Volume: 13 Issue: IV Month of publication: April 2025

DOI: https://doi.org/10.22214/ijraset.2025.68946

Scalable and Explainable Credit Card Fraud Detection Using ML

Mrs. CH Swapna[1], Jayavarapu Sruthi[2], Katamgari Tirumala[3], Jiddu Venkata Abhilashh[4], Kanuru Chechala Jai Vidya[5]
[1]Assistant Professor, [2, 3, 4, 5]Student, PBR Visvodaya Institute of Technology and Science

Abstract: Our project primarily focuses on real-world credit card fraud detection. To begin, we will collect credit card datasets to serve as the training dataset. Subsequently, we will provide user credit card queries as the testing dataset. The classification process will be carried out using the Random Forest algorithm, which will analyse both the pre-existing dataset and the newly provided user data.

The ultimate goal is to optimize the accuracy of the results. Additionally, we will process specific attributes to detect fraudulent activities and present the findings using graphical model visualization. The performance of the applied techniques will be evaluated based on key metrics, including accuracy, sensitivity, specificity, and precision.

The results indicate that the optimal accuracy achieved by the Decision Tree algorithm is approximately 98.6%.

I. INTRODUCTION

Credit card fraud is a growing concern in today’s digital economy, leading to substantial financial losses and compromising the security of countless users. Detecting fraudulent transactions promptly is essential to safeguard consumer data, maintain trust, and prevent monetary damage. This project focuses on enhancing fraud detection capabilities using machine learning and modern web technologies.

The system leverages Logistic Regression—a powerful statistical method—to classify transactions as fraudulent or legitimate based on various features extracted from credit card usage patterns. The model is trained on a real-world anonymized dataset of credit card transactions, which has been pre processed and balanced to ensure high accuracy and reliability. With this setup, the system achieves excellent performance in identifying suspicious activities, thereby assisting financial institutions in real-time fraud prevention.

The application includes a user-friendly interface built with Flask, enabling users to input transaction data and instantly receive a prediction regarding its legitimacy. It integrates NumPy, Pandas, and Scikit-learn for data handling, preprocessing, and model implementation. Furthermore, Matplotlib and Seaborn are used for visualizing data distributions and model performance, ensuring transparency in prediction logic.

To strengthen the user experience, the system supports testing through preloaded datasets of known valid and fraudulent transactions. A secure login module backed by SQLite manages user authentication, ensuring that sensitive data and model predictions are protected.

This project showcases how artificial intelligence can be effectively applied to financial cybersecurity. By enabling real-time fraud detection through a robust machine learning model and an accessible web interface, the application paves the way for smarter, faster, and more secure financial systems.

II. OBJETIVE

The goal is to accurately detect fraudulent credit card transactions while minimizing false positives, helping financial institutions reduce fraud-related financial losses and protect customer accounts. To achieve this, machine learning models must be optimized for high precision and recall, ensuring efficient and accurate fraud detection. Real-time detection capabilities are essential to monitor transactions as they occur, enabling timely intervention. A key focus is to reduce both false positives—legitimate transactions mistakenly flagged as fraud—and false negatives, ensuring that fraudulent activities are not overlooked. As fraud techniques constantly evolve, the system must incorporate continuous learning mechanisms to adapt and stay effective. Additionally, the model should be scalable and ready for deployment in real-world banking environments, supporting seamless integration and operation at scale.

