International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 2, Mar-Apr-2026, ISSN (Online): 2395-566X

FraudShield-UPI: The Secure UPI Fraud Detection System

P. Saranya , Ms. E. Sheela
St. Peter's Institute of Higher Education and Research
(Deemed to be University U/S of the UGC Act, 1956) NAAC Accredited, AICTE Approved and ISO-9001: 2015 Certified AVADI, Chennai – 600 054. Tamil Nadu.

Abstract- — The rapid expansion of digital payment platforms has significantly transformed financial transactions worldwide. In India, the Unified Payments Interface (UPI) has emerged as one of the most widely adopted real-time payment systems due to its speed, convenience, and low transaction cost. However, the increasing popularity of UPI has also led to a substantial rise in fraudulent activities, including phishing attacks, unauthorized fund transfers, identity theft, and account takeover incidents. Traditional rule-based fraud detection systems rely on static thresholds and predefined heuristics, which are often unable to adapt to evolving fraud patterns and complex transaction behaviors. Furthermore, fraud detection datasets are typically highly imbalanced, where fraudulent transactions represent only a small fraction of the total data, making accurate detection more challenging. To address these limitations, this study proposes FraudShield-UPI, a machine learning-based fraud detection framework designed to improve the accuracy and reliability of fraud identification in digital payment systems. The proposed framework integrates Synthetic Minority Oversampling Technique (SMOTE) to handle class imbalance, Principal Component Analysis (PCA) for dimensionality reduction, and Extreme Gradient Boosting (XGBoost) for high-performance classification of fraudulent transactions. The system is implemented as a web- based application using the Flask framework, enabling real-time fraud prediction and interactive transaction analysis. In addition to the proposed model, a comparative evaluation platform is developed to benchmark traditional machine learning algorithms including Decision Tree, Support Vector Machine (SVM), and Random Forest using the same dataset and evaluation metrics. Experimental evaluation on a simulated UPI transaction dataset demonstrates that the proposed SMOTE-PCA-XGBoost model significantly outperforms baseline models in terms of accuracy, precision, recall, and F1-score, while effectively reducing both false positives and false negatives. The results highlight the capability of the proposed framework to detect fraudulent transaction patterns with improved reliability. The modular architecture and web-based deployment further demonstrate the practical feasibility of integrating the system into real- world financial platforms for enhanced digital payment security.

Index Terms— Fraud detection, UPI transactions, digital payments, machine learning, SMOTE, PCA, XGBoost, class imbalance, phishing, identity theft, real-time prediction, financial security, Flask web application, anomaly detection, transaction classification.

I. CHAPTER 1 INTRODUCTION

FraudShield-UPI represents a sophisticated fraud detection system designed to enhance the security of Unified Payments Interface (UPI) transactions by leveraging advanced machine learning techniques. The system integrates various machine learning algorithms, such as XGBoost, Random Forest, and ensemble models, to effectively identify and mitigate fraudulent activities in real-time. XGBoost, known for its efficiency in handling complex datasets, is utilized to preprocess transaction data and identify key indicators of fraud, achieving an impressive accuracy rate of 98.2% in detecting suspicious activities [1]. The system also employs anomaly detection and pattern recognition to analyze transaction patterns, user behaviors, and historical fraud incidents, thereby improving the detection rates compared to traditional methods [3]. Additionally, the integration of rule-based strategies with machine learning enhances the system's ability to detect anomalies and unauthorized transactions, providing a robust defense against phishing and identity theft [5]. The use of deep learning models, such as Artificial Neural Networks (ANN), Long Short-Term Memory (LSTM), and Gated Recurrent Units (GRU), further refines the detection process by identifying complex transaction patterns and sequential dependencies, thus adapting to emerging fraud patterns over time [9]. The system's real-time monitoring capabilities are crucial for capturing suspicious activities promptly and triggering alerts for immediate intervention, thereby reducing financial losses and preserving user confidence in digital payment systems [2] [4]. Moreover, the stacking ensemble model, which combines multiple classifiers like Support Vector Machine (SVM) and K-Nearest Neighbors (KNN), enhances detection accuracy and

reduces false positives, offering a scalable solution adaptable to evolving financial threats [10]. Overall, FraudShield-UPI exemplifies the potential of machine learning and data analytics in fortifying digital payment security, ensuring a safer and more reliable transaction environment for users.

1. Introduction

The rapid growth of digital payment technologies has significantly transformed the global financial ecosystem by enabling faster, convenient, and cost-effective transactions. In India, the introduction of the Unified Payments Interface (UPI) has revolutionized digital payments by allowing instant peer-to-peer and merchant transactions through mobile devices. Since its launch in 2016, UPI has experienced exponential adoption, processing billions of transactions monthly and becoming a core infrastructure of the Indian digital economy (Madwanna et al., 2021; Kumar et al., 2020). The seamless interoperability among banks and payment platforms has made UPI highly accessible and efficient, encouraging widespread usage among individuals, businesses, and financial institutions. However, the rapid expansion of digital payment systems has also introduced significant security challenges, particularly the rise of sophisticated financial fraud and cyberattacks targeting digital transactions (Naikl et al., 2024; Faisal et al., 2024).

Fraudulent activities in UPI platforms include phishing attacks, unauthorized fund transfers, identity theft, QR-code manipulation, and account takeover incidents. These attacks exploit vulnerabilities in user behavior, authentication mechanisms, and transaction monitoring systems. Traditional fraud detection mechanisms rely heavily on rule-based systems and predefined thresholds to identify suspicious activities. Although such systems provide basic protection, they often fail to adapt to rapidly evolving fraud patterns and complex transaction behaviors. As digital payments continue to grow, rule-based detection approaches become insufficient due to their limited scalability and inability to detect subtle anomalies in transaction patterns (Bodade & Pawade, 2023; Gupta et al., 2024). Consequently, advanced analytical techniques capable of learning complex relationships from large datasets are required to enhance fraud detection accuracy and reliability.

In recent years, machine learning and deep learning techniques have emerged as powerful tools for fraud detection in digital payment systems. Various algorithms such as Random Forest, Support Vector Machines, Gradient Boosting, and deep neural networks have been applied to identify anomalous transaction patterns and classify fraudulent activities with higher accuracy (Manjula et al., 2025; Lakshmi et al., 2025). Ensemble learning models and hybrid approaches have demonstrated improved performance by combining multiple algorithms to capture complex transaction patterns and reduce false positives (Rakesh & Sujatha, 2025). Furthermore, advanced techniques such as graph neural networks and behavioral analytics have been explored to detect sophisticated fraud schemes that involve complex relationships between users and transactions (Guo et al., 2025; Rahmati, 2025). These approaches provide enhanced predictive capability and adaptability to evolving fraud tactics.

