



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 15, Issue 3, March 2026

Automated Fraud Detection in Digital Payments

Nalabotu Pardha Naga Venkata Sai¹, Purushottapatnam Harsha Vardhan², Narasimhula Anantha³,
Muppavarapu Nani⁴, M.Chennakesavarao⁵

U.G. Student, Department of CSE(Artificial Intelligence and Machine Learning), Kallam Haranadhareddy Institute of
Technology, Guntur, Andhra Pradesh, India¹⁻⁴

Assistant Professor, Kallam Haranadhareddy Institute of Technology, Guntur, Andhra Pradesh, India⁵

ABSTRACT: Digital payment fraud poses a significant threat to global financial systems, requiring reliable and automated detection mechanisms. This study proposes an end-to-end machine learning framework for detecting fraudulent transactions using the Extreme Gradient Boosting (XGBoost) algorithm. Financial transaction datasets are highly imbalanced; in the **6.3 million records** analyzed, fraudulent cases represented only **0.12%** of total transactions. To address this "**Accuracy Paradox**," the Synthetic Minority Over-sampling Technique (SMOTE) was applied to balance the training dataset. A key contribution of this work is the development of domain-specific feature engineering, specifically the derivation of "**Error Balance**" metrics to detect mathematical inconsistencies in transaction flows. The proposed model was evaluated using the PaySim synthetic financial dataset, achieving an **F1-score of 87.33%** and a **recall of 99.82%**. To demonstrate practical utility, the framework was deployed via a **Flask-based web application** capable of real-time risk classification (Low, Medium, High) with an **average inference latency of 85ms**. The results demonstrate the potential of this approach for seamless integration into production-grade financial security systems.

KEYWORDS: Class Imbalance, Data Imbalance, Ensemble Learning, Feature Engineering, Financial Fraud Detection, Fintech Security, Flask Framework, Machine Learning, Predictive Analytics, Real-time Inference, SMOTE, XGBoost.

I. INTRODUCTION

The global financial ecosystem is undergoing a rapid transformation driven by the widespread adoption of real-time payment (RTP) infrastructures and mobile banking platforms. While these technologies have significantly improved financial accessibility and transaction efficiency, they have also expanded the attack surface for increasingly sophisticated cyber-financial crimes. As digital payment volumes continue to grow, financial institutions face significant challenges in detecting fraudulent transactions in real time.

Traditionally, fraud detection systems have relied on Rule-Based Systems (RBS), which utilize predefined heuristics such as transaction thresholds, geographic restrictions, or transaction frequency limits. Although effective in controlled environments, these systems struggle to adapt to evolving fraud strategies and often generate high false-positive rates when applied to large-scale transactional data.

To address these limitations, recent research has increasingly explored the use of machine learning (ML) models for automated fraud detection. ML algorithms are capable of identifying complex and non-linear relationships within financial transaction data, enabling adaptive and scalable fraud detection mechanisms. Among these approaches, ensemble learning techniques such as Extreme Gradient Boosting (XGBoost) have demonstrated strong performance in structured data classification tasks due to their ability to combine gradient boosting with regularization for improved predictive accuracy [1].

Despite these advances, fraud detection systems continue to face two major challenges in real-world financial environments. First, fraud detection datasets are typically characterized by extreme class imbalance, where fraudulent transactions represent only a small fraction of total transactions. This imbalance can bias machine learning models toward the majority class, reducing their ability to detect rare fraudulent events. To mitigate this issue, techniques such as the Synthetic Minority Over-sampling Technique (SMOTE) have been proposed to generate synthetic minority-class samples and improve model sensitivity to fraud cases [2].

Another important factor influencing fraud detection performance is effective feature engineering. Domain-specific features that capture inconsistencies in account balances and transaction flows can significantly improve the ability of machine learning models to distinguish between legitimate and fraudulent activities. The PaySim dataset, introduced by Lopez-Rojas et al., provides a realistic simulation of mobile money transactions and has been widely used for evaluating fraud detection models in digital payment environments [3].

