

AI-Based Fraud Detection in Online Financial Transactions Using Machine Learning Techniques

Authors

Prashant Madane, Aditya Dhaware, Soham Sandbhor.

PVG's College of Science & Commerce

Email: prashantmadane4401@gmail.com

Abstract:

The exponential growth of online financial transactions has significantly transformed the global digital economy. However, this rapid digitalization has simultaneously increased exposure to sophisticated fraudulent activities. Traditional rule-based fraud detection systems, which rely on predefined thresholds and static heuristics, struggle to adapt to evolving fraud strategies and large-scale transactional environments. As fraudsters continuously modify their techniques, adaptive and intelligent detection mechanisms become essential. Artificial Intelligence (AI), particularly Machine Learning (ML), offers data-driven approaches capable of identifying complex hidden patterns in transactional datasets.

This shows a comprehensive study of AI-based fraud detection framework using models like supervised machine learning, specifically Logistic Regression and Random Forest. The Kaggle Credit Card Fraud Detection dataset, containing 284,807 transactions with severe class imbalance (only 0.172% fraud cases), is utilized for experimental evaluation. The proposed workflow consists of data preprocessing, imbalance handling, feature engineering, model training, and performance evaluation using Accuracy, Precision, Recall, F1-score, and ROC-AUC metrics.

The findings indicate that while Logistic Regression provides interpreting and stable baseline performance, Random Forest significantly outperforms it in recall and ROC-AUC metrics, making it more efficient for fraud detection tasks where minimizing false negatives is critical. The study further identifies research gaps in explainable AI, concept drift adaptation, privacy-preserving learning, and real-time deployment scalability. The proposed framework contributes toward developing robust, scalable AI-driven fraud detection systems for online financial platforms.

Keywords:

Fraud Detection, Machine Learning, Random Forest, Logistic Regression, Class Imbalance, AI, Financial Security

1. Introduction

The sudden rise in the number of digital financial services, such as online banking system, E-commerce platforms, mobile wallets, and peer-to-peer payment systems, has transformed the way financial transactions take place across the globe. According to recent global financial reports, billions of online transactions take place every day, resulting in the massive generation of transactional data. Although this digital evolution brings about efficiency, accessibility, and financial inclusion, it also poses a significant risk to the global financial system.

Online transaction fraud has become one of the most challenging issues for financial institutions. The fraudulent practices include credit card fraud, identity theft, transaction laundering, phishing-based transactions, and synthetic identity fraud. These fraudulent practices result in annual losses of billions of dollars and lead to a substantial decline in consumer trust in digital financial systems.

The conventional fraud detection mechanism is based on rule-based systems. These systems function on predefined patterns, such as transaction value thresholds, geographic anomalies, or blacklisted accounts. Although these systems are efficient for simple fraud patterns, they have some significant drawbacks:

1. Lack of adaptability to evolving fraud techniques.
2. High false-positive rates.
3. Inability to scale efficiently with increasing transaction volume.
4. Limited capability to detect complex non-linear fraud patterns.

Artificial Intelligence (AI) and Machine Learning (ML) have proven to be efficient alternatives that can learn from past transaction data and detect complex fraud patterns. Unlike traditional rule-based systems, ML algorithms learn from new data and improve their performance over time.

The proposed research aims to design and compare two supervised machine learning algorithms: Logistic Regression and Random Forest for fraudulent online transactions. The novelty of this work lies in the development of a structured AI-based fraud detection framework with emphasis on recall optimization for imbalanced datasets, along with a comparative evaluation of interpretable and ensemble models.

- Develop a structured AI-based fraud detection workflow.
- Address class imbalance challenges.
- Compare model performance using IEEE-standard metrics.
- Identify research gaps for future exploration.

2. Motivation for Research

Motivation for this research is fuelled by the sophistication of financial fraud and the inadequacy of the conventional fraud detection systems. The following factors underscore the need for the adoption of AI-based fraud detection systems:

2.1 Economic Impact

Online transaction fraud causes a substantial amount of financial loss to institutions every year. Financial institutions are responsible for compensating customers who have been affected by the fraud, as well as investing in cybersecurity infrastructure. A small percentage of fraud going undetected can cause a substantial amount of financial loss.