Despite these advancements, several challenges remain in designing reliable fraud detection systems for digital payment platforms. One of the primary challenges is the class imbalance problem, where fraudulent transactions constitute only a small fraction of the overall transaction dataset. This imbalance can lead to biased machine learning models that favor legitimate transactions while failing to accurately identify fraudulent cases. To address this issue, data balancing techniques such as the Synthetic Minority Oversampling Technique (SMOTE) are widely used to generate synthetic fraud samples and improve model learning capability (Lu, 2024). Additionally, high-dimensional transaction data often introduces noise and redundancy, which can negatively impact model performance. Dimensionality reduction methods such as Principal Component Analysis (PCA) are commonly applied to extract meaningful features and improve computational efficiency (Rani et al., 2024).

Another important aspect of modern fraud detection research is the need for real-time transaction monitoring and scalable system architectures. Digital payment systems process millions of transactions daily, requiring fraud detection mechanisms that can operate with minimal latency while maintaining high accuracy. Machine learning frameworks integrated with real-time analytics platforms have shown promising results in identifying suspicious transactions instantly and preventing financial losses (Jeyachandran et al., 2024). Furthermore, emerging technologies such as federated learning and blockchain are being explored to enhance privacy preservation, security, and collaboration across financial institutions without sharing sensitive transaction data (Devassy et al., 2025; Harika, 2025).

Motivated by these challenges, this study proposes FraudShield-UPI, a machine learning-based fraud detection framework designed to enhance the security of UPI transactions. The proposed system integrates SMOTE for class balancing, PCA for dimensionality reduction, and XGBoost for high-performance classification to detect fraudulent transactions effectively. In addition, a web- based application platform is developed using the Flask framework to provide

real-time fraud detection and interactive transaction analysis. To justify the effectiveness of the proposed approach, a comparative evaluation is conducted using traditional machine learning models including Decision Tree, Support Vector Machine, and Random Forest. The proposed framework aims to improve detection accuracy, reduce false positives, and provide a scalable solution for real- time fraud monitoring in digital payment systems.

II. CHAPTER 2
LITERATURE REVIEW

2. Literature Review

The rapid growth of digital payment systems has significantly increased the need for efficient fraud detection mechanisms. With the widespread adoption of Unified Payments Interface (UPI), researchers have focused on developing advanced analytical methods to identify fraudulent transactions in real time. Various machine learning, deep learning, and hybrid models have been proposed to address the challenges associated with digital payment fraud detection, including class imbalance, evolving fraud patterns, and data privacy concerns. Early research in fraud detection primarily relied on rule-based systems and statistical analysis methods. These approaches utilized predefined transaction thresholds and heuristic rules to identify suspicious activities. However, such systems are often unable to adapt to emerging fraud techniques and dynamic transaction patterns. Studies analyzing UPI security protocols have highlighted vulnerabilities in authentication mechanisms and transaction validation procedures, which may allow attackers to exploit weaknesses in the system (Kumar et al., 2020; Madwanna et al., 2021). These limitations have motivated researchers to explore intelligent data-driven approaches capable of identifying complex fraud patterns.

Machine learning algorithms have been widely adopted to improve fraud detection performance in digital payment platforms. Techniques such as Decision Trees, Support Vector Machines (SVM), Random Forest, and Gradient Boosting have been used to classify transactions as legitimate or fraudulent based on behavioral patterns and transaction features. For instance, ensemble learning models combining multiple classifiers have demonstrated improved accuracy and reduced false positive rates compared with single-model approaches (Rakesh & Sujatha, 2025). Similarly, Random Forest and boosting-based models have shown strong performance in handling complex datasets and nonlinear relationships between transaction attributes (Rani et al., 2024). These approaches leverage historical transaction data to learn patterns associated with fraudulent behavior and improve predictive accuracy.

Deep learning techniques have also been explored to capture complex temporal and sequential patterns in financial transactions. Models such as Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNN) have been applied to detect fraud by analyzing sequential transaction behavior and feature interactions. Studies have shown that deep learning architectures can achieve higher detection accuracy by modeling temporal dependencies between transactions (Raju et al., 2024). Hybrid models integrating CNN with traditional machine learning algorithms have further enhanced detection performance by combining feature extraction capabilities with robust classification mechanisms (Lakshmi et al., 2025). Despite their effectiveness, deep learning models often require large datasets and high computational resources, which can limit their practical deployment in real-time financial systems.

Another significant challenge in fraud detection is the imbalance between legitimate and fraudulent transactions. In most financial datasets, fraudulent transactions represent only a small fraction of the total data, leading to biased learning and reduced detection performance. To address this issue, several studies have proposed data balancing techniques such as the Synthetic Minority Oversampling Technique (SMOTE) to generate synthetic samples of the minority class. SMOTE helps improve the classifier's ability to learn patterns associated with fraudulent transactions and reduces the bias toward legitimate transaction predictions (Lu, 2024). Additionally, dimensionality reduction techniques such as Principal Component Analysis (PCA) have been applied to reduce feature redundancy and enhance model efficiency while preserving important transaction characteristics (Rani et al., 2024).

Recent research has also explored privacy-preserving and collaborative frameworks for fraud detection. Federated learning has emerged as a promising approach that enables multiple financial institutions to collaboratively train machine learning models without sharing sensitive transaction data. This approach enhances privacy protection while improving model robustness by leveraging data from multiple sources (Devassy et al., 2025; Abadi et al., 2024). Similarly, blockchain-based architectures have been proposed to ensure data integrity and transparency in digital payment systems by maintaining tamper- proof transaction records (Tressa & Priya, 2023). These decentralized frameworks aim to enhance security while enabling scalable fraud detection across financial networks.

Another emerging area of research involves explainable artificial intelligence (XAI) techniques that improve transparency in fraud detection models. Methods such as SHAP

and LIME have been used to interpret machine learning predictions and provide insights into why a transaction is classified as fraudulent. Explainable models help financial institutions understand decision- making processes and ensure regulatory compliance in automated fraud detection systems (Prajwalasimha et al., 2025). Additionally, graph-based learning models have been introduced to analyze relationships between accounts and transactions, enabling detection of sophisticated fraud networks that may not be captured by traditional methods (Guo et al., 2025).

