International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 12 | Dec - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

UPI Fraud Detection Using Machine Learning

Dr. Raja Meyyan[1], Deekshitha N[2], Bindu K[3], Guru Murthy R[4]

[1]*Professor & HOD, Dept of ISE, East West Institute of Technology, Bengaluru*

[2,3,4] *Student, Dept of ISE, East West Institute of Technology, Bengaluru*

Abstract – The project "UPI Fraud Detection Using Machine Learning" aims to provide an intelligent, real-time security layer for UPI-based digital payments by automatically identifying suspicious transactions before they are completed. An ensemble of machine learning models (including Random Forest, XG Boost, Light GBM and Gradient Boosting) is trained on a balanced fraud-non-fraud dataset, with standardized features and careful handling of class imbalance to improve recall on rare fraudulent cases while maintaining high precision. The system is deployed as a Flask-based UPI transaction portal, where users can register, log in, initiate payments, and receive instant feedback on the fraud risk for each transaction. For every payment request, the model outputs a fraud probability; high-risk transactions are automatically blocked, logged into the user's history, and accompanied by email alerts and downloadable CSV/PDF reports for audit and analysis.

Key Words: Blockchain Technology, Cybersecurity, Web Development, Machine Learning.

1.INTRODUCTION

The project "UPI Fraud Detection Using Machine Learning" addresses this need by designing and implementing a predictive model that can distinguish between legitimate and fraudulent UPI-style transactions based on historical data. The system preprocesses transaction records, engineers' relevant features such as amount behavior, balance changes, and transaction frequency, and applies mutual-information-based feature selection to retain the most informative attributes for classification. Multiple machine learning algorithms-including ensemble models like Random Forest, XG Boost, Light GBM, and Gradient Boosting—are trained and evaluated to handle the highly imbalanced nature of fraud data, with a focus on maximizing recall for fraudulent cases while maintaining acceptable precision and overall accuracy. Users can register, log in, and initiate payments, while the backend model evaluates each transaction in real time and predicts whether it is safe or potentially fraudulent. Suspicious transactions are blocked and logged, with options to generate downloadable reports and trigger email alerts to the user. Additionally, SHAP-based explainable AI techniques are used to visualize and interpret the most influential features behind the model's decisions, helping both end users and administrators understand why a transaction was flagged. This combination of robust machine learning, real-time deployment, and interpretability provides a comprehensive framework for enhancing security and trust in UPI-based digital payments.

1.1 PROBLEM STATEMENT

The problem statement for "Fraud Detection in UPI Transactions Using Ensemble Learning" centers on the critical need to overcome the inefficiencies and inaccuracies of traditional rule-based fraud detection systems in the financial technology sector. UPI Transactions, which are a cornerstone of the digital economy, require immediate and precise fraud analysis to protect user assets and ensure system reliability. However, current processes, which often rely on static thresholds and predefined rules, are not only reactive but also incapable of adapting to new fraud tactics, leading to undetected breaches and false alarms. These shortcomings result in direct financial losses, eroded customer trust, and increased operational costs for the financial institutions. Therefore, there is a pressing need for an intelligent, ensemble-based machine learning system that can learn from complex, imbalanced data, provide real-time and explainable predictions, and dynamically adapt to the evolving landscape of financial cybercrime.

1.2 KEY OBJECTIVES

This study aims to achieve the following objectives:

To design a robust machine learning model that can accurately classify UPI-style transactions as legitimate or fraudulent using historical transaction data.

To perform effective data preprocessing and feature engineering, including handling missing values, encoding categorical attributes, computing derived features, and addressing severe class imbalance between normal and fraudulent transactions.

To evaluate and compare multiple ML algorithms (such as Random Forest, XG Boost, Light GBM, Gradient Boosting, etc.) using metrics like accuracy, precision, recall, F1-score and ROC-AUC and to select the best model for deployment with a focus on high recall for fraud cases.

To develop a user-friendly web interface (UPI transaction portal) using Flask, where users can register, log in, initiate transactions, and obtain instant feedback on whether a transaction is safe or suspicious.

To implement real-time fraud prediction by integrating the trained ML model with the transaction flow so that high-risk UPI transactions can be blocked or flagged before they are completed.

To provide transaction history and reporting features, enabling users and administrators to view past transactions, filter suspicious transactions, and download detailed reports for audit and analysis.

2. RELATED WORK AND LITERATURE SURVEY

2.1 UPI Fraud Detection Using Machine Learning

This study focuses specifically on UPI transactions and proposes a fraud detection system that combines traditional machine learning and deep learning to handle large-scale digital payment data. The authors highlight how UPI's explosive growth in India has led to sophisticated fraud attempts and argue that static rule engines cannot cope with evolving patterns. Their proposed system uses supervised models such as Random Forest and XG Boost, along with LSTM-based deep networks and anomaly detection, to flag suspicious transactions in real time. The study emphasizes the importance of capturing temporal behavior, transaction context, and user habits and reports that the hybrid design improves accuracy and robustness over single-model baselines.