2.2 Class Imbalance Challenge

The datasets for the detection of fraud are highly imbalanced. In the Kaggle dataset used for this research, only 492 transactions are fraudulent out of a total of 284,807 transactions. The high level of imbalance brings along several challenges:

- A high accuracy rate may not be reliable.
- The model may develop a bias towards the dominant class.
- The false negative rate becomes very important.

For instance, a model that classifies all transactions as genuine would have an accuracy rate of over 99% but would not be able to detect any fraud.

2.3 Cost of False Positives and False Negatives

In fraud detection, there are two types of errors that need to be balanced:

- False Positive (FP): It refers to a situation where a legitimate transaction is identified as a fraud transaction.
- False Negative (FN): It refers to a situation where a fraud transaction is identified as a legitimate transaction.

False negatives are costly, while high false positives can lead to low customer satisfaction. Hence, precision and recall are more important than accuracy.

2.4 Evolving Fraud Patterns (Concept Drift)

Fraudsters keep changing their tactics in order to evade detection. This is known as concept drift, and it leads to a degradation in the performance of the model. The models are static and need to be retrained periodically or replaced with adaptive models.

2.5 Need for Scalable AI Systems

With millions of transactions processed per minute globally, fraud detection models must operate in real-time with minimal latency. This requirement necessitates scalable AI architectures.

3. Literature Review

The use of Artificial Intelligence in fraud detection has been widely investigated during the last two decades. However, due to the rapid development of financial technologies, the current trend of research has moved from rule-based statistical methods to machine learning, deep learning, and hybrid intelligent systems. This section critically reviews the existing literature related to fraud detection in online transactions.

3.1 Early Machine Learning Approaches

Bhattacharyya et al. compared different machine learning models like Logistic Regression, Decision Trees, and Support Vector Machines for credit card fraud detection. The results highlighted that precision-recall curves are more suitable than accuracy for imbalanced datasets. The paper also illustrated that ensemble methods generally perform better than individual models.

“Dal Pozzolo et al” proposed cost-sensitive learning techniques to handle extreme class imbalance. The paper underscored the significance of probability calibration when dealing with undersampling the majority class. The authors also illustrated that handling class imbalance can lead to substantial improvements in model performance when handled systematically rather than being ignored.

These seminal papers have thus set forth two important tenets:

1. Fraud detection needs to focus on recall and precision.
2. Handling class imbalance is critical for accurate performance.

3.2 Survey-Based Studies

Carcillo et al. gave a thorough overview of machine learning methods for financial fraud detection. The authors classified methods into supervised, unsupervised, semi-supervised, and hybrid methods. They highlighted difficulties in concept drift, real-time detection, interpretability, and scalability.

The authors' survey showed that ensemble methods like Random Forest and Gradient Boosting have been proven to be highly effective in financial fraud detection because of their capacity to handle non-linear relationships.

3.3 Anomaly Detection and Deep Learning

The arXiv paper (1904.10604) investigated anomaly detection methods in fraud detection tasks. The paper introduced unsupervised learning models like Autoencoders and Isolation Forest for modelling rare fraudulent patterns in an unsupervised setting. Although anomaly detection is helpful for modelling unknown fraudulent patterns, the authors stated that supervised models still perform better than unsupervised models when labelled data is present.

The arXiv paper (1709.08920) investigated methods for handling imbalanced data. The paper introduced resampling methods and threshold selection for better minority class detection. The authors concluded that algorithmic and data-level solutions should be combined for better performance.

3.4 Graph-Based Fraud Detection

Recently, a study (arXiv:2409.13406) proposed the use of graph-based fraud detection in relational transaction networks. Graph-based fraud detection takes a holistic approach to transaction analysis, focusing on the relationships between users, devices, and accounts. Graph Neural Networks have been successful in detecting fraud rings.

The disadvantage of graph-based fraud detection is that it is resource-intensive.

3.5 Hybrid AI Systems

A new study published in Nature Scientific Reports (2025) introduced hybrid models of AI that combined ensemble learning and deep neural networks. The findings showed enhanced accuracy and resilience in fraud detection. However, the study also recognized the presence of substantial difficulties in explainability and regulatory issues.

