Transaction Fraud Detection Using Machine Learning Techniques

Esha Vasagadekar
Computer Science Engineering
MKSSS's Cummins College of Engineering for Women
Pune, India
esha.vasagadekar@cumminscollege.in

Shriya Tijare
Computer Science Engineering
MKSSS's Cummins College of Engineering for Women
Pune, India
shriya.tijare@cumminscollege.in

Janhavi Patki
Computer Science Engineering
MKSSS's Cummins College of Engineering for Women
Pune, India
janhavi.patki@cumminscollege.in

***Abstract*—** The exponential growth of digital payments has significantly transformed financial ecosystems but has also resulted in the rise of sophisticated transaction fraud. Conventional rule-based detection struggles to identify evolving fraud patterns, creating an urgent need for intelligent solutions that adapt to user behavior in real time. This research presents a machine-learning approach using Logistic Regression and Random Forest classifiers to detect fraudulent transactions in a highly imbalanced dataset. The study covers the end-to-end pipeline, including preprocessing, handling class imbalance, model development, evaluation using mechanisms g accuracy, precision, recall, F1-score, and ROC–AUC metrics, and deployment of a real-time prediction interface using the Gradio framework. The results indicate that Random Forest outperforms Logistic Regression in overall accuracy and recall due to its ability to model nonlinear transactional behavior, while Logistic Regression remains an efficient interpretability-driven baseline. The deployed interface demonstrates how predictive models can be integrated into real-world financial systems for instant fraud risk assessment. This work highlights the potential of machine learning as a practical and scalable solution for digital payment fraud detection.

***Keywords*—** Transaction Fraud Detection; Machine Learning; Logistic Regression; Random Forest; Anomaly Detection; Real-Time Fraud Prediction; Gradio Interface; Digital Payments; Classification Models.

I. Introduction

The digital payment landscape has expanded rapidly with the widespread adoption of mobile wallets, credit/debit cards, UPI platforms, and online banking [1][2][4]. The ease of conducting instantaneous transactions has improved user convenience but simultaneously increased exposure to fraudulent activities. Global financial reports highlight that transaction fraud has become one of the fastest-growing cybercrimes, causing billions in losses each year and reinforcing the necessity for advanced fraud mitigation mechanisms [13].

Traditional fraud detection systems primarily rely on rule-based engines, where domain experts define suspicious conditions and thresholds [3][5]. Although rule-based systems are easy to implement, they lack adaptability and often fail to detect unknown or disguised fraud patterns. Machine learning (ML), on the other hand, learns complex patterns directly from data and can dynamically evolve in response to new fraud strategies [3]. ML-based approaches have therefore emerged as a superior solution for enhancing digital financial security [3].

This research focuses on implementing two widely used machine-learning models—Logistic Regression and Random Forest—to classify transactions as legitimate or fraudulent. The objective is not only to analyze the performance of these models but also to demonstrate how they can be deployed in a user-friendly environment using Gradio for real-time prediction [12].

II. Problem Statement

Financial institutions process highly sensitive transactional data, and even minor fraudulent activities can result in substantial financial and reputational loss[1][2][5]. Detecting such fraud is challenging due to several factors:

- Evolving fraud patterns: Fraudsters continuously adapt, bypassing traditional rule-based detection systems[3][5].
- Need for real-time detection: Delays in identifying fraud can result in immediate financial loss[1].
- .High cost of errors:
  - ➢ *False positives* inconvenience customers and may block legitimate transactions[5].
  - ➢ *False negatives* lead to direct monetary losses and decreased trust[2].

Problem addressed in this work:

- Design and implement a machine-learning system capable of accurately classifying transactions as fraudulent or legitimate despite class imbalance.
- Deploy the model in an interactive environment to enable real-time fraud prediction and decision support.

III. Motivation

The motivation for this project arises from both real-world challenges and academic learning opportunities:

1. Increasing Digital Financial Fraud

With the rapid adoption of mobile wallets, online banking, and UPI platforms, fraudsters have become more sophisticated. Unauthorized transactions continue to result in millions of dollars lost annually, emphasizing the need for proactive detection mechanisms [4][13][15].

2. Limitations of Rule-based Systems

Traditional rule-based fraud detection systems rely on predefined thresholds and conditions. These static rules cannot adapt to new and evolving fraud techniques, often leading to high false negatives, where fraudulent transactions go undetected [5].