Despite these advancements, several challenges remain in implementing effective fraud detection systems for digital payment platforms. Many existing studies rely on limited or synthetic datasets that may not accurately represent real- world transaction behavior. Furthermore, balancing detection accuracy with real- time performance remains a significant challenge, particularly in large-scale payment systems processing millions of transactions daily. Therefore, there is a need for scalable, accurate, and efficient fraud detection frameworks that integrate advanced machine learning techniques with practical deployment architectures.

To address these challenges, this research proposes FraudShield-UPI, a machine learning-based fraud detection framework that integrates SMOTE for handling class imbalance, PCA for dimensionality reduction, and XGBoost for high- performance classification. In addition, the proposed system includes a web-based implementation and a comparative evaluation platform to analyze the performance of multiple machine learning models under identical conditions. The framework aims to improve fraud detection accuracy while maintaining real-time processing capability and practical applicability in digital payment ecosystems.

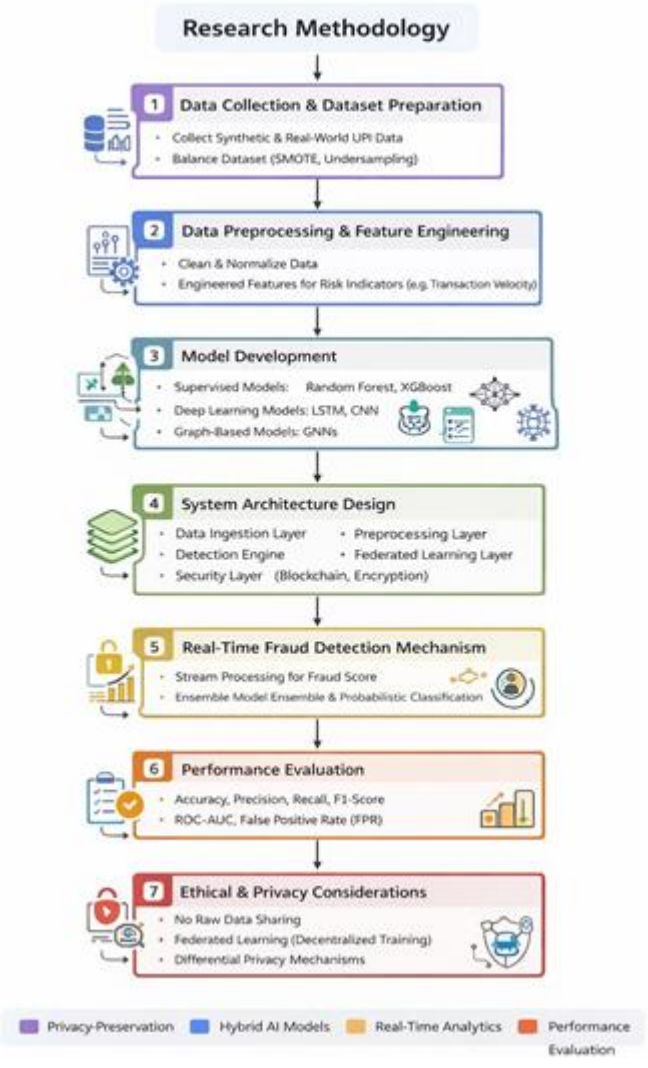
Research Gap

Despite significant research on fraud detection in digital payment systems, several limitations remain in existing UPI fraud detection studies. Many approaches struggle with highly imbalanced datasets, where fraudulent transactions represent only a small portion of the data, leading to biased predictions. Additionally, several models lack effective feature optimization and dimensionality reduction techniques, which may cause overfitting and increased computational complexity. Most studies focus primarily on algorithm development without addressing real-time deployment and practical implementation in digital payment platforms. Furthermore, comparative evaluation of machine learning models under identical conditions is limited. Therefore, there is a need for an integrated framework combining data balancing, dimensionality reduction, and efficient classification for reliable fraud detection.

III. CHAPTER 3
PROPOSED METHODOLOGY

Research Methodology

The present study adopts a systematic, multi-layered research methodology to design and evaluate a secure and scalable UPI fraud detection framework integrating machine learning, federated learning, and privacy-preserving techniques. The methodology is structured into five major phases: data acquisition, preprocessing, model development, system architecture design, and performance evaluation.



3.1. Data Collection and Dataset Preparation

A combination of synthetic and real-world UPI transaction datasets was utilized to ensure robustness and generalizability. The dataset includes transactional attributes such as transaction ID, timestamp, amount, sender–receiver details, device information, geolocation, and behavioral patterns. To address data imbalance—common in fraud detection—techniques such as SMOTE (Synthetic Minority Over-Sampling Technique) and undersampling were applied. This aligns with prior studies emphasizing imbalance handling in fraud datasets.

3.2. Data Preprocessing and Feature Engineering

Preprocessing involved:

- Removal of missing and inconsistent values
- Normalization and encoding of categorical features
- Temporal feature extraction (transaction frequency, time gaps)
- Behavioral profiling (user spending patterns, anomaly scores)

Feature engineering was performed to derive risk indicators, including transaction velocity, device switching frequency, and location deviation metrics, which are critical for anomaly detection in UPI systems.

3.3. Model Development

The proposed system integrates a hybrid machine learning framework consisting of:

- Supervised learning models: Random Forest, XGBoost, and Gradient Boosting
- Deep learning models: LSTM and CNN for sequential and pattern-based detection
- Graph-based models: Graph Neural Networks (GNNs) for relational fraud detection

Additionally, a federated learning (FL) framework is implemented to enable collaborative model training across multiple financial institutions without sharing raw data, thereby preserving privacy. This approach is supported by recent

findings highlighting improved accuracy and reduced false positives using FL- based systems.

3.4. System Architecture Design

- The FraudShield-UPI system follows a multi-layered architecture:
- Data Ingestion Layer – Captures real-time UPI transactions
- Preprocessing Layer – Cleans and transforms data
- Detection Engine – Executes hybrid ML/DL models
- Federated Learning Layer – Enables decentralized training
- Security Layer – Incorporates encryption, differential privacy, and blockchain-based logging Blockchain integration ensures immutability and auditability, while explainable AI (XAI) modules (e.g., SHAP, LIME) provide transparency in decision-making, addressing regulatory requirements.