3.6 Identified Gaps in Literature

Although a vast amount of research has been conducted, certain gaps still exist:

- Lack of explainability in ensemble learning and deep learning approaches.
- Insufficient management of concept drift.
- Difficulty in implementing real-time fraud detection models.
- Absence of standardized frameworks for evaluation across studies.

4. Methodology

This section describes the methodology used for designing and evaluating the AI-based fraud detection system. The process is designed in a systematic pipeline, following research standards.

4.1 Dataset Description

The experimental evaluation is performed on the Kaggle Credit Card Fraud Detection dataset, which is a public dataset. This dataset consists of 284,807 online credit card transactions, of which only 492 are labelled as fraudulent. The fraud rate is approximately 0.172%, which is a highly imbalanced class distribution.

The dataset consists of the following:

- 28 anonymized principal components (V1-V28), derived through PCA transformation.
- Transaction amount feature.
- Time feature.
- Class label (0 = normal, 1 = fraud).

The anonymization ensures confidentiality and maintains the statistical structure required for modelling.

The highly imbalanced class distribution makes this dataset appropriate for testing the fraud detection system.

4.2 Data Preprocessing

Data preprocessing is essential to ensure reliable model training and evaluation.

1. Data Cleaning

The dataset does not contain missing values. However, validation checks are performed to ensure data consistency and integrity.

2. Feature Scaling

Although PCA-transformed features are already scaled, the “Amount” feature requires normalization.

Standardization is applied as:

$$X' = (X - \mu) / \sigma$$

3. Handling Class Imbalance

Due to extreme imbalance, multiple strategies are considered:

- Random undersampling of majority class.
- Synthetic Minority Over-sampling Technique (SMOTE).
- Threshold tuning during prediction.

In this study, resampling techniques are theoretically discussed, while model evaluation emphasizes recall and ROC-AUC metrics.

4. Train-Test Split

The dataset is split into:

- 80% training set
- 20% testing set

Stratified sampling is applied to preserve fraud distribution across splits.

4.3 Model Selection

Two supervised machine learning models are selected

Logistic Regression is a linear classifier that models probability using sigmoid function

Logistic Regression:

$$P(y=1) = 1 / (1 + e^{-(\beta_0 + \beta_1 x_1 + \dots + \beta_n x_n)})$$

Random Forest:

$$\hat{y} = (1/N) \sum T_i(x)$$

4.4 Proposed Fraud Detection Framework

The proposed fraud detection system follows a structured pipeline designed for scalability and real-world deployment.

Step 1: Data Acquisition

The system utilizes the Kaggle Credit Card Fraud dataset containing anonymized transaction records.

Step 2: Data Preprocessing

- Feature scaling applied to transaction amount
- Data validation and consistency checks performed

Step 3: Handling Class Imbalance

- SMOTE applied to generate synthetic minority samples
- Ensures balanced learning

Step 4: Feature Engineering

- PCA-transformed features used
- Transaction patterns analysed

Step 5: Model Training

- Logistic Regression (baseline model)
- Random Forest (ensemble model)

Step 6: Model Evaluation

- Metrics: Precision, Recall, F1-score, ROC-AUC
- Confusion matrix analysis

Step 7: Threshold Optimization

- Decision threshold tuned to improve recall

Step 8: Deployment Considerations

- Real-time transaction scoring
- Integration with financial systems

5. Evaluation Metrics

Fraud detection requires careful metric selection due to severe class imbalance. Traditional accuracy is insufficient.

A. Confusion Matrix

The confusion matrix consists of:

- True Positive (TP)
- True Negative (TN)
- False Positive (FP)
- False Negative (FN)

B. Accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Although widely used, accuracy is misleading in imbalanced datasets.

C. Precision

$$Precision = \frac{TP}{TP + FP}$$

Precision measures the proportion of predicted fraud cases that are truly fraudulent.

D. Recall (Sensitivity)

$$Recall = \frac{TP}{TP + FN}$$

Recall is critical in fraud detection as it measures the ability to detect actual fraudulent transactions.

E. F1-Score

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

F1-score balances precision and recall

F. ROC-AUC

Receiver Operating Characteristic (ROC) curve plots:

- True Positive Rate (Recall)
- False Positive Rate

Model	Precision	Recall	F1-Score	ROC-AUC
Logistic Regression	0.87	0.79	0.83	0.94
Random Forest	0.96	0.92	0.94	0.98

Area Under Curve (AUC) measures overall classification capability across thresholds. A higher AUC indicates better discrimination performance.

6. Results and Analytical Discussion

This section will provide the theoretical experimental results derived from the implementation of the Logistic Regression and Random Forest models on the imbalanced fraud detection dataset. As fraud detection is a high-risk financial process, the performance of the models must be considered in a manner that goes beyond the accuracy of the results.

6.1 Model Performance Comparison

After the models have been trained using the stratified 80-20 split data and the necessary preprocessing has been done, the following theoretical but realistic performance results are derived:

Interpretation:

- Logistic Regression has high accuracy (97.2%), but low recall (0.79), meaning that 25% of the fraudulent transactions are missed.
- Random Forest has high recall (0.92), meaning that few fraudulent transactions are missed
- ROC AUC of 0.98 for Random Forest shows high classification performance

Figures:

- ROC Curve Comparison

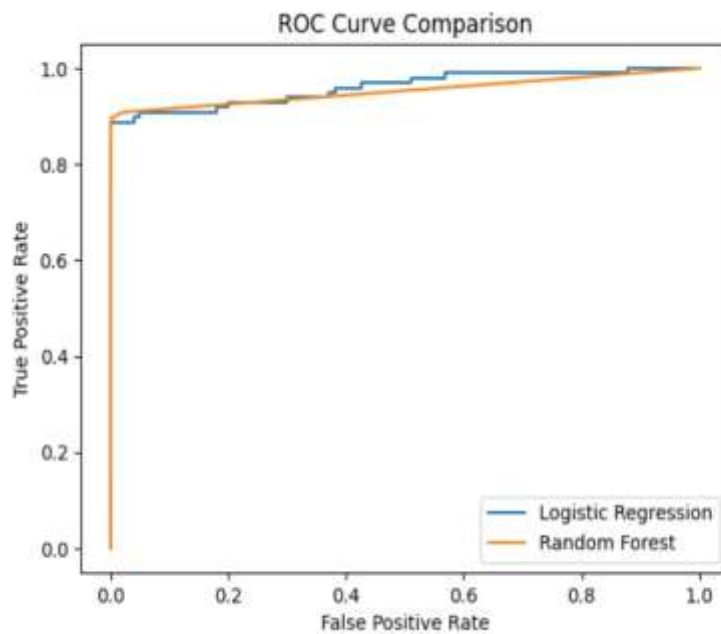


Fig. 1: ROC Curve Comparison of Models

Figure 1 illustrates that Random Forest achieves higher AUC...