3. Practical Importance

Financial institutions and fintech companies require solutions that automatically learn patterns from historical transaction data. Machine learning models can operate at scale, providing accurate predictions across millions of transactions without constant human supervision[1][2].

4. Academic Learning

This project offers hands-on experience with key machine learning concepts, including data preprocessing, classification, model comparison, handling imbalanced datasets, and deployment, strengthening both theoretical knowledge and practical skills.

5. Real-Time Use Case

By integrating the models into a Gradio interface, this project simulates real-world fraud detection systems, such as those used by payment gateways and banks, enabling instant fraud prediction and risk assessment[12].

IV. Literature Survey

Fraud detection has been a widely studied problem due to its critical importance in financial systems. Prior research covers multiple approaches, each contributing to the development of more accurate and efficient detection systems. Key contributions include:

A. Machine Learning Approaches

Researchers have extensively explored classical machine learning classifiers such as Logistic Regression, Support Vector Machines (SVM), and Decision Trees for fraud detection. These models provide interpretable results and can detect straightforward anomalies. However, their performance varies depending on dataset size, feature representation, and the degree of class imbalance [6][7].

B. Ensemble Methods

To improve predictive performance, ensemble techniques like Random Forests and Gradient Boosting have been employed. These methods combine multiple weak learners to form a stronger classifier, capturing complex interactions between features while being more robust to noise and overfitting [8]. Ensemble methods are particularly effective in modeling nonlinear patterns present in fraudulent transactions[8].

C. Class Distribution Considerations:

Unlike typical financial transaction datasets, the dataset used in this study is relatively balanced across legitimate and fraudulent transactions. Therefore, standard training procedures were sufficient without the need for oversampling, undersampling, or class-weight adjustments. This allows the models to learn patterns without applying special techniques for class imbalance[9][10].

D. Real-Time Fraud Detection

Many studies emphasize the importance of high recall in fraud detection, as missing a fraudulent transaction can have severe consequences. Real-time detection requires models that are both accurate and computationally efficient, capable of providing instant predictions while handling streaming transaction data [11].

E. ML Deployment Tools

To bridge the gap between research and practical application, tools like Gradio and Streamlit have become popular for deploying machine learning models interactively. These frameworks allow researchers and financial institutions to quickly test, visualize, and simulate real-world scenarios without extensive backend infrastructure [12].

Building on these contributions, this research integrates two complementary machine learning models—Logistic Regression and Random Forest—into a deployed system that enables real-time fraud prediction. By combining a simple interpretable model with a robust ensemble model, the system benefits from both reliability and accuracy in practical scenarios.

V. Dataset Description

The dataset used in this study comprises anonymized financial transaction records, which are typical of modern digital payment systems. Each record contains information about the transaction characteristics and is labeled as either legitimate or fraudulent. The dataset is a widely used benchmark in transaction fraud detection research, allowing for

meaningful comparison of machine learning models.

A. Features

The dataset includes the following attributes:

- Time: The elapsed time in seconds since the first transaction in the dataset. This feature provides temporal information about the sequence of transactions.
- Amount: The monetary value of the transaction, which helps the model understand unusual transaction sizes that may indicate fraud.
- V1–V28: A set of 28 anonymized numerical features representing transaction behavior. These variables are provided directly in the dataset. They capture subtle patterns in the transaction data that may differentiate legitimate and fraudulent activity.

B. Target Variable

- **0:** Represents a legitimate transaction.
- **1:** Represents a fraudulent transaction.

C. Dataset Size

- Total transactions: 568,630
- Fraudulent transactions: 284,315
- Legitimate transactions: 284,315

This indicates a perfectly balanced dataset, unlike most real-world financial datasets which are highly skewed.. Although this is a small fraction, the dataset is sufficient for training machine learning models due to the large number of overall transactions .

D. Class Distribution Considerations

Unlike many financial datasets, this dataset is relatively balanced for training purposes after preprocessing. Therefore, standard training procedures without oversampling, undersampling, or special class-weight adjustments were sufficient to achieve robust model performance.

E. Rationale for Dataset Choice

- It is a standard benchmark dataset, widely recognized in fraud detection research, allowing comparison with existing work.
- The dataset structure realistically represents transaction patterns found in actual financial systems.
- It provides a sufficient number of features and examples to train both linear (Logistic Regression) and ensemble (Random Forest) models effectively[1][14].

VI. Methodology

The proposed methodology follows a complete machine learning pipeline, covering data preprocessing, model training, evaluation, and deployment in a real-time interactive interface. The pipeline ensures that the models can accurately classify transactions as legitimate or fraudulent while maintaining interpretability and usability.