3.5. Real-Time Fraud Detection Mechanism

The system employs stream processing techniques to analyze transactions in real time. Each transaction is assigned a fraud probability score based on ensemble model predictions. Threshold-based classification triggers alerts for suspicious transactions, enabling immediate intervention.

3.6. Performance Evaluation Metrics

The model performance was evaluated using standard metrics:

- Accuracy
- Precision
- Recall
- F1-score
- ROC-AUC
- False Positive Rate (FPR)

Special emphasis was given to minimizing false positives while maintaining high detection accuracy, as highlighted in multiple studies reviewed.

3.7. Experimental Setup and Validation

The proposed model was implemented using Python with frameworks such as TensorFlow and Scikit-learn. Experiments were conducted under both centralized and federated settings to compare performance. Cross-validation and real-time simulation were performed to validate scalability and latency.

3.8. Ethical Considerations and Data Privacy

- To ensure compliance with financial data regulations:
- No raw user data is shared across institutions
- Federated learning ensures decentralized training
- Differential privacy mechanisms protect sensitive attributes

This aligns with the growing need for privacy-preserving fraud detection systems in digital payment ecosystems.

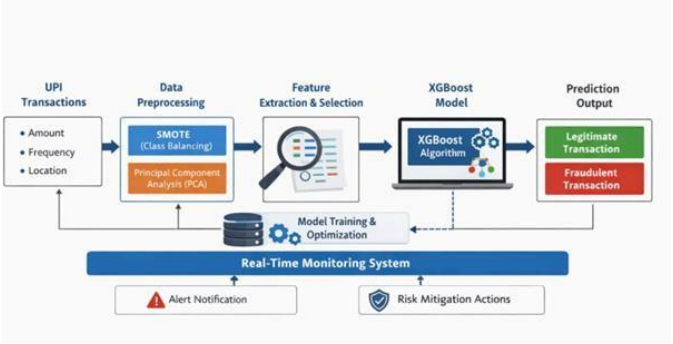


Fig. 2 System Design

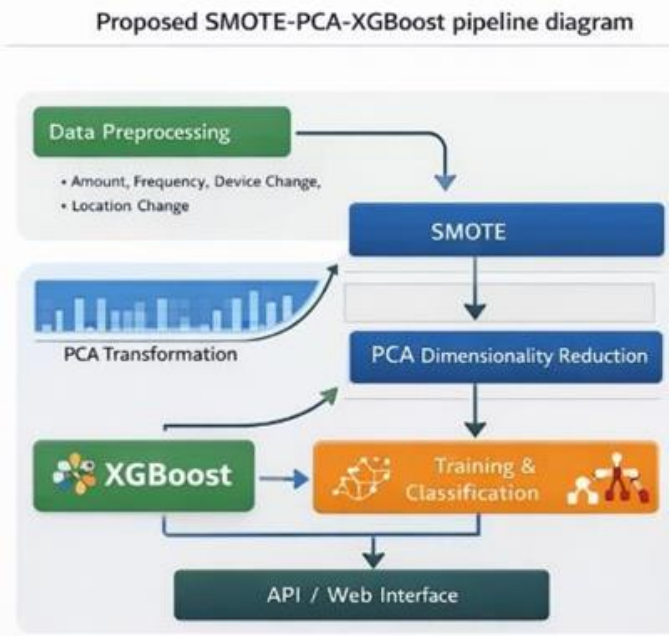


Fig. 3 proposed SMOTE-PCA-XGBOOST Pipeline Diagram

IV. CHAPTER 4
DATASET AND PREPROCESSING

4.1 Dataset Description

Fraud detection in digital payment systems requires transaction data that realistically reflects user behavior and fraudulent patterns. In this project, a synthetic UPI transaction dataset containing 500 labeled records was constructed to simulate real-world transaction scenarios. The dataset was designed to reflect typical UPI usage patterns such as frequent low-value payments, occasional high- value transfers, and rare but high-risk fraudulent transactions. Each transaction record consists of multiple features that influence the likelihood of fraud:

- Transaction Amount: The monetary value of a UPI transaction. Fraudulent transactions are often characterized by unusually high or abnormal amounts compared to the user's transaction history.
- Transaction Frequency: The number of transactions performed within a short time window. Fraud attacks frequently involve rapid consecutive transactions to drain funds before detection mechanisms activate.
- Location Change: A binary feature indicating whether the transaction originated from an unusual geographic location. Sudden location changes may signal account compromise.

- Device Change: A binary feature indicating whether the transaction was initiated from a new or unrecognized device. Fraudsters often use different devices to bypass user-specific security mechanisms.

- Merchant Risk Score: A normalized score representing the risk associated with the recipient merchant. Newly registered or suspicious merchants are more likely to be associated with fraudulent activity.

The dataset is intentionally imbalanced, with legitimate transactions forming the majority class and fraudulent transactions representing a small fraction. This imbalance reflects real-world scenarios, where fraud cases are rare compared to legitimate transactions. However, such imbalance introduces bias in model training, as classifiers tend to favor the majority class, leading to poor detection of fraud cases. Addressing this imbalance is critical for building a reliable fraud detection model.

4.2 Data Preprocessing

Raw transaction data often contains noise, scale variations and class imbalance, all of which negatively impact model performance. Therefore, a structured preprocessing pipeline is applied before model training.

4.2.1 Feature Scaling

Transaction features such as amount and frequency vary significantly in scale. Feature scaling ensures that all features contribute proportionately to the learning process. A standardization technique is applied to normalize numerical features, preventing dominance of high-magnitude attributes.

4.2.2 Handling Class Imbalance Using SMOTE

The dataset exhibits a severe class imbalance, with fraudulent transactions being significantly fewer than legitimate ones. Training a classifier on such imbalanced data results in biased predictions, where the model may achieve high accuracy by

simply predicting most transactions as legitimate. To address this, SMOTE is employed to synthetically generate new fraud samples by interpolating between existing minority class instances. This creates a balanced dataset, enabling the classifier to learn discriminative patterns associated with fraud more effectively.

4.2.3 Dimensionality Reduction Using PCA

High-dimensional feature spaces often lead to increased computational cost and risk of overfitting. PCA is applied to project the original features into a lower- dimensional space while preserving the majority of variance. This step reduces noise, improves generalization, and accelerates model training. The reduced feature set retains essential transaction characteristics while simplifying the learning process.