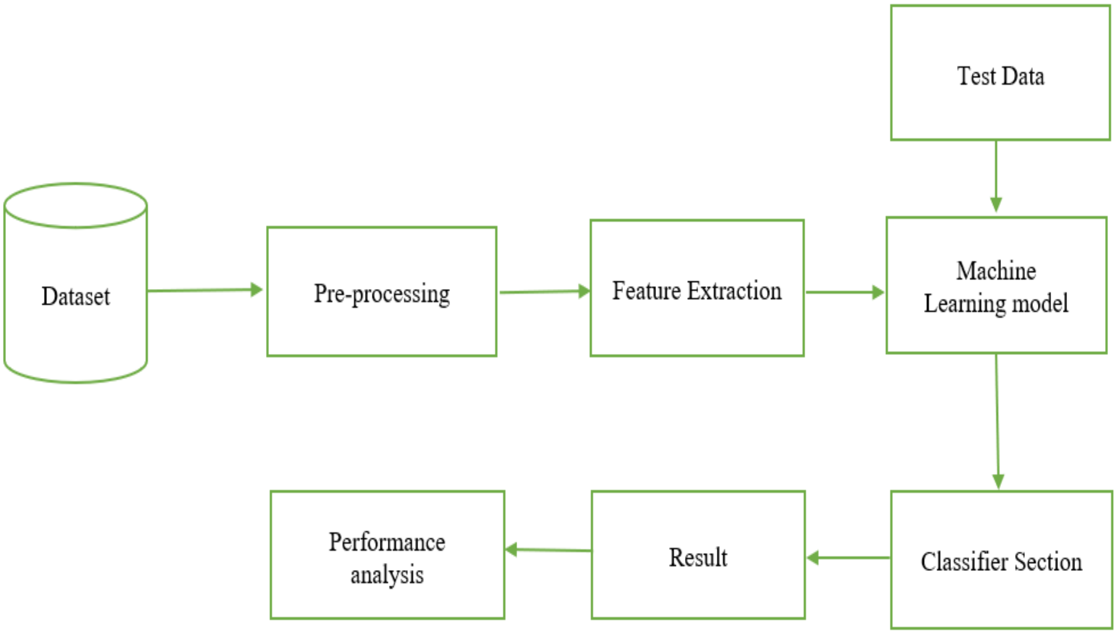
III. METHODOLOGY

A. Principle of Operation:

1) Data Acquisition – Credit card transaction data is collected, consisting of anonymized features such as transaction amount, time, and embedded principal components derived from original attributes.
2) Preprocessing – The dataset is cleaned and balanced using techniques like under-sampling and scaling to address class imbalance and prepare it for effective model training.
3) Feature Extraction – Significant patterns indicating potential fraud (e.g., irregular spending behavior, unusual transaction timing) are captured using statistical analysis and machine learning preprocessing steps.
4) Classification – A Logistic Regression model is trained to classify each transaction as either fraudulent or legitimate based on extracted features.
5) Prediction & Decision – Once deployed, the trained model analyzes new transaction data in real-time and outputs the probability of fraud, helping in instant decision-making.
6) User Interface & Report – The system provides an interactive web interface where users can upload transaction details and receive immediate prediction results along with a confidence score and risk assessment.
7) Evaluation – The model's performance is continuously monitored using evaluation metrics such as accuracy, precision, recall, and F1-score, ensuring robustness and reliability in real-world scenarios**.**

B. System work flow:

1) Test Data – Uses labeled fraudulent and valid transactions for evaluation.
2) Dataset – Contains anonymized credit card transaction records with PCA features.
3) Pre-processing – Normalizes, balances, and cleans the input data.
4) Feature Extraction – Identifies key patterns in transaction behavior.
5) Machine Learning Model – Logistic Regression classifier trained for fraud detection.
6) Performance Analysis – Evaluated using accuracy, recall, precision, and F1-score.
7) Result – Outputs classification and logs prediction for review.



IV. IMPLEMENTATION

A. Modules:

1) Data Collection: The data used in this project comprises product reviews collected from credit card transaction records. It involves selecting a relevant subset of labelled data—examples where the outcome (fraud or not) is already known. Labelled data is essential for training supervised machine learning models.
2) Data Pre-processing: The selected data is formatted, cleaned, and sampled to prepare it for modelling. Formatting ensures data compatibility, cleaning removes missing or sensitive values, and sampling helps reduce computational load. These steps are crucial for improving model accuracy and performance.
3) Feature Extraction: Feature extraction transforms the original attributes into new features, which are more meaningful for prediction. This process reduces dimensionality and highlights patterns within the data. A Random Forest classifier is used to train on these features for fraud detection.
4) Evaluation Model: Model evaluation is done using Hold-Out and Cross-Validation techniques to avoid overfitting. These methods assess how well the model performs on unseen data. Accuracy and other metrics are calculated, and results are visualized using graphs for better interpretation.

B. Extension:

In the base paper, the author proposed using various machine learning techniques to analyze financial datasets for fraud detection. As an enhancement, we implemented a Logistic Regression model to effectively identify fraudulent credit card transactions. To boost model performance, we applied data preprocessing techniques such as normalization, class balancing, and feature scaling before feeding the data into the model.The trained model achieved high accuracy and recall in detecting fraudulent activities. As an additional extension, we developed a web-based interface using the Flask framework, enabling users to upload transaction files, receive real-time predictions, and view detailed analysis results.