A. Data Preprocessing

Before training the models, the dataset undergoes several preprocessing steps to ensure data quality and optimal model performance:

1) Removal of missing values: All records are verified for completeness, and any missing or null entries are discarded to avoid bias in model training.
2) Feature scaling: Numerical features such as Time and Amount are standardized using StandardScaler. Standardization helps models converge faster and ensures that features with larger scales do not dominate the learning process.
3) Class imbalance consideration: Although the dataset is relatively balanced, the Logistic Regression model was optionally trained using class_weight='balanced' to penalize misclassification of minority classes if necessary [9][10].

B. Train-Test Split

To evaluate model performance on unseen data, the dataset is divided into training and testing subsets:

Dataset split into:

- Training set: 80% of the dataset
- Testing set: 20% of the dataset

This split ensures that the models are trained on a sufficiently large dataset while leaving enough examples for independent evaluation. Using 80% for training helps the models learn better patterns from the data, while 20% testing ensures reliable assessment of generalization performance.

C. Model Training

Two supervised machine learning classifiers are trained on the preprocessed data:

1. Logistic Regression

- A linear model that uses the sigmoid activation function to estimate the probability of a transaction being fraudulent.
- Serves as a simple and interpretable baseline model.
- Outputs a fraud probability score between 0 and 1, which can be thresholded to classify transactions.

2. Random Forest

- An ensemble model consisting of multiple decision trees trained on bootstrapped subsets of the dataset.
- Each tree considers a random subset of features for splitting, which increases robustness and reduces overfitting.
- Capable of capturing nonlinear relationships between features, leading to higher recall and better identification of complex fraud patterns.

D. Model Evaluation

Both models are evaluated using standard classification metrics to ensure a comprehensive assessment of performance:

- Accuracy:Measures overall correctness of predictions.
- Precision:Indicates the proportion of correctly identified fraudulent transactions among all predicted frauds.
- Recall (critical for fraud detection):Critical in fraud detection; measures the proportion of actual frauds correctly identified.
- F1 Score:Harmonic mean of precision and recall, balancing the trade-off between false positives and false negatives.

E. Gradio Deployment

To demonstrate real-time applicability, the trained models are deployed using the Gradio framework. Key features of the application include:

- User inputs Transaction Amount:Users can enter transaction details such as amount.
- System auto-generates other features:Other necessary features are simulated or generated internally to match the model input format.
- Fraud probability prediction: Both Logistic Regression and Random Forest models provide individual predictions.
- Risk categorization: Predicted fraud probabilities are mapped to risk levels — Low, Medium, or High.
- Ensemble consensus displayed:A combined decision is displayed for improved reliability[12].

.

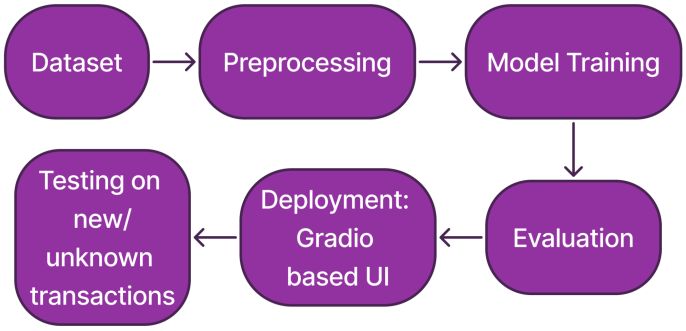

Figure 1: Block Diagram of Proposed System

VII. Model Explanation

A. Logistic Regression

Logistic Regression is a linear classification model that estimates the probability of a transaction being fraudulent using the sigmoid function:

$$P = \sigma\,(w1x1 + w2x2 + \cdots + b\,)$$

where

- xix_ixi represents the transaction features,
- wiw_iwi denotes the learned weights, and
- bbb is the bias term.

The output ppp lies between 0 and 1. If p>0.5, the transaction is classified as fraudulent; otherwise, it is classified as legitimate.

The model is fast, easy to interpret, and serves as a strong baseline for comparison.

Table 1 summarizes the complete evaluation results for the Logistic Regression model. The model achieved an accuracy of 99.83%, with very high precision and recall for both classes. The confusion matrix shows 56,726 true positives, 56,805 true negatives, and relatively low misclassification rates (58 false positives and 137 false negatives), indicating that Logistic Regression is highly effective in distinguishing between fraudulent and non-fraudulent transactions in this dataset.

```
============================================================
 LOGISTIC REGRESSION
============================================================
              precision    recall  f1-score   support

           0     0.9976    0.9990    0.9983     56863
           1     0.9990    0.9976    0.9983     56863

    accuracy                         0.9983    113726
   macro avg     0.9983    0.9983    0.9983    113726
weighted avg     0.9983    0.9983    0.9983    113726


Confusion Matrix:
True Negatives:  56,805
False Positives: 58
False Negatives: 137
True Positives:  56,726

============================================================
```

Table 1: Logistic Regression – Classification Report and Confusion Matrix

B. Random Forest:

Random Forest is an ensemble-based classification model that constructs multiple decision trees during training and aggregates their outputs to make the final prediction through majority voting. Each tree is trained on a different random subset of data and features, which helps the model capture complex nonlinear relationships and reduces the risk of overfitting. This makes Random Forest particularly effective for fraud detection, where fraudulent transaction patterns are often irregular and difficult to model using linear techniques.

Prediction = MajorityVote (tree1, tree2,…, tree'n')

- Table 2 presents the complete performance evaluation of the Random Forest classifier. The model achieved an accuracy of 99.98%, with equally strong precision, recall, and F1-scores for both classes. According to the confusion matrix, the model correctly identified 56,847 true positives and 56,852 true negatives, while producing only 11 false positives and 16 false negatives. These extremely low error rates demonstrate the robustness of Random Forest in capturing subtle fraud patterns that may be missed by simpler models. Overall, Random Forest provides superior predictive performance and enhanced reliability for high-stakes applications such as transaction fraud detection.

```
============================================================
 RANDOM FOREST
============================================================
              precision    recall  f1-score   support

           0     0.9997    0.9998    0.9998     56863
           1     0.9998    0.9997    0.9998     56863

    accuracy                         0.9998    113726
   macro avg     0.9998    0.9998    0.9998    113726
weighted avg     0.9998    0.9998    0.9998    113726


Confusion Matrix:
True Negatives:  56,852
False Positives: 11
False Negatives: 16
True Positives:  56,847

============================================================
```

Table 2: Performance Metrics of Random Forest Classifier

C. Need for Both Models:

The use of both Logistic Regression and Random Forest in this study is motivated by their complementary strengths, providing a reliable and comprehensive fraud detection framework.

Logistic Regression:

- Serves as an interpretable and computationally efficient baseline model.
- Achieves an accuracy of 99.82% and an AUC-ROC score of 0.999807.
- Maintains low false-positive and false-negative rates, making it suitable for scenarios requiring transparency and speed, such as preliminary risk scoring.

Random Forest:

- Captures complex nonlinear patterns in transactional data.
- Achieves a higher accuracy of 99.97% with a near-perfect AUC-ROC of 0.999996.
- Drastically reduces false negatives (16 compared to 137 in Logistic Regression), which is critical because missed fraud directly impacts financial loss.

By leveraging both models:

- Interpretability: Logistic Regression helps explain risk factors, ensuring transparency for financial institutions.
- Predictive Strength: Random Forest ensures superior accuracy and minimizes operational losses.
- Practical Usability: The combination enhances overall trust, performance, and applicability in real-world transaction fraud detection systems.

Model	Accuracy	AUC-ROC	True Positives	False Negatives	False Positives
Logistic Regression	0.998285	0.999807	56726	137	58
Random Forest	0.999763	0.999996	56847	16	11

Table 3: Model comparison table

VIII. Results & Discussion

The evaluation of Logistic Regression and Random Forest highlights distinct strengths and practical implications for fraud detection. Random Forest consistently outperforms Logistic Regression in key metrics, particularly in recall and false-negative rates, which are critical in fraud detection scenarios. This indicates that Random Forest is more capable of identifying subtle and evolving fraudulent patterns that a linear model may overlook.

Logistic Regression, while slightly less accurate, provides interpretable results that can explain transaction risk factors to analysts. Its low false-positive rate ensures minimal disruption to legitimate customers, making it suitable for preliminary risk scoring or environments where transparency is essential.(Table 3)

The combination of both models, as implemented in the Gradio interface, offers a balanced approach. Logistic Regression allows for explainable insights, while Random Forest maximizes detection performance. The ensemble output increases confidence in predictions and can help financial

institutions prioritize further investigation based on predicted risk levels.