2.2 Unified Payment Interface Fraud Detection Using Machine Learning

This study provides an application-oriented overview of how ML can be integrated into UPI fraud detection workflows. It starts with a review of existing ML-based financial fraud studies, then narrows down to UPI, and discusses dataset imbalance, feature engineering, and model evaluation. The proposed methodology uses supervised learning to classify transactions as fraudulent or legitimate, focusing on algorithms such as Random Forest and SVM. This study stresses the handling of class imbalance and explains how ML models can classify, predict patterns, and work with highly skewed datasets.

2.3 UPI Fraud Detection Using Machine Learning

This study is similar to an implementation-centric project report. It begins with a discussion of UPI growth and common attack vectors such as fake ID's, phishing links, and unauthorized access. The authors built a machine-learning model that analyses transaction attributes such as amount, time, device details, and user frequency to predict whether a transaction is safe or fraudulent. The final system includes analytics to show the distribution of detected frauds and model performance, positioning the solution as a smarter alternative to rule-based detection in UPI applications.

2.4 UPI Fraud Detection Using Machine Learning

This open-access chapter provides a complete ML pipeline for detecting UPI fraud. It describes UPI's role in India's digital economy, and lists common fraud patterns, and constructs a dataset with transaction amounts, timestamps, sender/receiver behavior, and device indicators. Several supervised algorithms, including Logistic Regression, SVM and Random Forest, were evaluated for fraud classification. The study reports that Random Forest outperforms the other models, especially on imbalanced data, and is recommended as a strong baseline for UPI fraud-detection deployments.

2.5 Real-Time Fraud Detection Using Machine Learning

This study is centered on credit card data rather than UPI and contributes important ideas for real-time financial fraud detection. Using a publicly available Kaggle dataset, the author compared several ML algorithms—Logistic Regression, KNN, Naïve Bayes, SVM, Random Forest, XG Boost and Light GBM—to classify transactions as genuine or fraudulent. Class imbalance was handled with SMOTE, and extensive metrics such as AUC, precision-recall AUC, F1, KS statistic and recall were reported. The Random Forest emerged as the best performer. This study also integrates SHAP-based explainability to show which features drive each prediction, bridging the gap between accuracy and interpretability.

3. METHODOLOGY

The proposed system, "UPI Fraud Detection Using Machine Learning," provides a data-driven fraud detection layer on top of the UPI payment workflow. Historical UPI-style transaction data were collected and preprocessed to derive meaningful features, such as transaction amount behavior, balance changes, transaction type, frequency, time patterns, and user behavior indicators. Using these features, multiple machine learning models (e.g., Random Forest, XG Boost, Light GBM, Gradient Boosting) are trained to distinguish between genuine and fraudulent transactions, with special handling for class imbalance so that rare fraud cases are not ignored.

The trained model was integrated into a Flask-based web application that simulated a UPI portal. When a user initiates a transaction, the system extracts the relevant features, passes them to the ML model, and obtains a fraud

probability score in real-time. If the transaction is predicted to be safe, it is allowed; if it is flagged as suspicious, it is blocked or marked as high risk, and the details are logged in the user history. The system also supports generating reports (CSV/PDF) and visual analytics, and can incorporate SHAP-based explainable AI to show which features contribute most to the decision, thereby increasing transparency for both users and administrators.

4. SYSTEM ARCHITECTURE

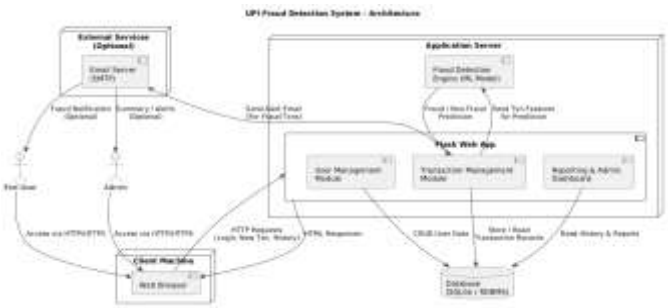

Fig: Architecture Design

4.1 MAIN COMPONENTS

1. Client (Web Browser)
   - User/Admin accesses the UPI Fraud Detection portal using a browser.
   - Sends HTTP requests (login, new transaction, view history) and receives HTML pages.
2. Web Application Layer (Flask App)
   - Handles routing, forms, validation, and session management.
   - Contains modules for:
     - User Authentication &Authorization
     - Transaction Management
     - Fraud Detection Integration
     - Reporting & Admin Dashboard
3. Fraud Detection Engine (ML Model Service)
   - Pre-loaded trained ML model (Random Forest / XG Boost / etc.).
   - Feature-engineering logic that converts raw transaction input into model-ready features.
   - Returns prediction: Fraud / Non-Fraud.
4. Database Layer (SQLite / RDBMS)
   - Stores user details, hashed passwords, roles.
   - Stores transaction records with prediction label (safe / fraud).
   - Stores logs and optionally model metadata / statistics.
5. Reporting & Analysis Module (Inside Flask)
   - Generates summary views, graphs and downloaded reports.
   - Allows admin to monitor suspicious transactions.