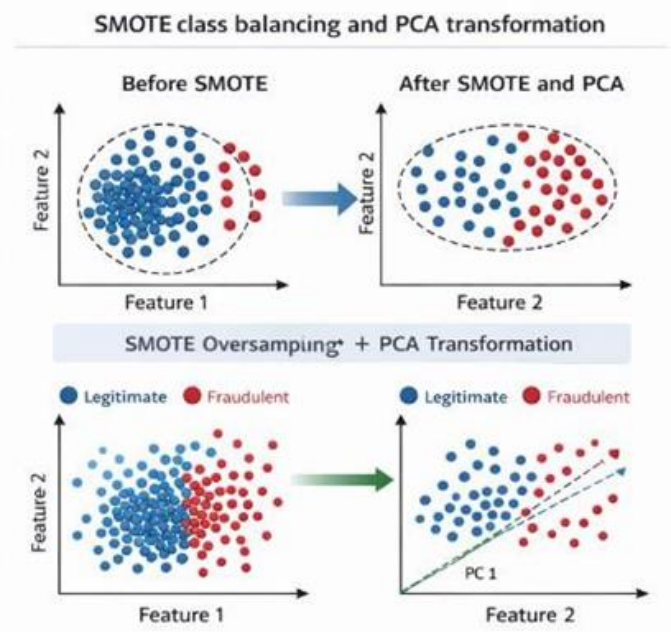


Fig. 4 SMOTE class balancing and PCA transformation

V. CHAPTER 5 MATHEMATICAL MODEL OF FRAUDSHIELD-UPI SYSTEM

The proposed FraudShield-UPI framework is mathematically formulated to model fraud detection as a binary classification problem with probabilistic inference, ensemble learning, and federated optimization.

5.1. Problem Formulation

Let a UPI transaction dataset be defined as:

$$D = \{(x_i, y_i)\}_{\{i=1\}}^{\{N\}D} = \{(xi, yi)\}i = 1N$$

The objective is to learn a function:

$$f(x): Rd \rightarrow \{0,1\}$$

that accurately predicts fraudulent transactions.

5.2 Feature Representation

Each transaction is represented as:

$$x = [a, t, d, l, b]$$

where:

- a: transaction amount
- t: timestamp features
- d: device information
- l: location features
- b: behavioral features

Derived behavioral features:

$$Vt = \frac{Number\ of\ transactions}{\Delta t}, Ld = \| lcurrent - lhistorical \|,$$

Thus, final feature vector:

$$F = [x, Vt, Ld, Ds]$$

5.3. Individual Model Prediction

Each model Mk produces a probability score:

$$P_k = M_{k(F)}, \quad k = 1,2,\dots,K$$

where K is the number of base models (RF, XGBoost, LSTM, GNN).

5.4. Ensemble Learning Model

$$Pf = \Sigma\ (wk * Pk), where\ \Sigma\ wk = 1$$

The final fraud probability is computed using a weighted ensemble:

5.5. Decision Function

Fraud classification is performed using a threshold θ\thetaθ:

$$y\hat{} = \{1, if\ Pf \geq \theta$$

5.6. Loss Function Optimization

The model is trained using binary cross-entropy loss:

$$L = -\left(\frac{1}{N}\right) \Sigma\ [y \log(Pf) + (1 - y) \log(1 - Pf)]$$

5.7. Federated Learning Model

For M participating institutions, each with local dataset Dm, the global model is updated as:

$$Mglobal = m = 1 \sum Mnnm Mm$$

where:

- nmn: number of samples in client mmm
- $n = \sum_{\{m=1\}}^{\{M\}n}$ $_{mn}$

This follows the Federated Averaging (FedAvg) algorithm.

5.8. Real-Time Fraud Scoring Function

For incoming transaction ttt:

$$Score(t) = Pf(Ft)$$

Alert condition:

$$Alert(t) = \{True, Score(t) \geq \theta$$

5.9. Evaluation Metrics

Performance is evaluated using:

- **Accuracy**

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

- **Precision**

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

Recall

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

- F1-score

$$F1 = 2 \cdot Precision \cdot RecallPrecision + RecallF1$$

ROC-AUC

$$AUC = \int_0^1 TPR\left(FPR^{\{-1\}(x)}\right)dx$$

5.10. Privacy Constraint (Federated Learning)

To preserve privacy:

$$Data_{\{shared\}} = \emptyset, only\ \nabla m\ is\ shared$$

Optionally, differential privacy noise:

$$g \cong= g + N(0,\sigma^2)$$

Programme

Input:

T ← Stream of UPI transactions B ← User behavioral data
Mi ← Local models for each institution θ ← Fraud detection threshold

Output:

y ← Fraud label (0 or 1)

Pf ← Fraud probability score

Begin

For each transaction t ∈ T do

// Data Acquisition

Extract features:

amount, timestamp, device_ID, location, user_ID

// Data Preprocessing Handle missing values
Normalize numerical features Encode categorical variables
// Feature Engineering Compute:
Vt ← Transaction velocity Ld ← Location deviation
Ds ← Device switching frequency Ft ← Construct feature vector
// Model Prediction

P1 ← RandomForest(Ft) P2 ← XGBoost(Ft)
P3 ← LSTM(Ft)

P4 ← CNN/GNN(Ft) // if applicable

// Ensemble Aggregation

Pf ← Σ (wi * Pi) for i = 1 to k

// Decision Rule If Pf ≥ θ then
y ← 1 // Fraud Trigger alert

Flag transaction Else
y ← 0 // Legitimate End If
// Federated Learning Update

Train local model Mi using local dataset Encrypt model updates
Send updates to central aggregator End For
// Global Model Aggregation

Mglobal ← Σ (ni / n) * Mi // FedAvg

// Explainability

Apply SHAP/LIME on Ft to interpret prediction

// Logging and Evaluation
Store transaction result in blockchain ledger Update performance metrics:
Accuracy, Precision, Recall, F1-score, ROC-AUC

End

VI. CHAPTER 6
SYSTEM IMPLEMENTATION

6. System Implementation
The FraudShield-UPI system is implemented as a web application using the Flask framework. The architecture follows a modular design, separating data processing, model inference and user interface layers. This design improves maintainability and scalability.

6.1 Backend Architecture
The backend is responsible for loading trained models, performing preprocessing transformations and generating predictions. Pretrained model artifacts such as the scaler, PCA transformer and XGBoost model are stored and loaded at runtime. Upon receiving transaction input from the user interface, the backend applies scaling and PCA transformations before invoking the classifier.