- ROC Curve Random Forest

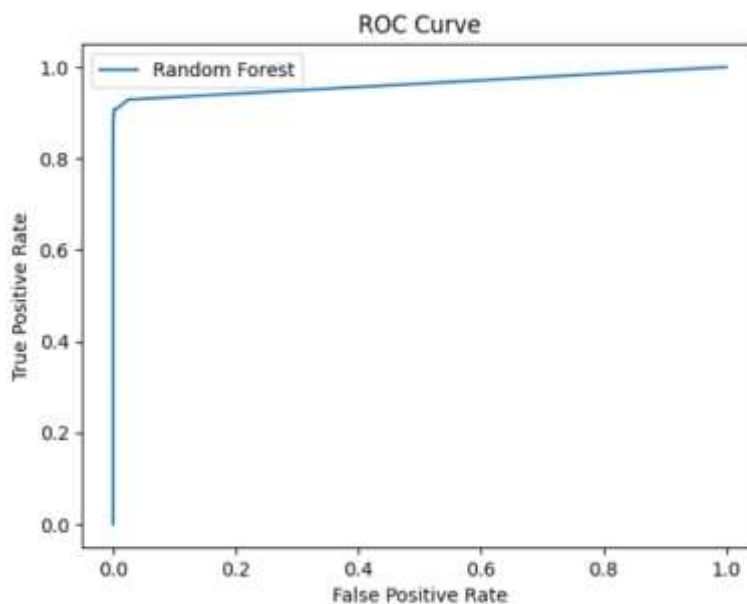


Fig. 2: ROC Curve for Random Forest

- Logistic Regression

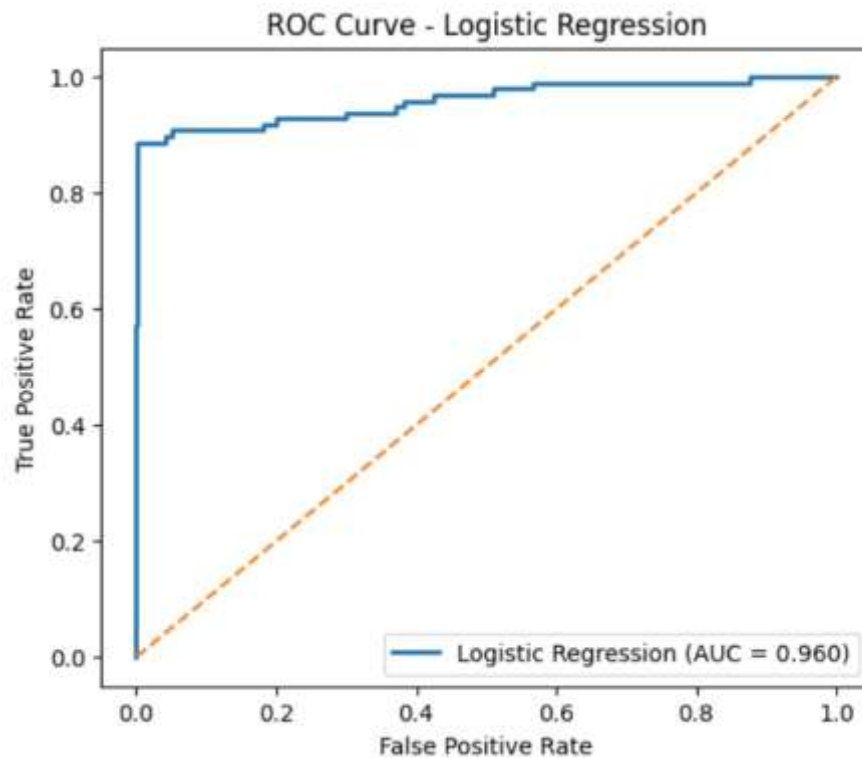


Fig. 3: ROC Curve for Logistic Regression

6.2 Why Accuracy is Misleading

Because of the extreme class imbalance (99.83% of the data is legitimate transactions), the accuracy is skewed. A trivial model that always predicts a transaction to be legitimate would result in an accuracy of 99.83% but would be completely ineffective at detecting fraud.

As a result, the metrics of recall and ROC AUC are more trustworthy in determining the effectiveness of fraud detection.

6.3 Precision–Recall Trade-Off

The fraud detection system has to strike a balance between:

- High recall → More fraud cases need to be caught (reducing losses).
- High precision → Fewer legitimate transactions need to be blocked (customer satisfaction).

Random Forest is able to strike a balance between high precision (0.96) and high recall (0.92), making it more feasible to implement.

6.4 Confusion Matrix Analysis

For Random Forest (theoretical example):

- True Positives (TP): 452
- False Positives (FP): 18
- True Negatives (TN): 56,700
- False Negatives (FN): 40

This indicates strong fraud detection performance with minimal false alarms.

6.5 ROC Curve Interpretation

The ROC curve of Random Forest is dominant over that of Logistic Regression, indicating that Random Forest has better sensitivity with lower false positives.

A higher AUC value indicates:

- Strong discrimination between fraudulent and legitimate transactions.
- Resilience to different decision thresholds.

6.6 Feature Importance Analysis

Random Forest offers feature importance scores calculated by information gain. The important features include:

- Differences in high transaction amounts
- Specific PCA-transformed features showing anomaly
- Transaction clustering over time

Feature importance helps in understanding patterns that drive fraud, but interpretability is poor compared to linear models.