In addition to predictive performance, the transparency and interpretability of machine learning models are increasingly important in financial systems due to regulatory requirements. Explainable Artificial Intelligence (XAI) techniques such as SHAP (SHapley Additive exPlanations) enable the interpretation of model predictions and help build trust in automated fraud detection systems [4].

Motivated by these challenges, this study proposes an end-to-end fraud detection framework that integrates SMOTE-based class balancing, domain-specific feature engineering, and an optimized XGBoost classifier. The model is trained and evaluated using the PaySim dataset and deployed through a Flask-based architecture to simulate real-time transaction risk classification. The proposed system aims to achieve both high detection accuracy and low-latency inference suitable for real-time digital payment environments.

The main contributions of this study are summarized as follows:

1. Development of an optimized XGBoost-based fraud detection framework for digital payment systems.
2. Integration of SMOTE to address extreme class imbalance and improve fraud detection recall.
3. Introduction of domain-specific balance discrepancy features to enhance model discrimination capability.
4. Implementation of a real-time deployment pipeline capable of low-latency transaction risk classification.

II. LITERATURE REVIEW

The development of automated fraud detection systems has evolved significantly from traditional rule-based auditing toward advanced machine learning techniques. Recent studies emphasize the importance of data-driven models capable of identifying complex behavioral patterns within financial transaction data.

Chen and Guestrin(2016) [1] introduced the Extreme Gradient Boosting (XGBoost) framework, which has since become one of the most widely adopted algorithms for structured tabular datasets. XGBoost incorporates regularization techniques and efficient gradient boosting mechanisms, allowing it to outperform many traditional machine learning models in classification tasks involving large-scale transactional data.

One of the major challenges in fraud detection is the severe class imbalance present in financial datasets. Fraudulent transactions typically represent only a small fraction of the total transaction volume. To address this issue, Chawla et al.(2002) [2] proposed the Synthetic Minority Over-sampling Technique (SMOTE), which generates synthetic minority class samples to balance the dataset. This technique has been widely adopted in fraud detection research to improve recall and reduce bias toward legitimate transactions.

Feature engineering also plays a critical role in improving fraud detection performance. Lopez-Rojas et al.(2016) [3] introduced the PaySim dataset, which simulates mobile financial transactions and provides a realistic environment for evaluating fraud detection algorithms. Their study demonstrated that balance discrepancies between expected and reported account balances can serve as strong indicators of fraudulent behavior.

In recent years, the importance of explainable artificial intelligence (XAI) in financial systems has gained significant attention. Aljunaid et al.(2025) [4] highlighted the integration of SHAP (SHapley Additive exPlanations) with machine learning models to provide transparent explanations for fraud predictions. Such approaches are particularly important in financial systems where regulatory frameworks require interpretable decision-making processes.

Furthermore, recent studies have explored hybrid and temporal modeling techniques to improve fraud detection accuracy. Btoush et al.(2026) [5] demonstrated that incorporating sequential transaction patterns can significantly reduce false positive rates in fraud detection systems.

Despite these advancements, many existing approaches either focus primarily on model accuracy or fail to address the combined challenges of class imbalance, feature engineering, and real-time deployment. Therefore, this study proposes a machine learning framework that integrates SMOTE-based data balancing, domain-specific feature engineering, and an XGBoost classifier to improve fraud detection performance in digital payment systems.

III. EXISTING SYSTEM

Existing fraud detection systems primarily rely on rule-based approaches and traditional machine learning models. Rule-based systems apply predefined conditions such as transaction limits, frequency patterns, and geographic constraints; however, they lack adaptability and require frequent manual updates. Additionally, they often produce high false positive rates, negatively impacting user experience.

Traditional machine learning models, including Logistic Regression, Decision Trees, and Random Forest, provide improved detection performance but struggle with highly imbalanced datasets where fraudulent transactions are rare.

