International Journal of Science, Strategic Management and Technology
Volume **02** Issue **04** April-**2026** | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

AI-Driven UPI Fraud Detection for Customer Safety

D. Kalpana Devi Dept. of CSE
M.I.E.T Engineering College
(Affiliated to Anna University) Trichy,TamilNadu
kalpanadurairaj12@gmail.com

A.Aneesa Banu Dept. of CSE
M.I.E.T Engineering College
(Affiliated to Anna University) Trichy,
TamilNadu aneesabanu251@gmail.com

M.Geetha Dept. of CSE
M.I.E.T Engineering College
(Affiliated to Anna University) Trichy, TamilNadu
mahadevangeetha786@gmail.com

S.Dhivya Dept. of CSE
M.I.E.T Engineering College
(Affiliated to Anna University) Trichy,
TamilNadu
Dhivyasundarraj06@gmail.com

https://doi.org/10.55041/ijsmt.v2i4.202

Cite this Article: Banu, A., M.Geetha, & S.Dhivya, (2026). AI-Driven UPI Fraud Detection for Customer Safety. International Journal of Science, Strategic Management and Technology, *02*(04). https://doi.org/10.55041/ijsmt.v2i4.202

Abstract— In this project, an artificial intelligence-based technique to detect fraud in Unified Payments Interface (UPI) networks to ensure that users' transactions are protected is described. With the increased popularity of online payments, there is the need to detect such frauds in real-time to minimize any loss and increase the safety of consumers. Data used in this research are obtained and pre-processed to make sure the data does not contain any noise or null values, thus ensuring data accuracy. Various features associated with such transactions are identified and utilized to classify them. Using the identified features, a classifier using random forest algorithm is built and used to distinguish between fraudulent and non-fraudulent transactions. The random forest model enhances the performance of the classifier by using more than one decision tree hence improving prediction accuracy and avoiding overfitting. The proposed method can detect frauds in real- time hence helping to reduce the associated costs.

Keywords— Unified Payment Interface (UPI), Artificial Intelligence (AI), Random Forest, Fraud Detection.

I. INTRODUCTION

With the advent of Unified Payment Interface (UPI) and QR code-based payment system, financial services have become faster, more convenient, and widely accessible. Nonetheless, due to the fast development of payment platforms, there has been a significant rise in cases of fraud, scams, and malicious activity on such websites. The problem lies in the fact that numerous individuals are becoming the victims of attacks based on fake QR code and phishing links without even knowing about it, which leads to financial damages and a lack of trust in payment platforms. Existing rule-based security frameworks can hardly detect the evolution of threats because of the lack of adaptability.

Thus, the present project suggests creating an AI-powered link classifier for detecting potential fraud in UPI and QR code transactions. Based on the analysis of several essential features, such as URL length, transaction time, number of special characters used, domain reliability, and scan frequency, this system allows one to classify links into 'safe' and 'dangerous' categories and gives users opportunity to block the transaction and receive an AI-based voice warning prior to its automatic blocking.

The new architecture will increase the level of security for transactions through a combination of conventional security measures and feature-based classification techniques. This will be achieved by introducing real-time fraud detection and user decision-making, as well as blocking transactions that pose risks to users.

II. LITERATURE REVIEW

A. Fraud Detection in Digital Payments

UPI and QR codes have brought revolutionary changes to banking and finance through their efficiency and convenience. Yet, with such rapid progress, there has been an alarming rise in frauds like QR code scams, phishing, and other digital

payment scams. The traditional systems based on rules cannot cope with the constantly developing methods of fraud. In order to solve the issue, machine learning approaches like Random Forest, XGBoost, and Isolation Forest were introduced to identify abnormal transactions through analyzing transaction metadata, web page information, and behavioral patterns[1]. Although the accuracy can be increased by using an ensemble of models, a lot of papers use artificially generated data without applying it to real-time usage scenarios[2]. These techniques cannot adapt to changing fraud tactics. In order to address these issues, scientists have recommended the use of machine learning algorithms such as Random Forest, XGBoost, and Isolation Forest for detecting anomalies and fraudulent transactions[3]. These algorithms utilize transaction metadata, URL characteristics, and user activity to determine whether the transaction is valid or fraudulent[4]. Although ensemble learning can enhance performance, most of the research studies utilize simulated data and do not deploy the solution in real time[2].