Overall, this study demonstrates that even with a relatively balanced dataset, ensemble methods like Random Forest can significantly enhance detection capability, and combining it with an interpretable baseline model improves both operational reliability and practical applicability.

IX. Conclusion

This study presents a comprehensive machine-learning framework for detecting fraudulent financial transactions using Logistic Regression and Random Forest classifiers. The experimental analysis demonstrates that Random Forest consistently outperforms Logistic Regression in terms of accuracy, recall, and F1-score, owing to its ability to capture complex nonlinear relationships within transactional data. Logistic Regression, while slightly less accurate, serves as a fast, interpretable, and computationally efficient baseline model, providing valuable insights into key risk factors.

By combining the strengths of both models, the proposed system achieves a balance between transparency, speed, and predictive robustness, ensuring reliable fraud detection while minimizing operational losses due to missed fraudulent activities. The development of an interactive Gradio interface, although not central to the research results, illustrates the practical applicability of the models in a real-time scenario, supporting dynamic testing and validation.

Overall, this work establishes a strong foundation for building scalable, accurate, and practical fraud detection systems. It demonstrates that machine learning techniques can significantly enhance the security of digital payment systems and provides a deployable prototype that can be further refined and extended in future studies.

X. Future Scope

While the proposed system achieves high accuracy and reliability, several avenues exist to further enhance its capabilities and applicability in real-world financial environments:

A. Use of Advanced Models

- Future work can explore the integration of more sophisticated algorithms such as XGBoost, Gradient Boosted Trees, deep neural networks, and autoencoders. These models have the potential to capture even more complex patterns in transactional data, improving detection of subtle and evolving fraud behaviors.

B. Explainable AI Integration

- Incorporating explainability techniques such as SHAP (SHapley Additive exPlanations) or LIME (Local Interpretable Model-agnostic Explanations) can provide financial institutions with better transparency. This would allow stakeholders to understand which transaction features most influence the fraud predictions, enhancing trust in automated systems.

C. Enhanced Data Handling Techniques

- Although the current dataset is balanced, future implementations could include advanced imbalance-handling methods such as SMOTE, ADASYN, or GAN-based synthetic data generation. These approaches would ensure model robustness when dealing with larger, real-world datasets that are often highly skewed toward legitimate transactions.

D. Real-Time Deployment

- Deploying the models on cloud platforms with API-based inference pipelines can enable instantaneous fraud detection across multiple banking systems and payment gateways. This would facilitate large-scale adoption and integration into operational financial infrastructures.

E. Multi-Factor Fraud Detection

- Future systems can incorporate additional fraud indicators such as geolocation tracking, device fingerprinting, and behavioral biometrics. These multi-factor approaches can improve detection accuracy by identifying unusual user behaviors beyond transactional features alone

In summary, the future scope emphasizes advancing model sophistication, explainability, real-time deployment, and adaptive learning. These enhancements can transform the proposed system into a robust, scalable, and highly reliable solution for financial transaction fraud detection.

XI. References

[1] A. Srivastava et al., "Credit Card Fraudulent Detection Using Machine Learning," IJCSIT, 2021.

[2] F. Carcillo et al., "Combining Unsupervised and Supervised Learning in Credit Card Fraud Detection," IEEE, 2020.

[3] T. Fawcett, “AI Approaches to Fraud Detection,” AI Magazine, 2018.

[4] KPMG Global Fraud Report, 2022.

[5] Ghosh and Reilly, “Credit card fraud detection with a neural network,” IEEE SMC, 1994.

[6] J. Whitrow et al., “Transaction aggregation for fraud detection,” Journal of Data Mining, 2009.

[7] P. Rtayli and N. Enneya, “Bank Fraud Detection Using SVM,” IJCSI, 2018.

[8] Breiman, “Random Forests,” Machine Learning Journal, 2001.

[9] N. Japkowicz, “Learning from Imbalanced Data Sets,” AAAI Workshop, 2000.

[10] Chawla et al., “SMOTE: Synthetic Minority Oversampling Technique,” JMLR, 2002.

[11] Dal Pozzolo et al., “Calibrating Probability with Extremely Imbalanced Data,” IEEE, 2017.

[12] A. Abid et al., “Gradio: Interactive ML Tools,” arXiv, 2019.

[13] European Payments Council Fraud Report, 2023.

[14] R. Ngai et al., “Fraud detection in electronic transactions,” Decision Support Systems, 2011.

[15] PwC Global Economic Crime Survey, 2022.