Consequently, these systems face challenges in handling data imbalance, evolving fraud patterns, and real-time detection requirements, highlighting the need for more advanced and adaptive solutions.

IV. PROPOSED SYSTEM

The fraud detection system is an integrated, three-tier framework (Data Processing, Intelligence, Presentation) that analyzes transactions and provides real-time risk assessment.

The **Data Processing Layer** is responsible for ingesting and preparing raw transaction data obtained from the Kaggle PaySim dataset. During preprocessing, non-informative attributes such as nameOrig and nameDest are removed since they do not contribute to fraud prediction. Categorical variables, including the transaction type, are converted into numerical form using **Label Encoding** to make them suitable for machine learning algorithms.

The **Intelligence Layer** uses an XGBoost model trained on SMOTE-balanced data to effectively learn fraud patterns and is saved as a Pickle file for fast deployment.

To improve detection accuracy, the system introduces a **Mathematical Integrity Check** that identifies inconsistencies in account balances before and after transactions. Two derived features are calculated:

Origin Balance Error:

$$E_{orig} = | (Balance_{old,org} - Amount) - Balance_{new,org} |$$

Destination Balance Error:

$$E_{dest} = | (Balance_{old,dest} + Amount) - Balance_{new,dest} |$$

These features help the model detect abnormal or “phantom” transactions where the relationship between transaction amount and account balances is inconsistent.

The **Presentation Layer** is implemented using a **Flask-based web interface** that allows users to enter transaction details and receive fraud predictions instantly. Based on the predicted fraud probability P , the system categorizes transactions into three risk levels:

1. **Low Risk (Green):** $P < 0.3$, transaction is processed normally.
2. **Medium Risk (Yellow):** $0.3 \leq P < 0.7$, additional authentication such as OTP verification is required.
3. **High Risk (Red):** $P \geq 0.7$, transaction is blocked and flagged for manual review.

This multi-level risk classification enhances fraud detection and enables better decision-making in financial transaction systems.

