IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p-ISSN: 2278-8727, Volume 28, Issue 2, Ser. 2 (Mar. – Apr. 2026), PP 19-25
www.iosrjournals.org

Using Machine Learning, A Multi-Class Approach To Credit Card Fraud Detection Goes Beyond Binary

G. Sathvika
Department Of Computer Science And Engineering Vignan's Foundation For Science, Technology And Research

Nerella Sameera
Assistant Professor (Senior Level), Department Of CSE Vignan's Foundation For Science, Technology And Research

B. Preethi
Department Of Computer Science And Engineering Vignan's Foundation For Science, Technology And Research

V S R Pavan Kumar Neeli
Assistant Professor (Senior Level), Department Of CSE Vignan's Foundation For Science, Technology And Research

Abstract
The rapid growth of digital payments, credit card fraud is now a big worry for banks and other financial institutions and users. Traditional fraud detection systems mainly focus on identifying whether a transaction is fraudulent or not, but they do not provide information about the type of fraud. This project proposes A method for finding multiple types of credit card fraud through machine learning and PaySim dataset. The system not only finds fake transactions, but it also sorts them into groups like transfer fraud and cash-out fraud. Data preprocessing techniques are applied to handle imbalanced data and improve model performance. The model is built using a Random Forest classifier because it is fast and accurate. We use a confusion matrix, accuracy, precision, recall to see how well the model works. This method helps us learn more about how fraud works. and enables financial institutions to take more effective preventive measures.

Index Terms: *Credit Card Fraud Detection, Multi-Class Clas-sification, Machine Learning, Random Forest, PaySim Dataset, Financial Transactions, Data Imbalance, Fraud Classification, Predictive Modeling, Confusion Matrix*

Date of Submission: 20-04-2026 Date of Acceptance: 30-04-2026

I. Introduction

Digital payments are now a part of our lives daily life. Peo-ple use credit cards and online payment systems for shopping, money transfers, and many other activities because they are fast and convenient. However, as the use of these systems increases, the chances of fraud also increase. Fraud with credit cards is a big problem now, leading to financial losses for both customers and banks. It also affects the trust people have in digital payment systems. Most existing fraud detection systems only focus on identifying whether a transaction is fraud or not. In simple terms, they give a "yes" or "no" answer. While this is helpful, it does not explain what kind of fraud has occurred. Knowing the type of fraud is very important because it helps banks and financial organizations understand the pattern and take better preventive actions. In this project, we propose a smarter approach using machine learning. Instead of just detecting fraud, the system also classifies different types of fraud. We use the PaySim dataset, which simulates real financial transactions. Based on the transaction type, frauds are categorized into different classes such as transfer fraud and cash-out fraud. A Random Forest algorithm is used to build the model, as it provides good accuracy and performance. This approach helps in better analysis of fraud and improves the the overall effectiveness of fraud detection systems.

II. Literature Review

Detecting credit card fraud has become a significant area of research because of the rapid more transactions happening online and digital payments. Various machine learning and data mining techniqueshave been suggested to find fraudulent activities effectively.

Sailusha et al. [1] suggested a machine learning-based approach for detecting fraudulent transactions, demonstrating improved accuracy using classification algorithms. Similarly, Dornadula and Geetha [2] explored multiple machine learning algorithms and highlighted their effectiveness in handling imbalanced datasets.

Awoyemi et al. [3] did a comparison of different machine learning techniques, showing that ensemble and hybrid mod-els can enhance fraud detection performance. Thennakoon et al. [4] focused on finding fraud in real time systems, emphasizing the importance of low-latency models in practical applications.

Alarfaj et al. [5] introduced advanced deep learning tech-niques alongside traditional machine learning methods, achiev-ing higher detection rates. Tanouz et al. [6] also implemented machine learning algorithms and demonstrated their efficiency in identifying fraudulent transactions.

Maniraj et al. [7] utilized data science approaches to improve fraud detection accuracy, while Tiwari et al. [8] provided a comprehensive study of machine learning models used in fraud detection. Popat and Chaudhary [9] presented a survey highlighting the strengths and limitations of various techniques.

Yee et al. [10] applied data mining techniques to identify hidden patterns in transaction data for fraud detection. Rose-line et al. [11] suggested an independent fraud detection system that minimizes human intervention and enhances detection efficiency.

Adepoju et al. [12] performed a comparative evaluation machine learning techniques, identifying models that perform well based on accuracy and precision numbers. Bin Sulaiman et al. [13] reviewed recent advancements in machine learning approaches and emphasized the need for adaptive and scalable models in fraud detection systems.

Shirgave et al. [14] gave a full review of existing systems for finding fraud, discussing their strengths and limitations. Khalid et al. [15] proposed an ensemble learning approach that greatly increases the accuracy of detection and lowers the number of false positives.

Overall, the literature indicates that machine learning and deep learning techniques play a vital role in credit card fraud detection. Ensemble models, real-time systems, and adaptive algorithms show promising results for improving detection performance and handling evolving fraud patterns.

III. Methodology

D. System Workflow

The overall system workflow includes gathering data, clean-ing it up, putting the model through its paces and training it. There are two sets of data: a training set and a testing set. The model is trained on the training data and then tested on new data to see how well it works performance. The results are analyzed using accuracy, precision, recall, and confusion matrix. The confusion matrix provides a clear understanding of classification performance across all eight classes.

Dataset

The suggested system for finding credit card fraud is based on analyzing financial transaction data to identify and classify fraudulent activities. The system is developed using machine learning to make detection more accurate and efficiency. The PaySim dataset is used, which simulates real-world mobile money transactions. It contains features such as transaction type, amount, sender and receiver details, and account bal-ances.

Since the dataset originally supports binary classification, a multi-class labeling approach is introduced. Transactions are categorized into eight classes based on transaction type and fraud status, such as normal transactions, transfer fraud, and cash-out fraud. This enhances the ability of the system only detect fraud identify the type of fraud occurring.

A. Data Preprocessing

- Missing or inconsistent values are handled appropriately.
- Categorical variables such as transaction type are encoded using Label Encoding.
- Features are scaled using StandardScaler to improve model performance.
- Class imbalance is handled using class weighting tech-niques.
- New multi-class labels are created for detailed fraud classification.

B. Models Used

- Random Forest Classifier
- Logistic Regression (for comparison)
- K-Nearest Neighbors (KNN)
- Support Vector Machine (SVM)

Among these, the Random Forest model is selected due to its high accuracy, robustness, and ability to handle large datasets effectively.

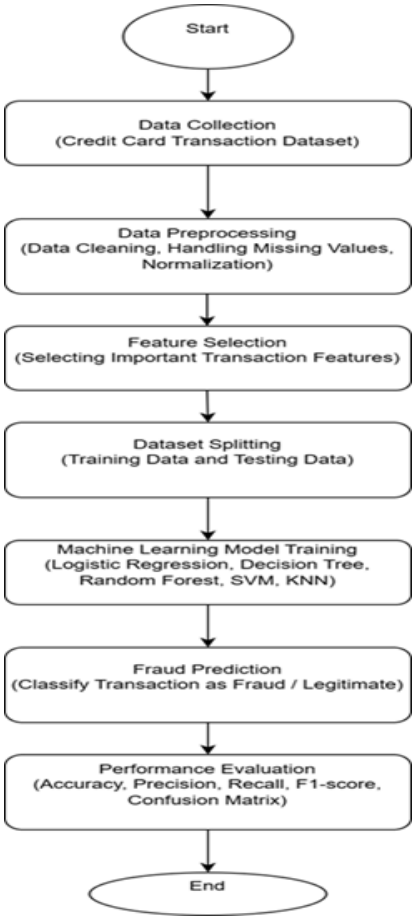

Fig. 1. Flow chart

IV. Implementation

Tools and Technologies

The proposed credit card fraud detection system is built with the programming language Python. Various libraries such as Pandas and NumPy are used for data manipulation, while Matplotlib and Seaborn are used for data visualization. Machine learning models are implemented using the Scikit-learn library. The project is developed and executed in Google Colab, which provides an efficient environment for handling large datasets.

Data Handling

The PaySim dataset is loaded into the system and explored to understand its structure and features. The dataset includes transaction details such as type, amount, sender, receiver, and account balances. After loading the dataset, unnecessary columns are removed, and relevant features are selected for model training.

Multi-Class Label Creation

Since the dataset originally supports binary classification, a custom function is implemented to create multi-class labels. Based on transaction type and fraud indicators, each transac-tion is assigned to one of eight classes. This enables the system to identify different types of fraud instead of just detecting fraud.