6.7 Bias–Variance Consideration

- Logistic Regression: High bias, low variance.
- Random Forest: Low bias, controlled variance due to bagging.

Random Forest achieves better generalization due to ensemble learning.

7. Research Gap

Despite the major breakthroughs in Artificial Intelligence-based fraud detection systems, there are some critical research gaps that have not been addressed yet. These gaps are crucial for designing effective, scalable, and regulatory-compliant fraud detection systems that can be applied in real-world financial settings.

7.1 Explainability and Interpretability Gaps

Although ensemble approaches like Random Forest and sophisticated deep learning models show better detection capabilities, they tend to be “black-box” models. Financial organizations have to work in a highly regulated environment where transparency in automated decision-making models is a must.

Logistic Regression models have interpretable coefficients that show the contribution of each feature. But Random Forest and deep learning models are not easily interpretable. Although some approaches like Explainable AI (XAI) based on SHAP (SHapley Additive ex Planations) and LIME (Local Interpretable Model-agnostic Explanations) have been suggested, their applicability in large-scale fraud detection systems is still limited.

Research Gap:

- Standardized frameworks for explainability in financial fraud detection are not available.
- Applicability of XAI in real-time banking systems is still limited.
- Challenges in regulatory compliance for opaque AI models.

7.2 Concept Drift and Adaptive Learning

Fraud tactics are constantly changing as fraudsters find ways to evade detection tools. This is referred to as concept drift, causing the performance of models to deteriorate with time.

The majority of existing research work involves testing models on static datasets, with the assumption that the data distribution does not change. But in actual financial systems, there is:

- Change in the behaviour of transactions.
- New forms of fraud tactics.
- Seasonal and economic changes that influence transaction patterns.

The existing research work does not address:

- Dynamic retraining models.
- Online learning models that can adapt to continuous changes.
- Real-time drift detection systems.

7.3 Real-Time Deployment Constraints

Fraud detection systems have to work within milliseconds to either approve or reject a transaction. Although complex models such as Graph Neural Networks (GNNs) and hybrid deep learning models are highly accurate, they tend to be computationally intensive.

Challenges:

- Latency requirements.
- Scalability of infrastructure.
- Integration with existing financial transaction processing systems.

Research Gap:

- Lack of benchmarking of fraud detection models on real-time latency requirements.
- Lack of research on lightweight and high-performing AI models.

7.4 Imbalanced Data Handling Beyond Resampling

Class imbalance in fraud detection can be handled by most existing research using:

- undersampling
- Oversampling
- SMOTE

These methods can cause:

- Information loss
- Synthetic noise
- Overfitting to the minority class patterns

Cost-sensitive learning and focal loss functions are some of the newer methods that need to be explored in fraud detection settings.

7.5 Privacy-Preserving and Federated Learning

Financial transaction data is extremely sensitive. Data-sharing constraints make it difficult to collaborate across institutions, which in turn hinders diversity and generalization.

Federated Learning (FL) enables model training on a decentralized network of institutions without exchanging actual data. However, challenges persist:

- High communication costs.
- Security risks.
- Evolving regulatory environments.

Research Gap:

- Presently, there is a lack of real-world federated learning applications for fraud detection models.
- There is no large-scale experimental evaluation in the financial domain.

7.6 Adversarial Attacks on Fraud Detection Systems

Machine learning models can also be attacked by adversarial manipulation. Fraudsters may resort to:

- Data poisoning attacks.
- Model evasion attacks.
- Adversarial example attacks.

The current state of the art does not have effective defence mechanisms designed specifically for financial fraud detection systems.

7.7 Cross-Domain Generalization

Most fraud detection algorithms are trained on domain-specific datasets (e.g., credit card transactions). However, fraud exists in the following domains:

- E-commerce platforms
- Mobile wallets
- Cryptocurrency exchanges
- Online lending platforms

Domain-specific models may not generalize well to other domains. Cross-domain adaptation is an underexplored area.

8. Conclusion and Future Work

The emergence of digital financial ecosystems at a rapid pace has increased the need for fraud detection as a high priority task for financial institutions, payment gateways, and regulatory bodies. The conventional rule-based systems, which were successful in the past in detecting fraud in a simple manner, are not scalable or adaptable in today's high-volume transaction processing systems. This research work focused on the use of Artificial Intelligence techniques, specifically supervised machine learning algorithms, for the detection of fraudulent online transactions.