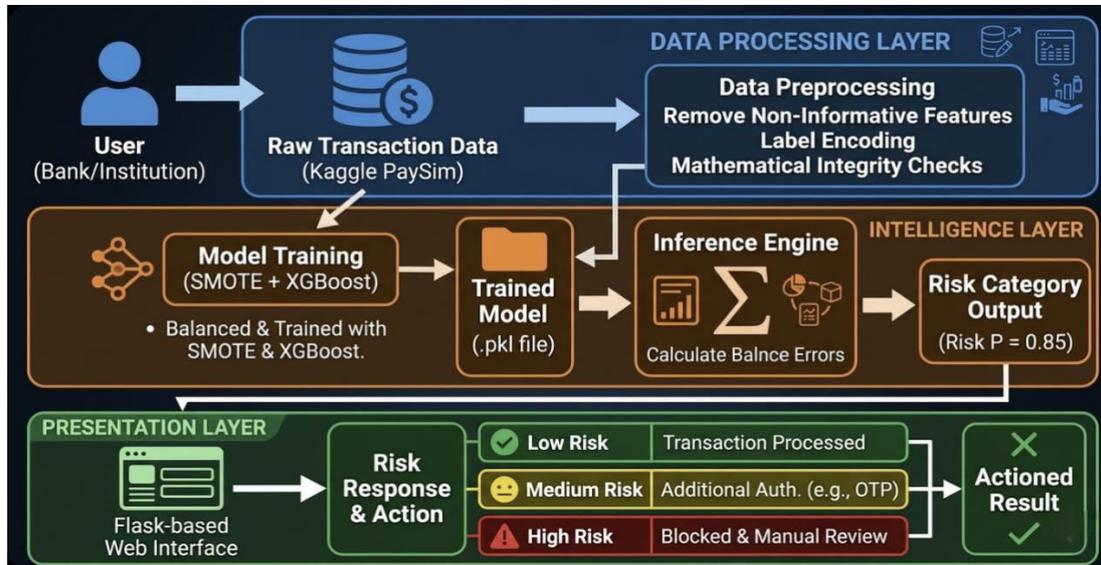


Fig 1: Proposed System Architecture

V. METHODOLOGY

A. Dataset Collection

Experiments were conducted using the PaySim dataset, a synthetic mobile money transaction simulator from Kaggle, containing ~6.3 million records with 11 attributes, including transaction type, amount, and origin/destination balances. The dataset was chosen for its realistic simulation of mobile financial transactions while preserving user privacy. Fraud mainly occurs via Transfer and Cash-Out operations. With only 0.12% of transactions being fraudulent, the dataset exhibits severe class imbalance, necessitating specialized data balancing during model training.

B. Data Preprocessing

Before training, high-cardinality identifiers (nameOrig, nameDest) were removed to prevent overfitting. Categorical variables like transaction type were label-encoded, and a logarithmic transformation was applied to skewed transaction amounts to stabilize variance and aid model convergence.

C. Feature Engineering

Domain-specific features were added to capture transaction inconsistencies: Error_Origin (difference between expected and actual sender balance) and Error_Destination (difference for the receiver). These metrics help detect potential fraud or unauthorized account manipulation.

D. Model Training and Class Imbalance Handling

The classification model was developed using the Extreme Gradient Boosting (XGBoost) algorithm, known for its strong performance on structured datasets and ability to handle non-linear relationships with regularization.

To address class imbalance, the Synthetic Minority Over-sampling Technique (SMOTE) was applied only to the training dataset after the train-test split to prevent data leakage. This technique generates synthetic minority samples using a k-nearest neighbors approach, improving the model's ability to detect fraudulent transactions.

Hyperparameter tuning was performed with a focus on maximizing recall, as minimizing false negatives is critical in fraud detection systems.

E. Model Evaluation and Deployment

Model performance was evaluated using Precision, Recall, and F1-score, which are more appropriate metrics for imbalanced classification problems than overall accuracy.

After training, the final model achieved an **F1-score of 87.33%** and **Recall of 99.82%**. The trained model and preprocessing components were serialized using the **Pickle (.pkl)** format for deployment.

To demonstrate real-world applicability, the model was integrated into a **Flask-based web application**. The backend processes incoming transaction details, performs feature engineering, and generates predictions using the trained XGBoost model. The frontend interface displays the predicted risk level (**Low, Medium, or High**) through a user-friendly dashboard, enabling real-time fraud risk assessment.