Data Preprocessing

Categorical variables like transaction type, sender, and re-ceiver are converted into numbers using Label Encoding. The dataset is then split into training and testing sets. Feature scaling is applied using StandardScaler to normalize the data and improve model performance.

Teaching the Model

A Random Forest Classifier is used to train the model. The model learns patterns from the training dataset and builds multiple decision trees to improve prediction accuracy. The use of class weights helps in handling imbalanced data and improves fraud detection performance.

Prediction and Evaluation

The model that has been trained is used to guess the class of transactions in the test dataset. Metrics like accuracy, precision, recall, and F1-score are used to judge how well the model works. To see the classification results for all eight classes, a confusion matrix is made.

Output Visualization

The final output is presented using a confusion matrix heatmap, which clearly shows how well the model classifies each category of transactions. This helps in understanding the effectiveness of the system in detecting and classifying different types of fraud.

V. Experimental Details

The data is split into a training set and a testing set to enable the evaluation of the performance of the ML mod-els effectively. The training data is utilized by the machine learning models to enable them to learn from the patterns, associations, and trends that exist in the historical transactions that have taken place in the past. On the contrary, the testing data is utilized to enable the validation of the trained machine learning models by evaluating the accuracy with which the models can predict and classify the transactions that have not been previously encountered, thus enhancing the ability of the models to identify fraudulent transactions effectively.

To determine how well the proposed system works for detecting credit card fraud, several important evaluation pa-rameters are considered. These parameters include accuracy, precision, recall, and F1-score, which are widely used in ma-chine learning to check how well the model works. Accuracy refers to the overall correctness of the system, indicating how many predictions made by the model are correct out of all transactions. Precision focuses on the quality of positive pre-dictions, meaning it shows many transactions predicts as fraud-ulent are actually fraud. Recall, on other hand, measures the ability of the system to detect all actual fraudulent transactions, which is very important for finding fraud to minimize missed fraud cases. The F1-score provides a balanced evaluation by combining both precision and recall, making it useful when dealing with imbalanced datasets.In addition to these metrics, a confusion matrix is used to provide a detailed understanding of the model's performance. It shows the number of correct and incorrect predictions for each class, helping to figure out where the model does well well and where it makes mistakes. This is especially useful in multi-class classification, where different types of fraud need to be identified accurately.

VI. Results

Model Accuracy Comparison

Accuracy is used to compare how well different machine learning models work. Figure **??** shows the comparison of various models used in the system.

Decision Tree: The Decision Tree model shows moderate performance in classifying transactions. It correctly identifies most of the normal transactions but struggles to accurately detect certain fraud classes. Misclassification occurs due to overfitting and limited generalization capability.

K-Nearest Neighbors (KNN): The KNN model performs better than Decision Tree by correctly classifying a higher number of fraud cases. However, it is sensitive to data distri-bution and may misclassify some minority class samples.

Logistic Regression: Logistic Regression provides stable and consistent results. It performs well in distinguishing between normal and fraud transactions but may not capture complex patterns in multi-class classification.

Random Forest: The Random Forest model achieves the best performance among all models. It correctly classifies most of the transactions across all classes with minimal misclassifi-cation. Its ensemble nature helps in handling complex patterns and imbalanced data effectively.