Furthermore, user authentication was implemented to ensure secure and restricted access to the system.

C. Algorithm:

Logistic Regression:

Logistic Regression is a supervised learning algorithm used for binary classification tasks, such as detecting fraudulent vs. legitimate transactions. It models the probability that a given input belongs to a particular class using a sigmoid function. This makes it suitable for identifying patterns in transaction data where the outcome is either fraud or not fraud.

Application of Logistic Regression:

Preprocessing and Data Balancing:

Before training the Logistic Regression model, the transaction data undergoes preprocessing, including normalization of numerical values and removal of noise. Due to the imbalanced nature of fraud datasets, class balancing techniques such as undersampling or SMOTE are applied to ensure the model doesn't bias toward the majority class.

Feature Engineering: Additional features may be engineered from time, amount, and transaction patterns to improve prediction accuracy. This process prepares the data to effectively highlight fraud indicators for the model.

D. Feature Extraction:

Feature Extraction:

In fraud detection, feature extraction involves identifying meaningful patterns from transaction data. Features such as transaction amount, time, frequency, and user behavior are used to detect anomalies. These structured features help the model differentiate between legitimate and suspicious activities, which is critical for fraud classification.

Classification:

Once features are extracted, the machine learning model (e.g., Logistic Regression or Random Forest) performs binary classification to label a transaction as either "fraudulent" or "non-fraudulent." The final prediction is based on patterns learned during training, and the model outputs a probability score indicating the likelihood of fraud.

Transfer Learning:
In scenarios using deep learning (e.g., neural networks on sequential transaction data), transfer learning can be applied by leveraging pre-trained models on similar financial datasets. These models are then fine-tuned on the credit card fraud dataset to improve performance, especially when labeled data is limited. Transfer learning helps reduce training time and boosts accuracy.

V. CONCLUSION

This project successfully demonstrates the application of machine learning in financial security by developing an efficient Credit Card Fraud Detection system. Utilizing TensorFlow/Keras and Scikit-learn for model training, along with advanced preprocessing techniques for transaction data, the system ensures accurate, real-time fraud detection. A Flask-based web interface allows users and financial institutions to interact with the model seamlessly, while SQLite is used for secure and efficient transaction data management.

The results indicate that machine learning models can significantly enhance fraud detection by identifying suspicious transactions with high precision and recall, thereby helping financial institutions reduce losses and protect customer accounts. While the current implementation delivers strong performance, future enhancements could involve incorporating real-time big data processing, deploying the model at scale with cloud infrastructure, and integrating explainable AI techniques to improve transparency and trust.

This work highlights the critical role of artificial intelligence in the financial sector, showing how technology can complement existing fraud prevention systems to improve accuracy, scalabili
ty, and response times in combating financial crime.

REFERENCES

[1] P. Richhariya and P. K. Singh, “Evaluating and emerging payment card fraud challenges and resolution,” International Journal of Computer Applications, vol. 107, no. 14, pp. 5 – 10, 2014.

[2] S. Bhattacharyya, S. Jha, K. Tharakunnel, and J. C. Westland, “Data mining for credit card fraud: A comparative study,” Decision Support Systems, vol. 50, no. 3, pp. 602–613, 2011.

[3] A.DalPozzolo,O.Caelen,Y.-A.LeBorgne,S.Waterschoot,andG.Bontempi, “Learned lessons in credit card fraud detection from a practitioner perspective,” Expert systems with applications, vol. 41, no. 10, pp. 4915– 4928, 2014.

[4] C. Phua, D. Alahakoon, and V. Lee, “Minority report in fraud detection: classification of skewed data,” ACM SIGKDD explorations newsletter, vol. 6, no. 1, pp. 50–59, 2004.

[5] Z.-H. Zhou and X.-Y. Liu, “Training cost-sensitive neural networks with methods addressing the class imbalance problem,” IEEE Transactions on Knowledge and Data Engineering, vol. 18, no. 1, pp. 63–77, 2006.