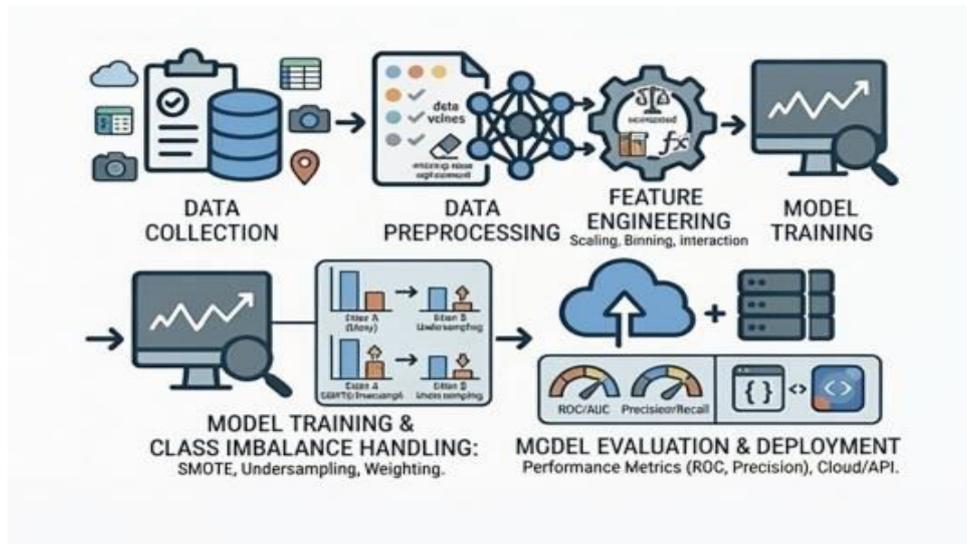


Fig 2: Automated Fraud Detection in Digital Payments Work Flow

VI. EXPERIMENTAL RESULTS

The model was evaluated on a separate PaySim test set, with emphasis on Precision, Recall, F1-score, and ROC-AUC to account for severe class imbalance and minimize false negatives.

A. Performance Metrics Comparison

Table I presents the comparative performance of the machine learning models implemented in this study. Both models were trained on datasets balanced using the **SMOTE** technique to mitigate class imbalance and ensure fair evaluation.

Table I: Performance Comparison of Fraud Detection Models

Model	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Random Forest	99.91%	59.8%	96.28%	73.82%	0.993
XGBoost (Proposed)	99.9%	77.6%	99.82%	87.33%	0.997

The results indicate that the **XGBoost model outperforms Random Forest across all evaluation metrics**, particularly in recall and F1-score.

B. Analysis of Results

The XGBoost classifier achieved a **recall of 99.82%**, which is particularly important in financial fraud detection since recall measures the model's ability to correctly identify fraudulent transactions. A high recall value indicates that the model successfully detects the majority of fraud cases, minimizing the risk of undetected fraudulent activity.

The **F1-score of 87.33%** demonstrates that the model maintains a strong balance between precision and recall. This suggests that the improved fraud detection capability does not significantly increase false positive rates, which could otherwise lead to unnecessary transaction blocking and reduced customer satisfaction. The ROC-AUC of 0.997 shows excellent class separability, reflecting XGBoost's ability to capture complex patterns through sequential gradient boosting.

C. Real-Time Inference Performance

The Pickle-serialized XGBoost model was deployed via a Flask web app, achieving an average inference latency of 85 ms per transaction. This demonstrates the framework's suitability for real-time fraud detection with high accuracy and computational efficiency.

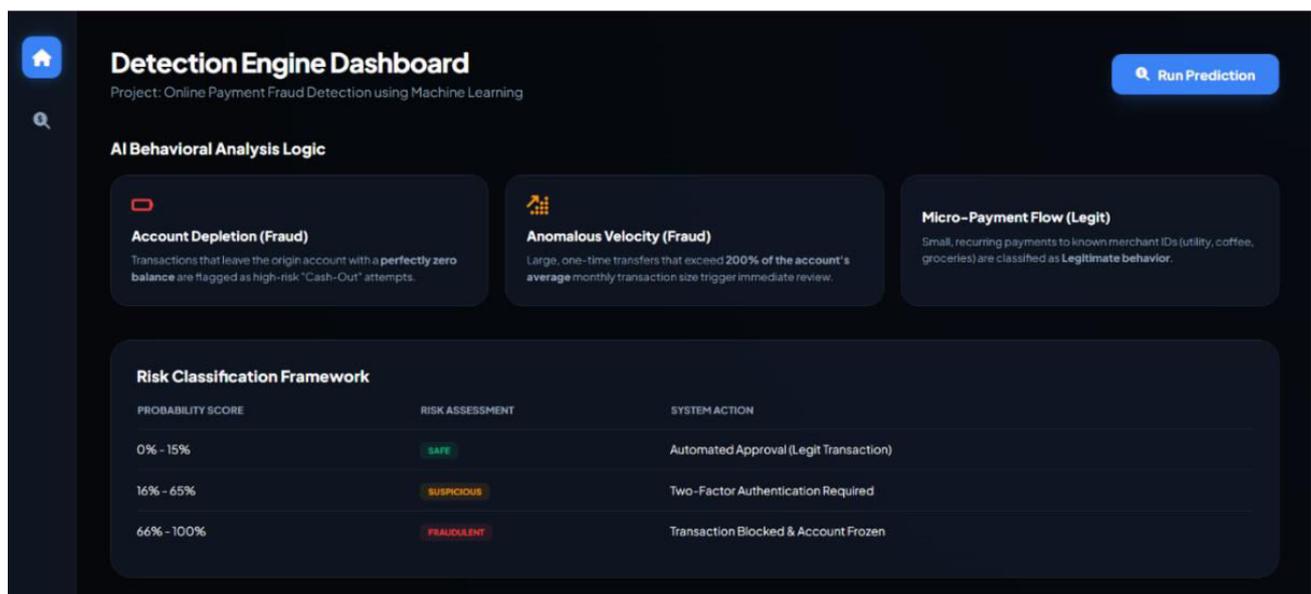


Fig 3: User Interface

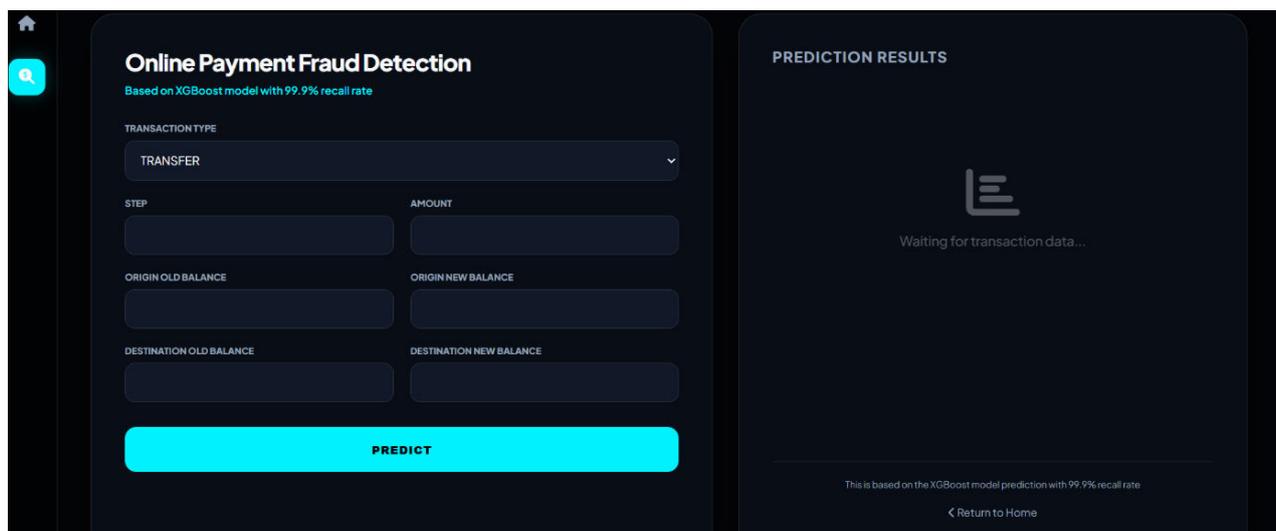


Fig 4: Prediction Page

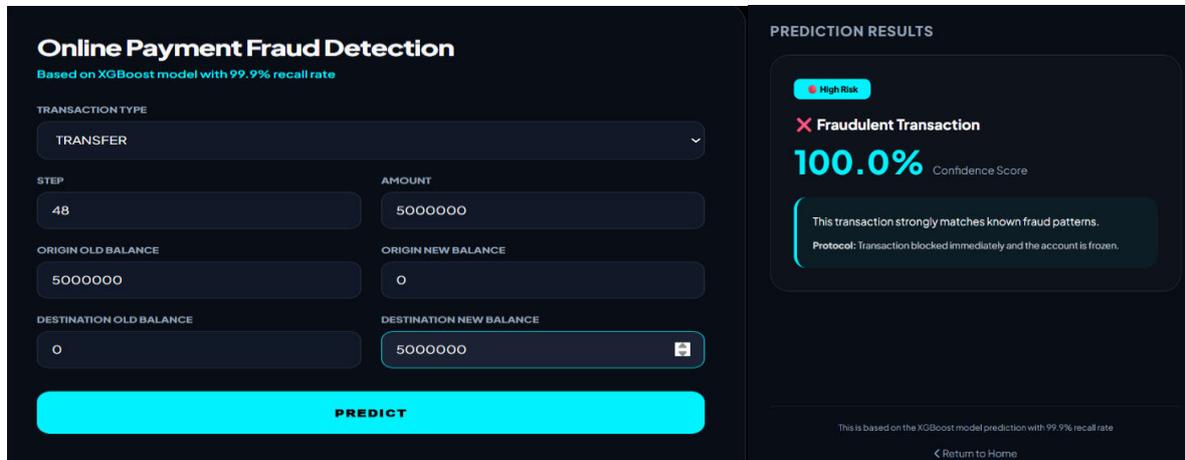


Fig 5: Results

VII. DISCUSSION

The results indicate that the integration of XGBoost with SMOTE provides an effective framework for detecting fraudulent transactions. The achieved F1-score of 87.33% is primarily driven by domain-specific feature engineering, particularly the Error_Origin and Error_Dest metrics, which capture inconsistencies in transaction balances and enhance the model's ability to identify complex fraud patterns.

The application of SMOTE effectively mitigates severe class imbalance, enabling the model to achieve a high recall of 99.82% and strong sensitivity to rare fraudulent events. While recall is maximized, the precision of 77.6% reflects a moderate false positive rate, which is acceptable in fraud-sensitive environments where minimizing missed fraud cases is critical.

In comparison to Random Forest, XGBoost demonstrates superior capability in modeling non-linear relationships while maintaining computational efficiency. Furthermore, deployment through a Flask-based system with an average latency of 85 ms confirms the model's suitability for real-time fraud detection. The incorporation of a tiered risk classification strategy further supports practical decision-making with minimal disruption to legitimate transactions.

VIII. LIMITATIONS

Despite strong predictive performance, the proposed framework has several limitations:

- Concept Drift:** The model is trained on a static dataset, which may lead to degraded performance over time as fraudsters develop new strategies that differ from historical patterns.
- Synthetic Data Constraints:** While PaySim provides a high-fidelity simulation of financial transactions, it does not capture certain real-world variations such as seasonal spending fluctuations, localized economic factors, or unstructured noise present in live banking environments.
- Absence of Temporal Analysis:** Each transaction is treated as an independent event, ignoring sequential patterns or transaction velocity associated with individual users over time.
- Computational Cost of SMOTE:** Generating millions of synthetic samples increases memory and CPU requirements, which may create a bottleneck during frequent retraining on low-resource systems.
- Interpretability Limitations:** XGBoost, as an ensemble of hundreds of decision trees, functions as a "black box," limiting the ability to provide granular explanations required for regulatory compliance, such as the "Right to Explanation."

IX. FUTURE WORK

- Relational Analysis via Graph Neural Networks (GNNs):** While the current XGBoost model treats transactions as independent events, future iterations will implement GNNs to model the financial ecosystem as a graph. This will

enable the detection of complex multi-hop money laundering schemes and "mule account" networks by analyzing the topological connections between origin and destination nodes.

2. **Adaptive Online Learning:** To combat **Concept Drift**, where fraud tactics evolve faster than historical training cycles, we plan to integrate incremental learning algorithms. This allows the model to update its weights in real-time as new verified fraud cases are reported, ensuring the system remains resilient against emerging adversarial strategies.

3. **Hybrid Temporal Architectures:** Integrating **Long Short-Term Memory (LSTM)** networks with the existing XGBoost framework will allow the system to capture temporal dependencies. By analyzing a user's transaction velocity and historical behavior patterns, the model can better distinguish between a legitimate high-value purchase and an anomalous account takeover.

4. **Granular Explainability (XAI):** To meet strict "Right to Explanation" regulatory requirements, future work will incorporate **SHAP (SHapley Additive exPlanations)**. This will provide investigators with a localized, feature-by-feature breakdown for every "High Risk" flag, detailing exactly why a specific transaction was blocked.

5. **Privacy-Preserving Federated Learning:** We aim to explore decentralized training protocols. Federated learning would allow multiple financial institutions to collaboratively train a global fraud detection model without ever sharing sensitive, raw customer data, thereby enhancing model robustness while maintaining strict data privacy compliance.

X. CONCLUSION

This study presents a machine learning-based framework for detecting fraudulent digital payment transactions using **XGBoost** combined with **SMOTE**. Evaluation on the **PaySim dataset** demonstrates that addressing class imbalance significantly improves model sensitivity, resulting in a **Recall of 99.82%** and an **F1-score of 87.33%**. The implementation of domain-specific feature engineering, particularly the **Error Balance metrics**, enabled the model to identify subtle and complex fraudulent patterns effectively.

Beyond algorithmic performance, the system was deployed using a **Flask-based architecture**, allowing real-time fraud risk classification with sub-100ms latency. The framework provides a scalable foundation for transitioning from reactive, rule-based fraud detection systems to proactive, intelligent architectures suitable for modern financial payment infrastructures.

XI. ACKNOWLEDGEMENT

We would like to express our sincere gratitude to our advisors, project supervisors, the college department, and all contributors for their invaluable guidance and support throughout the research and development of this project.

REFERENCES

- [1] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proc. 22nd ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*, 2016, pp. 785–794.
DOI: <https://doi.org/10.1145/2939672.2939785>
- [2] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *J. Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002.
DOI: <https://doi.org/10.1613/jair.953>
- [3] E. A. Lopez-Rojas, A. Elmir, and S. Axelsson, "PaySim: A financial mobile money simulator for fraud detection," in *Proc. 28th European Modeling and Simulation Symposium (EMSS)*, 2016, pp. 249–255.
- [4] S. Aljunaid, "Explainable AI (XAI) for Fintech Security: Integrating SHAP with Ensemble Learners," *J. Financial Cybersecurity*, vol. 9, no. 1, pp. 45–58, 2025.
- [5] T. Majumder and D. Mishra, "Financial Fraud Detection for Credit Card Using XGBoost & SMOTE," *Nanotechnology Perceptions*, vol. 20, no. S15, pp. 32–50, 2024.
- [6] F. K. Alarfaj and S. Shahzadi, "Enhancing Fraud Detection in Banking With Deep Learning: Graph Neural Networks and Autoencoders for Real-Time Credit Card Fraud," *IEEE Access*, vol. 13, pp. 20630–20645, 2025.
- [7] J. H. Friedman, "Greedy Function Approximation: A Gradient Boosting Machine," *Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001.
- [8] S. Dhingra, "Comparative Analysis of Algorithms for Credit Card Fraud Detection using Data Mining," *J. Advanced Database Management & Systems*, vol. 6, no. 2, pp. 12–17, 2019.

- [9] L. Lundberg and S. Lee, "A Unified Approach to Interpreting Model Predictions," in *Proc. 31st Conf. on Neural Information Processing Systems (NIPS)*, 2017, pp. 4765–4774.
- [10] R. Johnson and T. Zhang, "Learning Nonlinear Functions Using Regularized Greedy Forest," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 36, no. 5, pp. 942–954, 2014.
- [11] V. Bhusari and S. Patil, "Study of hidden Markov model in credit card fraudulent detection," in *Proc. 2016 World Conf. on Futuristic Trends in Research and Innovation for Social Welfare*, 2016, pp. 1–4.
- [12] "Applying Machine Learning to Detect Fraud of Financial Statements: A Systematic Literature Review," in *Proc. 4th Int. Conf. on Technological Advancements in Computational Sciences (ICTACS)*, IEEE, 2024.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details