Typical flow (High Level):

1. User opens portal →Flask serves UI.
2. User logs in →Flask validates credentials with DB.
3. User initiates transaction →Flask validates input and sends to Fraud Detection Engine.
4. Fraud Detection Engine engineer's features → ML model predicts fraud risk.
5. Flask records transaction + prediction in DB → returns result page:
   - If safe →"Transaction Successful / Safe".
   - If suspicious →"Transaction Blocked / Suspected Fraud".

6.User/Admin Module
   - Admin dashboard to view all users and all transactions.
   - Monitor suspicious / fraudulent activities.
   - Basic system statistics (total transactions, fraud percentage).

7.Notification Module
   - Send email alerts for flagged fraudulent transactions.
   - Can send periodic summary reports to admin.

5. RESULTS

The project "UPI Fraud Detection Using Machine Learning" was developed to address the growing problem of fraudulent transactions in UPI-based digital payments. The work started with a detailed study of UPI, existing fraud patterns and traditional rule-based detection mechanisms. A suitable dataset of UPI-style transactions was collected/assumed, preprocessed, and enriched through feature engineering. Multiple machine learning models were then trained and evaluated to classify transactions as legitimate or fraudulent, with special attention to handling class imbalance. The best-performing model was integrated into a Flask-based web application that simulates a UPI portal. The system allows users to register, log in, initiate transactions, and receive real-time fraud predictions. Each transaction and its prediction result are stored in a database, and users/admins can view transaction history and generate reports. Overall, the project successfully demonstrates how ML-based techniques can enhance security and trust in UPI transactions compared to static rule-based approaches.

6.CONCLUSION

The UPI Fraud Detection Using Machine Learning project successfully demonstrates how data-driven models can strengthen security in digital payment platforms. By combining machine learning, web technologies, and database management, the system is capable of providing

real-time risk assessment for UPI-style transactions and blocking suspicious activities before they cause financial loss. Although this project operates on a simplified dataset and a laboratory-scale deployment, the architecture, design, and implementation approach can be extended to real-world environments with larger datasets and more advanced models.

In future, the system can be enhanced by incorporating richer behavioral features, graph-based analysis of payer-payee networks, integration with real banking APIs, explainable AI visualizations, and deployment on a scalable cloud platform. Overall, the project meet its stated objectives and serves as a strong foundation for further research and development in the area of secure, intelligent digital payment systems.

ACKNOWLEDGEMENT

We wish to express our profound gratitude to our project guide, * Dr Raja Meyyan *, Professor and HOD, Department of Information Science and Engineering, East West Institute of Technology (EWIT), for his invaluable expertise, persistent encouragement, and meticulous guidance throughout this project. We also extend our sincere thanks to the Head of the Department, **Dr Raja Meyyan **, for providing the necessary facilities and support.

REFERENCES

1. S. Munirathinam and B. Ramadoss, “Predictive models for equipment fault detection in the semiconductor manufacturing process,” IACSIT International Journal of Engineering and Technology, vol. 8, no. 4, pp. 273-285, 2025.
2. A. Kaveri Sharma, “A study on effects of intrinsic characteristics of datasets on classification performance,” International Journal of Advanced Research in Computer Science and Software Engineering, vol. 6, no. 1, pp. 198-204, 2025.
3. J. Wang, Z. Yang, J. Zhang, Q. Zhang, and W.-T.-K. Chien, “AdaBalGAN: An improved generative adversarial network with imbalance learning for wafer defective pattern recognition,” IEEE Trans. Semiconductor Manuf., vol. 32, no. 3, pp. 310-319, Aug. 2025.
4. K.-J. Kim, K.-J. Kim, C.-H. Jun, L.-G. Chong, and G.-Y. Song, “Variable selection under missing values and unlabeled data in semiconductor processes,” IEEE Trans. Semiconductor Manuf., vol. 32, no. 1, pp. 121-128, Feb. 2024.
5. J.-S. Kim, S.-J. Jang, T.-W. Kim, H.-J. Lee, and J.-B. Lee, “A productivity-oriented wafer map optimization using yield model based on machine learning,” IEEE Trans. Semiconductor Manuf., vol. 32, no. 1, pp. 39-47, Feb. 2023.