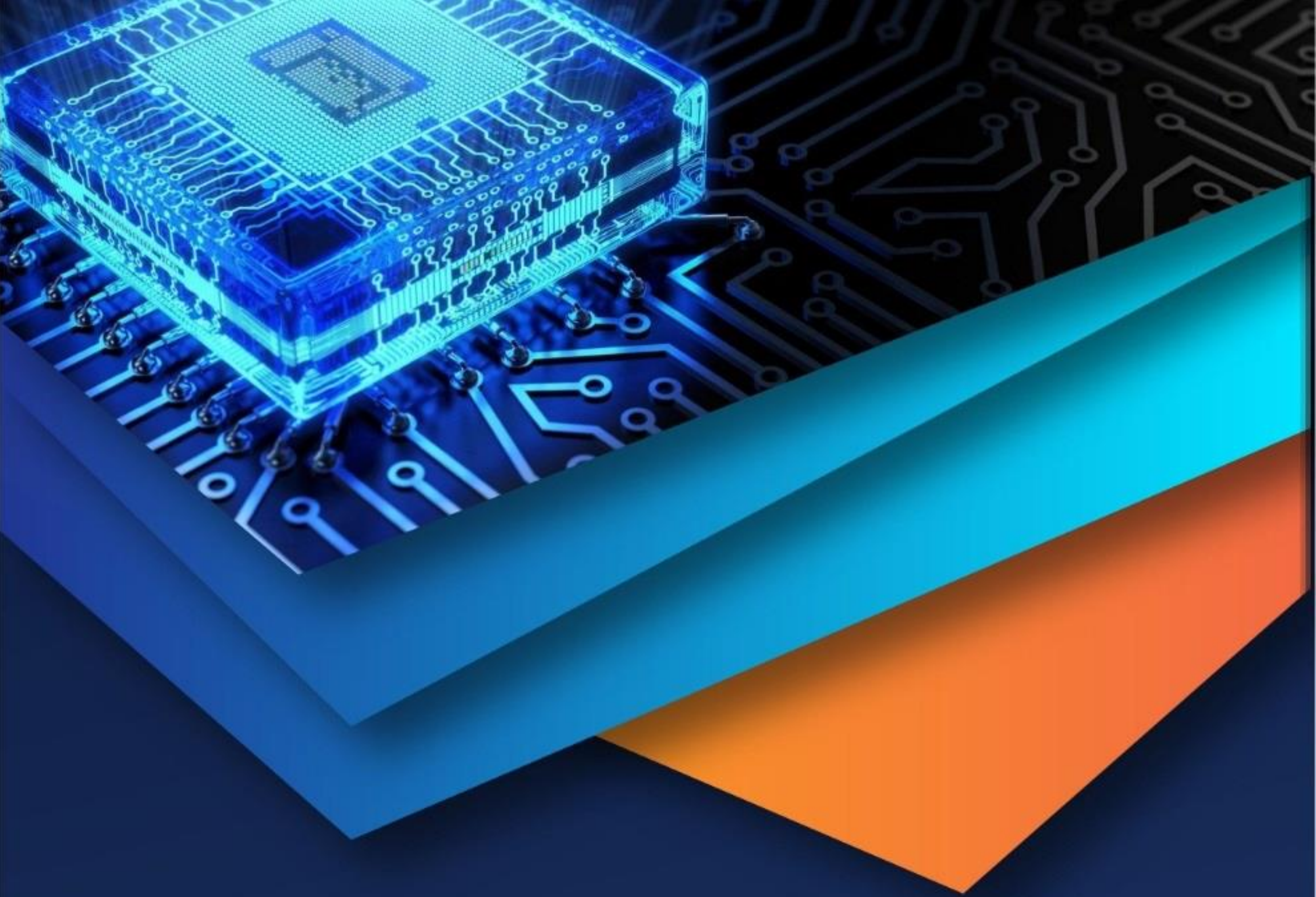
[6] S. Ertekin, J. Huang, and C. L. Giles, “Active learning for class imbalance problem,” The 30th annual international ACM SIGIR conference on Research and development in information retrieval, pp. 823–824, 2007.

[7] M.WasikowskiandX.-w.Chen,“Combatingthesmallsampleclassimbalance problem using feature selection,” IEEE Transactions on knowledge and data engineering, vol. 22, no. 10, pp. 1388–1400, 2010.

[8] S. Wang and X. Yao, “Multiclass imbalance problems: Analysis and potential solutions,” IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 42, no. 4, pp. 1119–1130, 2012.

[9] R. J. Bolton and D. J. Hand, “Statistical fraud detection: A review,” Statistical science, pp. 235–249, 2002.

[10] D. J. Weston, D. J. Hand, N. M. Adams, and C. Whitrow, “Plastic card fraud detection using peer group analysis,” vol. 2, pp. 45–62, 2008.



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY