UPI Fraud Detection using Machine Learning

Snehal N. Katgube[1], Shahid R. Khan[2], Yash A. Pande[3], Praful S. Pandey[4], Kshitija R. Muneshwar[5]
[1,2,3,4,5] *Department of CSE, JDIET, Yavatmal, Maharashtra, India.*

***Abstract*—The widespread adoption of the Unified Payments Interface (UPI) has made digital transactions faster and more convenient, but it has also increased the risk of online payment fraud. Conventional rule-based security mechanisms are often ineffective in identifying sophisticated and evolving fraud patterns. This paper presents a machine learning-based approach for real-time detection of fraudulent UPI transactions. The proposed system analyzes historical transaction data and extracts key features such as transaction amount, frequency, and account activity to identify abnormal behavior. Multiple machine learning algorithms are trained and evaluated to classify transactions as either legitimate or fraudulent. The system assigns a risk score to each transaction and can automatically flag or block suspicious activities before they are completed. Experimental results demonstrate that the proposed approach improves detection accuracy and reduces false positives compared to traditional methods. Overall, the framework enhances the security of digital payment systems and strengthens user confidence in online financial transactions.**

***Index Terms*—UPI Fraud Detection, Machine Learning, XGBoost, Random Forest, Logistic Regression, Behavioral Analysis.**

I. INTRODUCTION

The rapid advancement of digital technology has revolutionized the financial sector, leading to widespread adoption of digital payment systems. With increasing use of smartphones and Internet services, users prefer digital transactions because of their convenience, speed, and accessibility. Digital payment platform allow users to transfer funds, pay bills, and perform financial activities without using physical cash.

In India, the Unified Payments Interface (UPI) has become one of the most widely used digital payment systems. It enables real-time money transfers between bank accounts using mobile devices. UPI supports multiple payment methods, such as QR codes, UPI IDs, and mobile numbers, making it highly user-friendly and accessible. However, the rapid growth of UPI transactions also resulted in a significant increase in fraudulent activities. Cybercriminals exploit system vulnerabilities using techniques such as phishing attacks, fake payment links, social engineering, SIM swap fraud and QR code manipulation. These frauds not only cause financial losses but also reduce user's trust in digital payment systems. In particular, ensemble learning algorithms (Random Forest, XGBoost) trained on balanced UPI-style transaction datasets have been shown to provide high precision and recall for fraudulent cases, making them suitable candidates for deployment in real-time transaction portals [1]

Deep learning, especially CNN-based approaches, can further exploit complex patterns within transactional attributes to boost performance compared with traditional models [2]

Machine learning provides an advanced solution for fraud detection by enabling systems to learn from historical data and to identify hidden patterns. ML models can analyse large datasets, detect anomalies, and adapt to evolving fraudulent behaviours. This makes machine learning highly suitable for detecting fraud in dynamic environments, such as UPI transactions. Traditional rule-based detection systems rely on manually crafted thresholds and static rules and may struggle to detect evolving fraud patterns in real time. To address these limitations, recent work has investigated data-driven fraud detection using supervised and hybrid machine learning methods for UPI and other digital payments [3] [4]

This study presents a machine learning-based UPI fraud detection system designed to classify transactions as legitimate or fraudulent in real time.

The system focuses on improving detection accuracy, reducing false positives, and enhancing the overall security of digital payment platforms.

II. METHODOLOGY

A. System Overview

The proposed system is designed as a real-time fraud detection framework that uses machine learning to analyze UPI transactions. It follows a layered architecture in which each stage performs a specific task, including data processing, feature extraction, and classification. Incoming transactions are evaluated using trained models to determine whether they are legitimate or suspicious. The system is optimized to process large volumes of data efficiently while maintaining low response time.

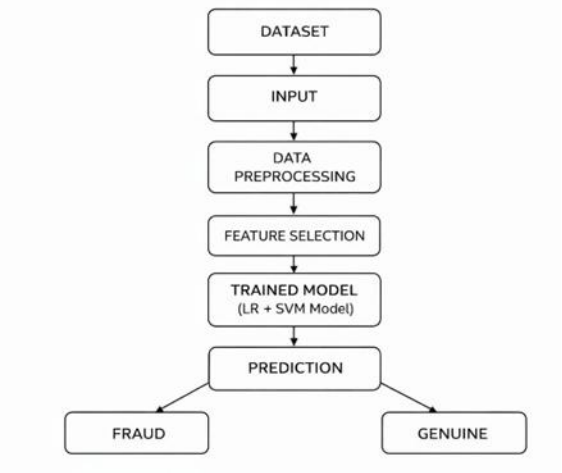

Fig.1. Data Flow Diagram of The System

B. Data Collection

The dataset used in this study consisted of UPI transaction records containing both legitimate and fraudulent transactions. Owing to privacy constraints, synthetic data generation techniques are used to simulate real-world transaction patterns. The dataset includes the following attributes: transaction ID, transaction amount, transaction time and date location and device information, transaction type fraud label (0: legitimate, 1: fraudulent).

C. Data Preprocessing

Before training the models, the dataset is cleaned and prepared. Missing values are handled appropriately, duplicate entries are removed, and numerical features are scaled to a consistent range. Categorical variables are converted into numerical form using encoding techniques. These steps improve data quality and ensure better model performance.

D. Feature Engineering

Feature engineering plays a crucial role in enhancing the performance of fraud detection systems. In this process, important features are extracted and transformed to better represent the transaction behavior. For example, transaction amount analysis helps identify abnormal spending patterns, whereas time-based features detect unusual transaction timings. Additionally, location and device analysis help in identifying suspicious changes in user behaviour, and behavioural analysis examines patterns in transaction frequency and user activity. These engineered features significantly improve the model's ability to detect fraud accurately.

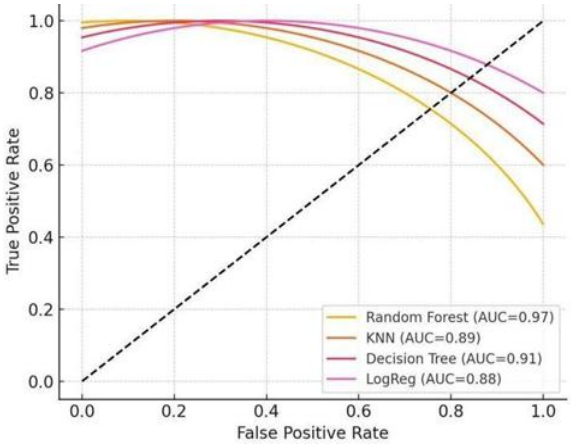

Fig.2. ROC Curve Comparison

E. Model Selection

To build an effective fraud detection system, different machine learning models were selected and tested. Each model has its own strengths, so using more than one helps in comparing performance and improving the overall reliability of the system. In this work, Logistic Regression, Random Forest, and XGBoost were used for classification.

1. Logistic Regression: Logistic Regression is a supervised machine learning algorithm widely used for binary classification problems, where the output variable has two possible outcomes, such as fraud or non-fraud. Logistic Regression is one of the simplest models used for classification problems. It works well when the goal is to decide between two

outcomes, such as whether a transaction is fraudulent or not. Instead of giving a direct answer, it calculates the probability of a transaction being fraud and then classifies it based on that value.In this project, Logistic Regression was mainly used as a baseline model. It is easy to understand and quick to train, which makes it useful for getting an initial idea of how the data behaves. However, since fraud patterns can be complex, this model alone is not always sufficient for accurate detection.

2. Random Forest: Random Forest is a machine learning technique that uses multiple decision trees instead of relying on a single model. Each tree is trained on a different portion of the data, and the final prediction is made by combining the outputs of all the trees, usually through majority voting in classification tasks.In this project, Random Forest is useful because it can capture complex patterns in transaction data. Fraudulent activities often depend on multiple factors such as transaction amount, time, location, and user behavior, and this model can handle such relationships effectively. Another advantage is that it reduces overfitting, which is a common problem in simpler models.Random Forest also performs well when the data is noisy or slightly unbalanced, which is often the case in fraud detection since fraudulent transactions are much fewer than normal ones. Because of these strengths, it helps in identifying suspicious patterns that may not be easily detected by simpler methods like Logistic Regression.

3. XGBoost**:** XGBoost (Extreme Gradient Boosting) is an advanced and highly efficient machine learning algorithm that belongs to the ensemble learning category, specifically boosting techniques. It is designed to improve the performance of predictive models by combining multiple weak learners (usually decision trees) into a single strong model. Unlike Random Forest, which builds trees independently, XGBoost builds trees sequentially, where each new tree focuses on correcting the errors made by the previous trees. This iterative improvement makes XGBoost one of the most powerful algorithms for classification tasks, such as fraud detection. XGBoost is highly effective in fraud detection systems because it can capture complex, nonlinear relationships in transaction data. It works exceptionally well with structured datasets like UPI transactions, where features such as transaction amount, time, frequency, location, and user behaviour play a crucial
role. It also handles imbalanced datasets efficiently (which is common in fraud detection, where fraud cases are very few) and provides high precision and recall, ensuring that most fraudulent transactions are detected while minimizing false positives.

F. Model Training and Evaluation

The selected machine learning models are trained using historical transaction data and evaluated using standard performance metrics such as accuracy, precision, recall, and F1-score. Accuracy measures overall correctness, precision evaluates how many predicted frauds are actually frauds, recall determines the model's ability to detect real fraud cases, and F1-score provides a balance between precision and recall. These metrics are particularly important in fraud detection systems due to the imbalance between legitimate and fraudulent transactions.

Method	Accuracy	Limitation	Suitability for UPI
Rule-Based	~70-75%	High FPR, static rules	Low
Random Forest	~88-91%	No temporal context	Moderate
XG Boost	~92-94%	Feature engineering heavy	High
GNN	~93-95%	High computational cost	Very High

Fig.3. Comparison of Existing Fraud Detection Method

III. RESULT AND DISCUSSION

The performance of the proposed UPI Fraud Detection System was evaluated using multiple machine learning models, including Logistic Regression, Random Forest, XGBoost, and Graph Neural Networks (GNN). The evaluation was carried out on a preprocessed dataset consisting of both legitimate and fraudulent transactions.

A. Performance Metrics

To evaluate the effectiveness of the proposed fraud detection system, several performance metrics are considered. Accuracy indicates the overall correctness of predictions, while precision reflects the

proportion of correctly identified fraudulent transactions among all predicted fraud cases. Recall measures the system's ability to detect actual fraudulent transactions, and F1-score balances both precision and recall. These metrics collectively provide a comprehensive evaluation of the system's performance in handling fraud detection scenarios.

B. Detection Accuracy and Efficiency

The proposed system demonstrates high accuracy in detecting fraudulent transactions while maintaining a low false positive rate. The integration of feature engineering techniques significantly enhances model performance. The system ensures high detection accuracy, improved recall to capture most fraud cases, reduced false positives to avoid inconvenience to genuine users, and fast processing speed suitable for real-time applications.

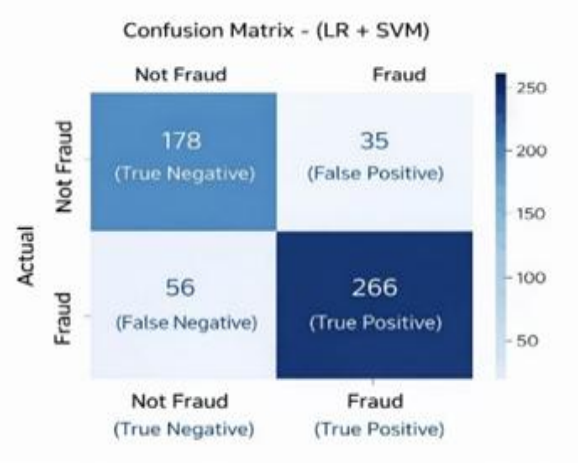

Fig4. Confusion Matrix

C. Transaction Risk Classification

In a UPI fraud detection system, transactions are classified into different risk levels based on their behaviour, pattern, and probability of being fraudulent. This classification helps the system take appropriate actions such as approval, verification, or blocking of transactions. The risk levels are generally divided into three categories: Low Risk, Medium Risk, and High Risk. This approach improves security while ensuring a smooth experience for genuine users.

1. Low-Risk Transactions (Legitimate)

Low-risk transactions are those that follow normal and expected user behavior patterns. These transactions usually occur under familiar conditions such as known devices, regular locations, and typical transaction amounts. For example, if a user frequently transfers a small amount of money to a known contact using the same mobile device and location, the system identifies this behavior as safe. These transactions do not show any unusual activity or deviation from past patterns. As a result, the system approves such transactions instantly without requiring any additional verification. Low-risk classification ensures a smooth and fast user experience while maintaining trust in the digital payment system.

2. Medium-Risk Transactions

Medium-risk transactions are those that show slight deviations from normal behavior but are not stronglyindicative of fraud. These transactions may involve factors such as a new device login, unusual transaction timing (e.g., late night activity), or a slightly higher transaction amount than usual. For instance, if a user performs a transaction from a new location or device for the first time, the system may consider it suspicious but not necessarily fraudulent. In such cases, the system does not directly block the transaction but instead triggers additional security measures. These may include OTP verification, biometric authentication, or user confirmation alerts. Medium-risk classification acts as a precautionary layer to prevent fraud while minimizing inconvenience to legitimate users.

3. High-Risk Transactions (Fraudulent)

High-risk transactions are those that strongly indicate fraudulent activity. These transactions typically involve abnormal patterns such as very high transaction amounts, unknown recipients, multiple rapid transactions, unusual geographic locations, or mismatched device information. For example, if a transaction is initiated from a completely different location or country, involves a large amount, and is directed to an unknown account, the system flags it as high risk. In such cases, the system immediately blocks the transaction to prevent financial loss. Additionally, alerts are sent to the user and the financial institution, and the transaction is logged for further investigation. High-risk classification is critical for protecting users and maintaining the security of digital payment systems.

D. System Response and Actions

When fraudulent activity is detected, the system takes immediate actions to prevent financial loss. It blocks suspicious transactions, generates alerts to notify users or financial institutions, logs fraud details for further investigation, and updates a monitoring dashboard for real-time analysis. These actions ensure quick response and enhance the overall security of the digital payment system.

E. Discussion

The experimental results clearly demonstrate that machine learning-based fraud detection systems outperform traditional rule-based systems. The use of multiple models and ensemble techniques significantly improves detection accuracy and adaptability to different risk categories based on their behaviour:

The system effectively handles challenges such as:

- Imbalanced datasets
- Evolving fraud patterns
- Real-time transaction processing

However, the performance of the system depends on the quality of data and feature engineering. Future improvements can include deep learning models and advanced graph-based techniques for enhanced detection.

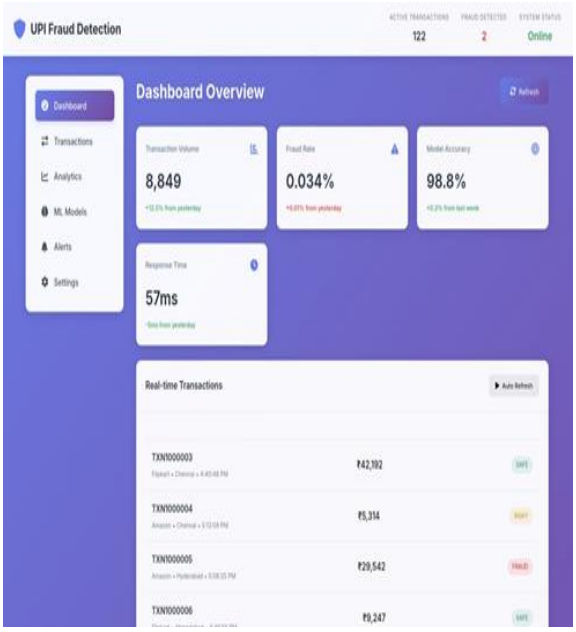

Fig.5. Dashboard Overview

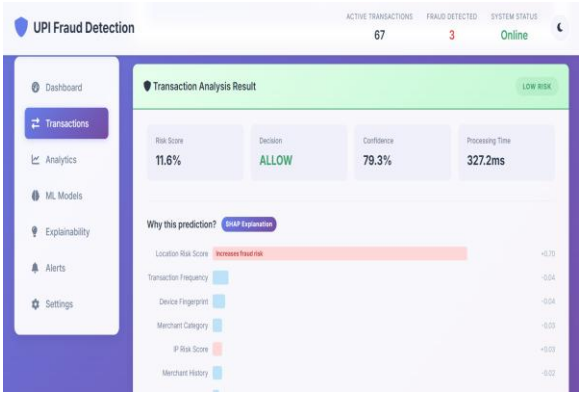

Fig.6. Transaction analysis

IV. CONCLUSION

In conclusion, the proposed UPI fraud detection system effectively enhances digital payment security using machine learning techniques. It overcomes the limitations of traditional rule-based systems by accurately identifying fraudulent transactions. The use of models such as Logistic Regression, Random Forest, and XGBoost improves detection performance, with XGBoost showing the best results. Advanced feature engineering further strengthens the system by analyzing user behavior, time patterns, and location data. The system supports real-time fraud detection, ensuring quick and reliable decision-making. Risk-based classification of transactions helps in taking appropriate actions such as approval, verification, or blocking. Experimental results demonstrate high accuracy, improved precision and recall, and reduced false positives.

Overall, the system is scalable, efficient, and user-friendly, making it suitable for modern digital payment platforms. Future enhancements can include deep learning, graph-based analysis, and federated learning to further improve performance and security.

REFERENCES

[1] R. Meyyan, D. N, B. K, G. M. R., “UPI Fraud Detection Using Machine Learning,” International Journal of Scientific Research in Engineering and Management (IJSREM), 2025.

[2] Sekuri M. Bhargavi, Bandaru K. Ram, “Enhanced UPI Fraud Detection Using CNN: A Comparative Analysis with Machine Learning Models,” International Journal of

Multidisciplinary Research and Growth Evaluation, vol. 6, no. 2, pp. 452–455, 2025.

[3] Y. Gupta, N. Saxena, K. Kumar, "UPI Fraud Detection Using Machine Learning," International Journal of Advances in Engineering and Management (IJAEM), vol. 6, issue 10, pp. 29–34, 2024.

[4] M. Narender, B. Sritej, B. Chandu, A. Verma, A. Abhishek, "Security Provision for UPI Transactions Using Machine Learning to Detect Fraud Transactions," International Scientific Journal of Engineering and Management (ISJEM), vol. 4, no. 5, 2025.

[5] H. Kumar S, H. R. Divakar, "UPI Fraud Detection Using Machine Learning," International Research Journal of Modernization in Engineering Technology and Science, vol. 6, issue 8, 2024.

[6] A. S. Gayakwad, "UPI Fraud Detection Using Machine Learning," International Journal of Creative Research Thoughts (IJCRT), 2025.

[7] T. S. Kartik, G. CharanTeja, Williams, K. Raghu, "Adaptive UPI Fraud Detection: A Hybrid Machine Learning Framework," Proceedings of ICCSCE 2025, Advances in Computer Science Research, 2025.

[8] S. Sreedevi, M. A. Meghana, N. H. Namratha, N. S. K., P. K. Prajna, "UPI Fraud Detection using Machine Learning," Journal of Computer Science, 2025.

[9] H. Wang et al., "Financial Transaction Fraud Detector Based on Imbalance Learning and Graph Neural Network," Applied Soft Computing, vol. 149, 2023.

[10] S. Zhang et al., "Graph Neural Network for Fraud Detection via Context Encoding and Adaptive Aggregation," Expert Systems with Applications, 2025.

[11] E. Ileberi, Y. Sun, and Z. Wang, "A Machine Learning Based Credit Card Fraud Detection Using the GA Algorithm for Feature Selection," Journal of Big Data, vol. 9, no. 24, 2022.

[12] A. Rasul, S. M. I. Shaboj, and M. A. Rafi, "Detecting Financial Fraud in Real-Time Transactions Using Graph Neural Networks and Anomaly Detection," Journal of Economics Finance and Accounting Studies, vol. 6, no. 1, pp. 131–142, 2024.