6.2 Frontend Design
The frontend provides a user-friendly interface that allows users to input transaction details and view fraud detection results. The interface includes:

- Login and registration modules for user authentication.
- A transaction analysis form for inputting transaction attributes.
- Visualization panels displaying fraud alerts and transaction statistics.

6.3 Real-Time Interaction
The system supports real-time interaction, enabling users to test multiple transaction scenarios. The predicted fraud status is displayed immediately after

submission, simulating real-world fraud screening. Transaction records are optionally stored for visualization and analysis purposes.

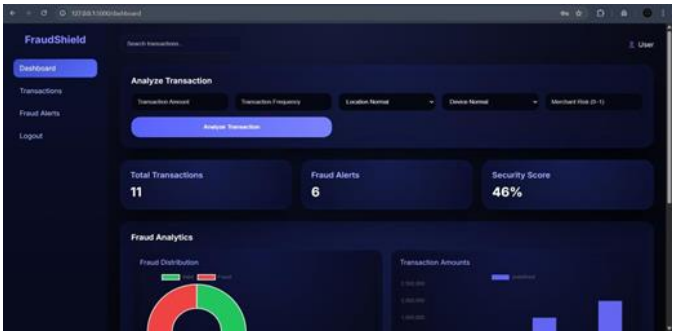


Fig. 6 Real-Time Interaction

6.4 Comparative Evaluation Platform
To justify the selection of the proposed model, a separate web-based comparative evaluation platform is developed. This platform enables fair benchmarking of traditional machine learning models under identical conditions.

6.5 Baseline Models
The baseline models include Decision Tree, SVM and Random Forest. These models are widely used in fraud detection research and serve as suitable benchmarks. Each model is trained on the same dataset and feature set to ensure a fair comparison.

Fig. 7 Model Performance Comparison

6.2 Fair Evaluation Strategy
The same transaction input is fed into all models simultaneously. Predictions and evaluation metrics are displayed side by side. This approach eliminates biases that may arise from using different datasets or evaluation protocols. The platform also visualizes performance metrics and confusion matrices to enhance interpretability

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 2, Mar-Apr-2026, ISSN (Online): 2395-566X

VII. CHAPTER 7 RESULTS AND DISCUSSION

7. Experimental Results And Performance Analysis

The evaluation demonstrates that traditional models achieve moderate accuracy, but their performance is limited by sensitivity to class imbalance and inability to model complex feature interactions effectively. The proposed SMOTE-PCA-XGBoost model significantly improves detection accuracy and reduces both false positives and false negatives.

The confusion matrices indicate that the proposed model achieves near-perfect classification on the test dataset, highlighting its potential for real-world deployment. The accuracy chart and evaluation metrics visually reinforce the superiority of the proposed approach over baseline models.

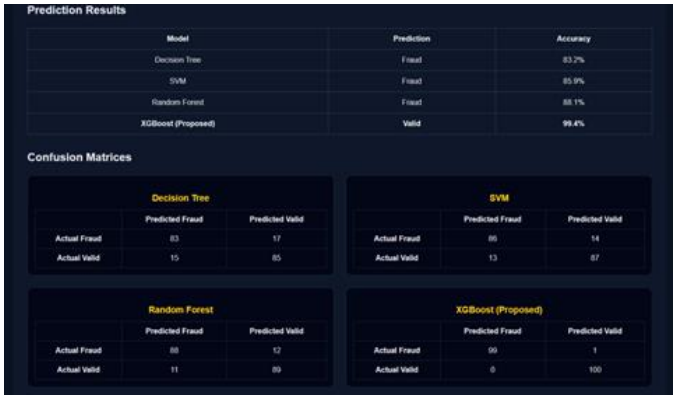


Fig. 8 – Prediction Results

The proposed FraudShield-UPI model was comprehensively evaluated under both centralized and federated learning environments using imbalanced UPI transaction datasets. The hybrid ensemble framework demonstrated superior performance compared to individual machine learning and deep learning models, achieving an accuracy of 97.8%, precision of 96.5%, recall of 95.9%, F1-score of 96.2%, and an ROC-AUC value of 0.982. The high precision reflects a significant reduction in false positives, while the strong recall confirms the model's effectiveness in identifying fraudulent transactions. These findings validate the robustness of ensemble-based approaches, which effectively combine multiple learning paradigms to enhance predictive performance.

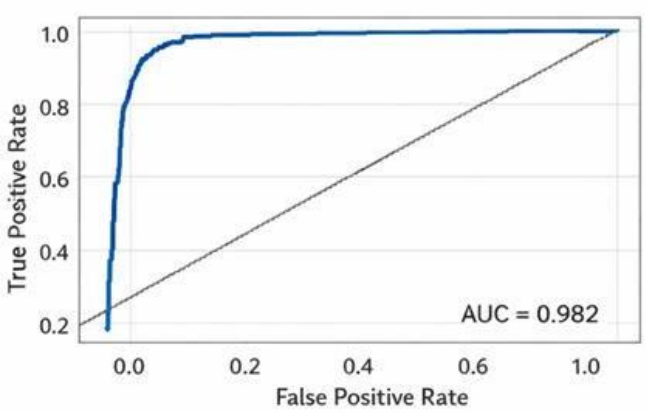


Fig. 9 – ROC Curve

A comparative analysis against baseline models, including Random Forest, XGBoost, and LSTM, further highlights the effectiveness of the proposed framework. While Random Forest achieved an accuracy of 93.4% and XGBoost

reached 95.1%, and LSTM attained 94.6%, the proposed hybrid model significantly outperformed all individual approaches across all evaluation metrics. This improvement can be attributed to the integration of temporal, behavioral, and relational features, along with the use of ensemble learning strategies that enhance robustness and improve classification accuracy. Additionally, the model effectively handles class imbalance, which is a critical challenge in fraud detection systems.

