, 2023.

[14] E. F. Malik, A. A. Bakar, and Z. Yusoff, "Credit card fraud detection using a new hybrid machine learning architecture," Sensors, vol. 22, no. 14, pp. 1–19, 2022.

[15] S. Khan, A. Rehman, and M. I. Khalil, "Developing a credit card fraud detection model using machine learning approaches," Int. J. Adv. Comput. Sci. Appl., vol. 13, no. 6, pp. 215–223, 2022.

[16] S. Nuthalapati, "Comparative analysis of ensemble learning methods for credit card fraud detection," Int. J. Data Sci. Anal., vol. 15, no. 2, pp. 89–101, 2023.

[17] S. Faraji, "Performance evaluation of KNN and ensemble models for financial fraud detection," J. Inf. Secur., vol. 13, no. 4, pp. 233–245, 2022.

[18] M. Islam, R. Hossain, and A. Rahman, "Rule-based and machine learning approaches for real-time fraud detection," Expert Syst. Appl., vol. 213, 2023.

[19] A. Bello, K. Adeyemi, and T. Ojo, "Big data analytics for real-time financial fraud detection," Future Gener. Comput. Syst., vol. 140, pp. 45–56, 2023.