B. Malicious URL and QR Code Detection

Research on malicious URL detection offers lessons that can be applied to QR code fraud prevention strategies. There are studies conducted based on machine learning models, which involve the use of lexical features such as the length of the URL, the number of special characters, and reputation of the domain for link classification purposes[5]. The use of ensemble modeling, such as the random forest algorithm and boosting methods, results in highly accurate detections,[6] and anomaly detection models such as the Isolation Forest model identify suspicious anomalies[7]. Deep learning models such as the LSTM neural network analyze transaction sequences to detect fraud.[8]

C. Application of Banking Fraud Detection Techniques to Digital Payments

Banking and credit card fraud detection systems have been adapted for digital payments as well. Logistic regression and boosting algorithms detect any unusual activity in the transactions[4], whereas random forest and XGBoost examine metadata related to transactions and interaction with merchants[1]. On the other hand, anomaly detection models like Isolation Forest detect any deviation in the normal pattern of spending[3], and the LSTMs detect any sequential activities of transactions[8]. Although useful in banking, using these models for UPI and QR-based payments poses several challenges[2].

D. Research Gaps and Contributions

In spite of advancements in detecting fraud, there are some limitations that still need to be addressed:

- Dataset Constraint – Previous research uses either synthetic or smaller datasets, limiting practical application[2][4].
- Lack of Real-Time Prevention – Currently available systems only consider detection and not prevention (Block/Cancelling Options)[1][3].
- Poor User Interactivity – Present models only offer fixed alerts without incorporating AI-based voice or interactive systems [3].
- Necessity for Learning Adaptation – Fraud detection models frequently require re- training[6].

Contribution of This Project:

This research focuses on the design of an artificial intelligence-powered fraud detection framework based on the Random Forest algorithm for the UPI and QR code payment systems[1]. The framework makes use of features like the length of the URL, timing of transactions, special characters present, trustworthiness of the domain, and scanning frequency to classify the transactions[5][7]. Apart from detecting fraud, this framework has also integrated features that allow users to block or cancel transactions in real-time and provide artificial intelligence-powered voice alerts[3].

III. SYSTEM ARCHITECTURE

The suggested system will ensure improved security for transactions using the Unified Payment Interface and QR codes through the application of machine learning approaches. The system includes several levels which combine their capabilities in order to identify any fraudulent behavior.

A. User Interaction Layer

The layer acts as the interface through which users conduct their transactions via the use of UPI and also scan the QR codes. The layer collects crucial transaction information like sender/receiver information, amount, transaction date/time, device information, and scanned QR code information.

B. Data Collection and Preprocessing

This is when the transaction data gathered is cleaned to enhance its quality by removing noise and inconsistencies. Data normalization is performed next to achieve consistency, making the machine learning algorithm perform better.

C. Feature Extraction

The important features which are extracted from the data to detect the pattern of fraud include the length of the URL, the age of the domain, the presence of HTTPS, the amount of the transaction, the time of transaction, device details, and location details.

These features help in distinguishing between legitimate and fraudulent transactions.