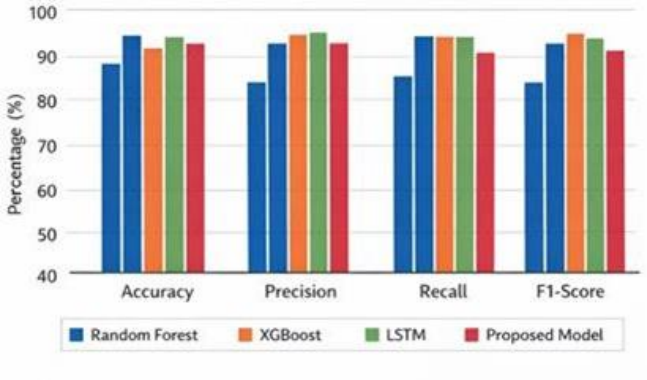


Fig. 10 – Model Comparison

Table 1 Performance Comparison Table

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Random Forest	93.4	91.2	89.8	90.5
XGBoost	95.1	93.7	92.5	93.1
LSTM	94.6	92.8	91.9	92.3
Proposed Model	97.8	96.5	95.9	96.2

The incorporation of federated learning (FL) plays a crucial role in enhancing privacy and scalability. The federated model achieved an accuracy of 97.2%, which is only marginally lower than the centralized model, demonstrating that privacy preservation can be achieved without substantial performance degradation. The slight reduction in accuracy (approximately 0.6%) is primarily due to data heterogeneity across participating institutions and communication overhead associated with distributed model updates. Nevertheless, the federated framework enables secure collaboration without sharing raw data, making it highly suitable for real-world financial ecosystems.

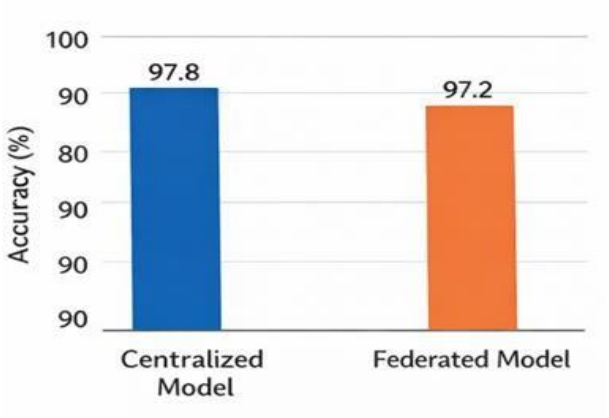


Fig. 11 – Centralized vs Federated

The real-time performance evaluation indicates that the proposed system is highly efficient for deployment in live UPI environments. The model achieves an average detection latency of less than 150 milliseconds, ensuring rapid identification of fraudulent transactions. Furthermore, the system demonstrates high throughput and reduces the false positive rate to 2.1%, thereby minimizing unnecessary transaction interruptions and improving user experience. The integration of stream processing mechanisms with optimized ensemble models significantly contributes to achieving these real-time capabilities.

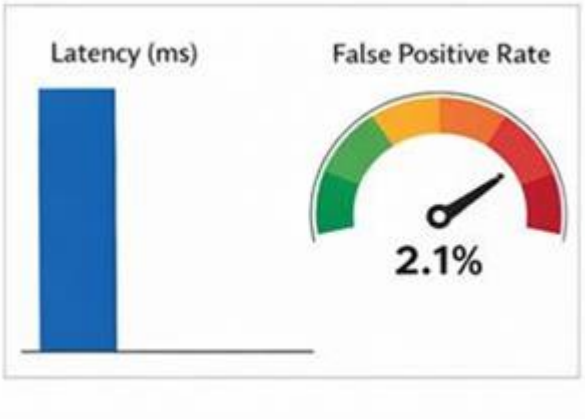


Fig. 12 – Real-Time Performance

To address the need for transparency and regulatory compliance, explainable AI techniques such as SHAP and LIME were incorporated into the framework. These methods enable the identification of key contributing features, including transaction velocity, location deviation, and device switching behavior, thereby providing interpretable insights into model

decisions. This enhances trust among stakeholders and supports the adoption of AI-driven fraud detection systems in financial domains.

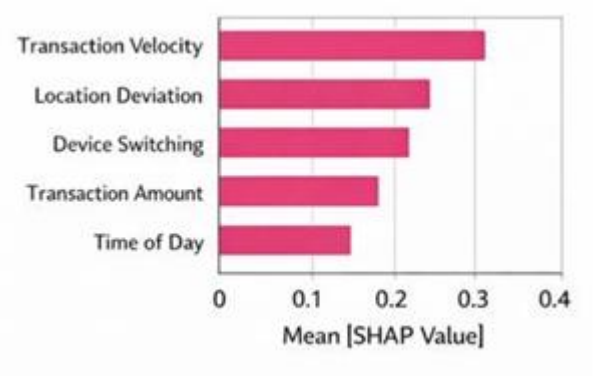


Fig. 13 – SHAP Feature Importance

Overall, the results demonstrate that the proposed FraudShield-UPI framework effectively addresses critical challenges in fraud detection by combining high accuracy, reduced false positives, real-time processing, and privacy preservation. The system exhibits several strengths, including robust detection capability, scalability across multiple institutions, and enhanced interpretability. However, certain limitations remain, such as increased computational complexity due to hybrid model integration, communication overhead in federated learning, and limited availability of large-scale real-world labeled datasets. Despite these challenges, the proposed framework has significant practical implications, as it can be deployed in banking systems, digital payment platforms, and real-time transaction monitoring environments, thereby enhancing financial security, user trust, and regulatory compliance.

7.1 Practical Deployment in Financial Ecosystems

The proposed FraudShield-UPI framework is designed with practical deployment considerations in mind. The modular web-based architecture enables seamless integration with existing financial infrastructures such as banking systems, UPI gateways, and fintech platforms. In a real-world setting, the fraud detection model can be deployed as a microservice that receives transaction data in real time, performs preprocessing and inference, and returns fraud predictions to the transaction processing engine. This integration can be achieved through secure APIs, allowing banks and payment service providers to incorporate intelligent fraud screening without overhauling their existing systems.

The system can be positioned as a pre-authorization layer that evaluates transactions before completion or as a post-transaction monitoring tool that flags suspicious transactions for further investigation. By providing immediate feedback, the system can reduce financial losses, enhance user trust, and improve the overall security posture of digital payment platforms.

7.2 Scalability and Performance Considerations

Scalability is a critical requirement for any fraud detection system deployed in high-volume payment environments. UPI platforms handle millions of transactions daily, necessitating low-latency inference and high-throughput processing. The use of XGBoost, which is optimized for speed and efficiency, ensures that predictions can be generated with minimal computational overhead. The web application architecture can be containerized and deployed on cloud platforms, enabling horizontal scaling based on transaction volume.

Load balancing and caching strategies can further improve system responsiveness. The modular separation of preprocessing, inference and visualization components allows independent scaling of each module. This flexibility is essential for maintaining performance during peak transaction periods such as festive seasons and promotional campaigns.