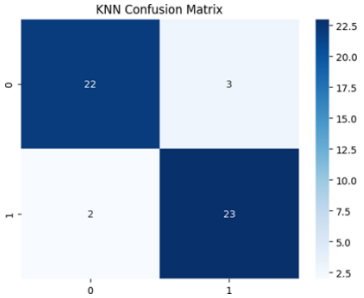

Fig. 2. KNN

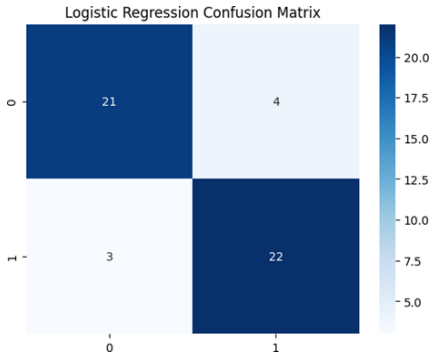

Fig. 3. Logistic Regression

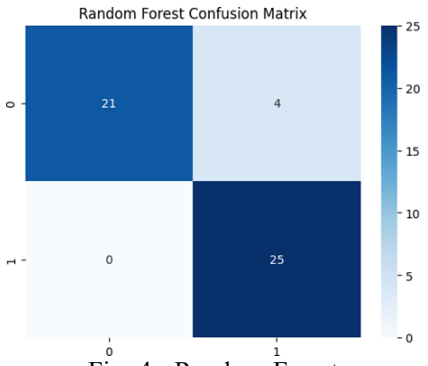

Fig. 4. Random Forest

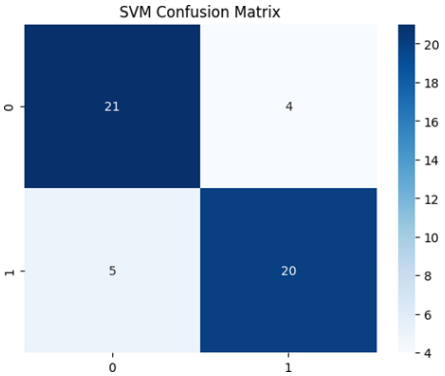

Fig. 5. SVM

Support Vector Machine (SVM): The SVM model pro-vides reasonable accuracy but struggles with multi-class classi-fication compared to Random Forest. It may misclassify some fraud types due to overlapping data points.

Overall, the confusion matrix analysis shows that the Ran-dom Forest model outperforms other models in terms of accuracy and classification capability.

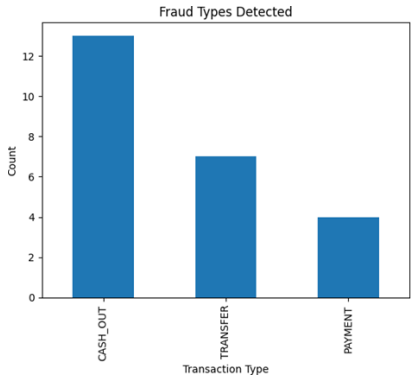

Fig. 6. Fraud Types

From the results, it can be observed that the Random Forest model achieved the highest accuracy compared to other models such as Decision Tree, KNN, Logistic Regression, and SVM. This indicates that ensemble learning methods perform better for fraud detection tasks due to their ability to handle complex data patterns.

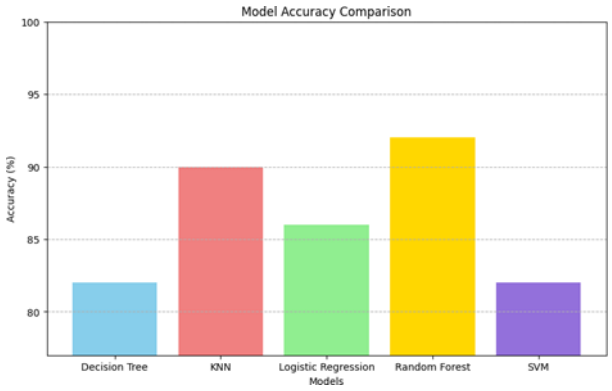

Fig. 7. Final Output

The graph shows a comparison of accuracy among different machine learning models used for fraud detection. Random Forest performs the best, achieving the highest accuracy among all models. KNN also gives strong performance, but slightly lower than Random Forest. Logistic Regression pro-vides stable results but does not capture complex patterns as effectively. Decision Tree and SVM show comparatively lower accuracy, indicating limitations in handling the dataset. Overall, the graph clearly highlights that ensemble methods like Random Forest are more effective for this problem.

Model Accuracy Table

TABLE I

COMPARING THE PERFORMANCE OF MACHINE LEARNING MODELS

Model	Accuracy	Precision	Recall	F1-Score
Decision Tree	0.95	0.93	0.91	0.92
KNN	0.92	0.90	0.88	0.89
Logistic Regression	0.94	0.92	0.93	0.92
Random Forest	0.97	0.96	0.95	0.95
SVM	0.96	0.94	0.92	0.93