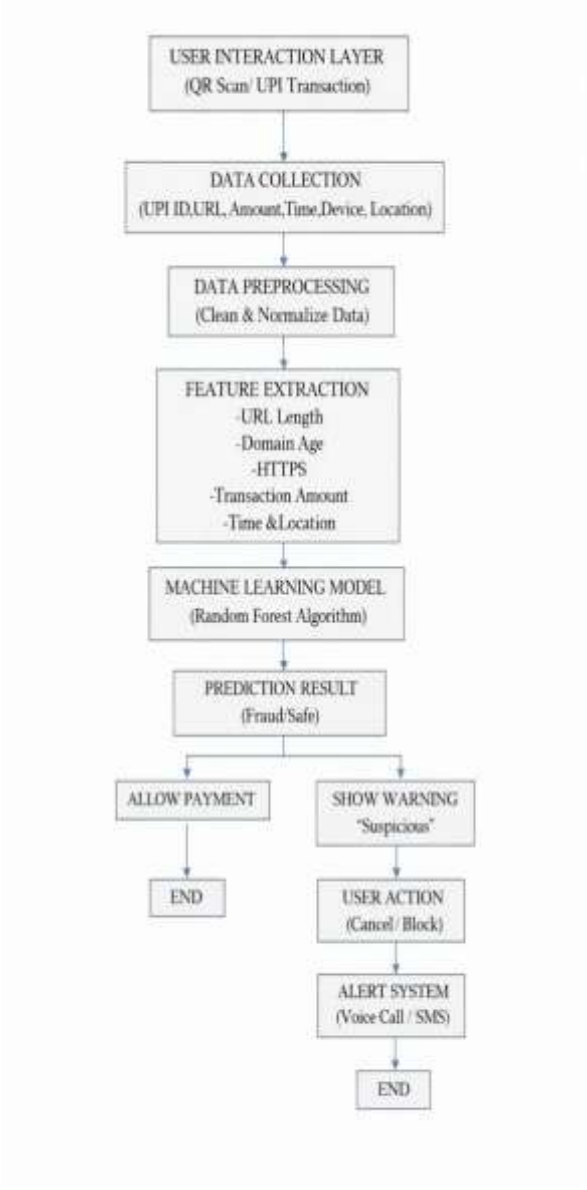

Fig.1 Block schematic of proposed model

D. Machine Learning Model

This process utilizes the Random Forest classifier for analyzing the feature vector. The classifier learns from past transactional data and classifies the transaction as either fraudulent or genuine. It is capable of uncovering patterns and anomalies that are difficult to find using conventional means.

E. Risk Assessment and Prediction

Based on the model output, risk scores are assigned to each transaction to determine whether it is legitimate or fraudulent. This step plays a crucial role in supporting accurate decision-making.

F. Alert and Decision System

According to the results of the model, risk scores are obtained for each transaction. safe or fraudulent This is an essential step in making decisions. If a transaction is detected as fraudulent, the system immediately alerts the user through warning messages. Additional alert mechanisms such as AI-based voice calls or SMS notifications are triggered. The user is given options to cancel or block the transaction. If the transaction is safe, it is allowed to proceed normally.

IV. FRAUD DETECTION MODEL FORMULATION

The fast adoption of UPI payment system transactions and QR-based payment methods has led to several security concerns, such as fraudulent acts and tampering with the QR codes. In order to resolve these problems, this paper presents a machine learning-based system for detecting fraud, which makes it possible to identify potentially suspicious transactions instantly. The system is based on the application of the Random Forest approach.

A. Feature Extraction and Preprocessing

This system first gathers data associated with the transactions, including UPI ID, transaction amount, time stamp, device-related information, geographical information, and QR code-related information. The data is then processed for noise reduction, handling missing values, and normalization of numeric variables. Feature extraction plays a crucial role in identifying fraud patterns:

- UPI Transactions: Behavioral characteristics such as transaction rate, variation in the amount transacted, and association between sender-receiver are considered.

-QR Code Transactions: Structural characteristics like length of embedded URL, domain trust, and integrity of encoded data are studied.

Such characteristics make it possible to differentiate between genuine and fraudulent transactions by detecting anomalies in patterns. Adequate preprocessing is essential for achieving accurate modeling and classification of risky transactions.