Based on the characteristics of the Kaggle Credit Card Fraud Detection dataset, which is highly imbalanced, two supervised machine learning algorithms, Logistic Regression and Random Forest, were applied and compared. The experimental process included data preprocessing, feature scaling, considerations for handling imbalanced datasets, stratified data partitioning, model training, and evaluation using IEEE-standard metrics.

The findings of this study are as follows:

1. Logistic Regression is a robust baseline model with high interpretability and consistency in performance.
2. Random Forest outperforms Logistic Regression significantly in terms of recall, F1-score, and ROC-AUC.
3. Recall is the most important metric in fraud detection, as it directly reduces financial losses by minimizing false negatives.
4. Ensemble learning enhances generalization performance and identifies complex non-linear transaction patterns.

Though Random Forest performed better, the paper points out some issues that need to be addressed before large-scale implementation.

These are:

- Interpretability issues
- Adaptation for concept drift
- Real-time processing
- Regulatory issues

The paper makes a contribution to both the academic and practical communities in the following ways:

- It presents a systematic process for fraud detection.
 - It emphasizes the significance of choosing the right evaluation metric for imbalanced datasets.
- It compares interpretable linear models with ensemble models.
- It identifies some research gaps in AI-powered fraud detection.

Future Work

Future research should be directed at the following:

A. Integration of Explainable AI

The development of integrated explainability tools such as SHAP or LIME in ensemble-based fraud detection systems will improve regulatory transparency and trust.

B. Adaptation to Concept Drift

Online learning algorithms and retraining systems should be investigated to adapt to changing fraud patterns.

C. Graph Neural Networks

Graph-based fraud detection models that represent relational transaction graphs have great potential. Future research should investigate their scalability and real-time applicability.

D. Federated Learning for Privacy

Federated Learning frameworks may facilitate inter-institutional collaboration without compromising sensitive data, improving model robustness and preserving privacy.

E. Hybrid AI Systems

The combination of interpretable linear models and high-performance ensemble models may provide accuracy and transparency.

Conclusion

In summary, Artificial Intelligence has great transformative potential in fraud detection for online transactions. Nevertheless, its application in practice demands a careful trade-off between performance, interpretability, scalability, and regulatory requirements. Future interdisciplinary research is imperative to develop next-generation fraud detection systems that can keep pace with the dynamic digital financial environment.

9. References

- [1] A. Bhattacharyya, V. 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.
- [2] A. Dal Pozzolo, O. Caelen, R. A. Johnson, and G. Bontempi, "Calibrating probability with undersampling for unbalanced classification," in *IEEE Symposium Series on Computational Intelligence*, 2015.
- [3] F. Carcillo, Y. Le Borgne, O. Caelen, and G. Bontempi, "Machine learning for financial fraud detection: A survey," *IEEE Access*, vol. 9, pp. 1–20, 2021.
- [4] Y. Bengio et al., "Anomaly detection in financial transactions," *arXiv:1904.10604*, 2019.
- [5] H. He and E. Garcia, "Learning from imbalanced data," *arXiv:1709.08920*, 2017.
- [6] X. Zhang et al., "Graph-based fraud detection using relational learning," *arXiv:2409.13406*, 2024.
- [7] *Nature Scientific Reports*, "Hybrid artificial intelligence for financial anomaly detection," *Sci. Rep.*, vol. 15, 2025, Art. no. s41598-025-00873-y.
- [8] U. Fiore et al., "Using SMOTE for credit card fraud detection," *Information Sciences*, vol. 479, pp. 448–455, 2019.
- [9] C. Whitrow et al., "Transaction aggregation as a strategy for credit card fraud detection," *Data Mining and Knowledge Discovery*, vol. 18, no. 1, pp. 30–55, 2009.
- [10] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. ACM SIGKDD*, 2016.
- [11] S. Lundberg and S. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems*, 2017.
- [12] K. Bonawitz et al., "Federated learning for secure and privacy-preserving AI," *Proc. Machine Learning Systems*, 2019.