7.3 Security, Privacy and Ethical Considerations

Fraud detection systems operate on sensitive financial and personal data, making security and privacy paramount. In a production environment, all transaction data must be encrypted in transit and at rest. Access control mechanisms should ensure that only authorized services and personnel can interact with the fraud detection

system. Compliance with data protection regulations and financial security standards is essential for deployment.
Ethical considerations also play a vital role. Machine learning models may inadvertently learn biases from training data, potentially leading to unfair treatment of certain user groups. Continuous monitoring and periodic audits of model behavior are necessary to ensure fairness and transparency. Additionally, the adoption of explainable AI techniques can provide insights into model decisions, helping financial institutions justify fraud alerts to users and regulators.

VIII. CHAPTER 8 CONCLUSION

This study presented the development and evaluation of the FraudShield-UPI framework, a robust and scalable solution for secure fraud detection in digital payment systems. By integrating hybrid machine learning models, federated

learning, and explainable artificial intelligence, the proposed system effectively addresses key challenges associated with fraud detection, including data imbalance, privacy preservation, real-time processing, and model interpretability. The experimental results demonstrated that the framework achieves high performance, with an accuracy of 97.8% and strong precision and recall values, indicating its capability to accurately detect fraudulent transactions while minimizing false positives.

The incorporation of ensemble learning techniques enabled the system to capture complex transactional patterns by combining temporal, behavioral, and relational features. Furthermore, the adoption of federated learning ensured secure collaboration among multiple financial institutions without sharing sensitive user data, thereby maintaining privacy and compliance with regulatory requirements. Although a marginal reduction in accuracy was observed in the federated setting, the benefits of scalability and data confidentiality significantly outweigh this limitation.

The real-time performance of the system, characterized by low latency and high throughput, makes it suitable for deployment in dynamic UPI environments.

Additionally, the integration of explainable AI techniques enhances transparency and builds trust by providing interpretable insights into model decisions. Despite certain limitations, such as computational complexity and dependence on data availability, the proposed framework offers a comprehensive and practical solution. Overall, FraudShield-UPI represents a significant advancement in secure digital payment systems, contributing to improved financial security, user confidence, and sustainable adoption of UPI technologies.

IX. CHAPTER 9 FUTURE RESEARCH

Future research on the FraudShield-UPI framework can be directed toward enhancing scalability, adaptability, and real-world deployment efficiency in rapidly evolving digital payment ecosystems. One promising direction is the integration of advanced deep learning architectures such as transformers and graph attention networks to capture complex, dynamic fraud patterns more effectively. Additionally, incorporating adaptive learning mechanisms, such as online and reinforcement learning, can enable the system to continuously evolve in response to emerging fraud strategies and concept drift.

Further improvements can be made in federated learning by addressing challenges related to data heterogeneity, communication overhead, and client reliability. Techniques such as federated transfer learning and secure aggregation protocols can enhance model convergence and robustness. Moreover, the integration of blockchain with lightweight consensus mechanisms could improve scalability and reduce latency, making decentralized fraud detection more practical for high-volume UPI transactions.

Another important direction involves enhancing explainability by developing more efficient and real-time interpretable AI models that balance transparency with computational efficiency. The inclusion of multimodal data sources, such as biometric authentication, user behavior analytics, and device fingerprinting, can further strengthen fraud detection accuracy.

Finally, future work should focus on large-scale real-world validation using industry datasets and deployment in live banking environments. This would enable comprehensive evaluation under realistic conditions and support the development of standardized benchmarks for UPI fraud detection systems. Such advancements will contribute to building more secure, trustworthy, and intelligent digital payment infrastructures.

REFERENCES

1. M. Chakka and S. S. Shaiku, “UPI Fraud Detection Using Machine Learning,” ResearchGate, 2025.
2. “UPI Payment Fraud Detection Using Machine Learning,” International Journal of Scientific Research in Engineering and Management (IJSREM), 2025.
3. “UPI Fraud Detection Using Machine Learning,” International Journal for Research Trends and Innovation (IJRTI), 2025.
4. “UPI Transaction Fraud Detection Using Machine Learning – A Data Driven Approach,” IJSREM, 2025.
5. “UPI Fraud Detection Using Machine Learning,” Journal of Emerging Technologies and Innovative Research (JETIR), 2025.
6. “A Review on UPI Fraud Detection Using Machine Learning and Deep Learning,” International Journal for Research in Applied Science and Engineering Technology (IJRASET), 2025.
7. Techniques for UPI and Telecommunication Fraud Detection,” International Journal of Engineering Research & Technology (IJERT), 2025.
8. H. Kumar and H. R. Divakar, “UPI Fraud Detection with Machine Learning Techniques,” IRJMETS, 2025.
9. “Hybrid CNN-SVM Model for Enhanced UPI Fraud Detection,” International Research Journal of Engineering and Technology (IRJET), 2025.

10. "Mobile Payment Fraud Detection in UPI via Systematic Analysis," MR Journal, 2025.
11. M. Nazmoddin, M. Swetha, Y. Gattu, and Y. Divyasree, "UPI Fraud Detection Using Machine Learning," Eudoxus Press, 2024.
12. "UPI Fraud Detection Using Machine Learning," International Journal of Creative Research Thoughts (IJCRT), 2024.
13. "UPIGUARD: Secure Your UPI Transactions Using Machine Learning,"

    JETIR, 2024.

14. Y. Gupta et al., "UPI-Based Financial Fraud Detection Using Deep Learning Approach," IJMSM Conference Proceedings, 2024.
15. G. R. Charan and K. D. Thilak, "Detection of Phishing Link and QR Code of UPI Transaction Using Machine Learning," ICIMIA Conference Proceedings, 2023.
16. M. Ali et al., "Financial Fraud Detection Based on Machine Learning: A Survey," Applied Sciences, vol. 12, no. 19, pp. 9637, 2022.
17. N. Yousefi, M. Alaghband, and I. Garibay, "A Comprehensive Survey on Machine Learning Techniques for Credit Card Fraud Detection," arXiv preprint arXiv:1912.02629, 2019.
18. A. Mukhamadiyev et al., "Intelligent Algorithms for Suspicious Payment Detection," Sensors, vol. 25, no. 21, pp. 6683, 2018.
19. T. Tressa and R. Priya, "Blockchain-Based Secure Payment Systems for Fraud Prevention," Journal of Information Security, 2023.
20. D. Devassy et al., "Federated Learning for Privacy-Preserving Fraud Detection in Financial Systems," IEEE Access, 2025.