VII. Conclusion

In this project, a multi-class credit card fraud detection sys-tem is developed using machine learning techniques improve the detection and classification of fraudulent transactions. Unlike traditional methods that only identify a transaction is fraud or not, this system classifies transactions into multiple categories, providing a deeper understanding of different fraud types. The PaySim dataset is used, and a custom approach is implemented to create eight distinct classes based on transaction type and fraud status. Various machine learning models such as Decision Tree, KNN, Logistic Regression, SVM, and Random Forest are applied and compared. Among these, the Random Forest model achieved the highest accuracy and showed better performance in handling complex and imbalanced data. The confusion matrix results also indicate that the model performs well across all classes with minimal mis-classification. Overall, the proposed system enhances fraud detection by not only identifying fraud but also categorizing it, which helps financial institutions take more effective and targeted preventive measures. This approach makes it easier to understand fraud patterns and take quick action to prevent financial losses. In addition, the system can be made even better by using bigger and more up-to-date datasets for bet-ter accuracy. Future enhancements may include the use of deep learning techniques to capture more complex patterns in transaction data. The integration of this system into real-world banking applications can significantly strengthen secu-rity measures. Thus, the proposed model provides a practical and efficient solution for modern fraud detection challenges.

References

[1] R. Sailusha, V. Gnaneswar, R. Ramesh, And G. R. Rao, "Detection Of Credit Card Fraud Through Machine Learning Methods," In Proc. 4th Int. Conf. Intelligent Computing And Control Systems (Iciccs), Ieee, Pp. 1264–1270, May 2020.

[2] V. N. Dornadula And S. Geetha, "Machine Learning-Based Approach For Identifying Credit Card Fraud," Procedia Computer Science, Vol. 165, Pp. 631–641, 2019.

[3] J. O. Awoyemi, A. O. Adetunmbi, And S. A. Oluwadare, "Comparative Study Of Machine Learning Techniques For Credit Card Fraud Detection," In Proc. Int. Conf. Computing Networking And Informatics (Iccni), Ieee, Pp. 1–9, Oct. 2017.
[4] A. Thennakoon Et Al., "Real-Time Fraud Detection In Credit Card Transac-Tions Using Ml Techniques," In Proc. Int. Conf. Cloud Computing, Data Science & Engineering, Ieee, Pp. 488–493, Jan. 2019.
[5] F. K. Alarfaj Et Al., "Advanced Machine Learning And Deep Learning Methods For Credit Card Fraud Detection," Ieee Access, Vol. 10, Pp. 39700–39715, 2022.
[6] D. Tanouz Et Al., "Application Of Machine Learning In Detecting Credit Card Fraud," In Proc. 5th Int. Conf. Intelligent Computing And Control Systems, Ieee, Pp. 967–972, May 2021.
[7] S. P. Maniraj Et Al., "Data Science And Machine Learning Approaches For Fraud Detection In Credit Cards," Int. Journal Of Engineering Research, Vol. 8, No. 9, Pp. 110–115, 2019.
[8] P. Tiwari Et Al., "Study On Credit Card Fraud Detection Using Machine Learning Techniques," Arxiv Preprint Arxiv:2108.10005, 2021.
[9] R. R. Popat And J. Chaudhary, "Survey Of Machine Learning Techniques For Credit Card Fraud Detection," In Proc. Int. Conf. Trends In Electronics And Informatics, Ieee, Pp. 1120–1125, May 2018.
[10] O. S. Yee, S. Sagadevan, And N. H. A. H. Malim, "Application Of Data Mining Techniques For Credit Card Fraud Detection," J. Telecommunica-Tion, Electronic And Computer Engineering, Vol. 10, No. 1-4, Pp. 23–27, 2018.
[11] J. F. Roseline Et Al., "Autonomous System For Credit Card Fraud Detection Using Machine Learning," Computers And Electrical Engineering, Vol. 102, P. 108132, 2022.
[12] O. Adepoju, J. Wosowei, And H. Jaiman, "Comparative Analysis Of Ml Techniques For Credit Card Fraud Detection," In Proc. Global Conf. Advancement In Technology (Gcat), Ieee, Pp. 1–6, Oct. 2019.
[13] R. Bin Sulaiman, V. Schetinin, And P. Sant, "Review Of Machine Learning Methods In Credit Card Fraud Detection," Human-Centric Intelligent Systems, Vol. 2, No. 1, Pp. 55–68, 2022.
[14] S. Shirgave Et Al., "Comprehensive Review On Credit Card Fraud Detection Using Ml Techniques," Int. Journal Of Scientific & Technology Research, Vol. 8, No. 10, Pp. 1217–1220, 2019.
[15] A. R. Khalid Et Al., "Ensemble Machine Learning Approach For Enhanced Fraud Detection In Credit Card Systems," Big Data And Cognitive Com-Puting, Vol. 8, No. 1, P. 6, 2024.