B. Fraud Detection Model Selection

The random forest method will be used in this system because of its high level of accuracy and effectiveness. The random forest is a type of ensemble learning model which uses many decision trees.

International Journal of Science, Strategic Management and Technology
Volume **02** Issue **04** April-**2026** | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

The model is capable of identifying transaction patterns and categorizing them into legitimate and fraud cases. This capability enables it to detect frauds that occur in real time in the Unified Payments Interface.

C. Risk Evaluation and Alert Mechanism

It has an in-built risk assessment process that helps determine the level of risk associated with each transaction. Once the transaction starts, a risk score is generated using the features obtained.

- In case the risk score surpasses a certain predefined value, an alert message will be generated informing the user about possible fraud.

- The user will be given two choices, which include either executing the transaction or cancelling it.

- The AI-based voice alert system will be activated before the execution of any high-risk transaction.

The system first checks for any tampering or manipulation of the QR code being read before processing the payment to ensure that there is no potential threat involved in the process.

D. Implementation Considerations

The intended fraud detection framework is meant to serve as a scalable and lightweight mechanism that would be able to interface with the UPI infrastructure currently in use. The proposed system works in real-time mode in order to ensure both timely operation and a high degree of precision.

This mechanism will perform its work independently and will analyze financial transactions separately from the main operations within the payment system itself. The proposed design is flexible, allowing future updates.

V. RESULTS AND ANALYSIS OF THE EXPERIMENT

A. Dataset Description

The proposed AI-powered UPI fraud detection system is developed using two primary datasets. The first dataset contains UPI transaction records, including features such as transaction amount, sender UPI ID, receiver UPI ID, transaction timestamp, and device-related attributes. These features are essential for identifying suspicious transaction patterns and abnormal user behavior.

Fig.2 Logo

The other set of data contains QR code images, both genuine and fake QR code images. Image processing methods like image resizing, normalizing, and feature extraction have been used to enhance the functioning of the model. The set of dataset help the system detect both financial fraud and QR code fraud.

Fig.3 Home Page

International Journal of Science, Strategic Management and Technology
Volume **02** Issue **04** April-**2026** | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

B. Performance Metrics

The fraud detection model's performance will be assessed using conventional metrics for classification:

Accuracy: The accuracy with which predictions are made.

Precision: The number of identified frauds that are actual frauds.

Recall: The model's capacity to detect all the frauds.

F1-Score: An assessment that balances precision and recall.

AUC-ROC: The model's effectiveness in identifying fraudulent and non-fraudulent transactions at various thresholds.

Fig .4 Transaction Status

These metrics ensure a reliable and comprehensive evaluation of the system's effectiveness.

C. Model Evaluation (Random Forest)

The Random Forest method serves as the principal classifier for detecting fraud because of its ability to handle complexity and ensure high precision. The Random Forest classifier can detect complex relationships between transactions.

It is capable of analyzing concealed patterns of user behavior and transaction flows, thus being able to detect fraud. Additionally, using several trees reduces the probability of overfitting.

Fig 5. Pin Authentication

The same algorithm is applied for the verification of the QR codes, where the extracted features will be analyzed and anomalies identified within the structure of the QR codes.

D. Discussion and Observations

The findings of the experiment indicate that the Random Forest model is effective in detecting fraudulent transactions on the UPI platform. The integration of QR code validation enhances system security by identifying counterfeit or tampered QR codes prior to transaction processing. This combined approach ensures reliable performance in both transaction monitoring and QR-based fraud detection. However, challenges related to the availability of real-world fraud data remain a limitation for further improvement.

.

Fig 6. Smart Voice Alert

VI. CONCLUSION

In this paper, an AI-based fraud detection system is presented that aims at providing greater protection to the transactions carried out using the UPI network and QR codes. The application of Random Forest algorithm ensures that any form of fraud is detected successfully.

This model detects fraud in real time and boosts user trust in online payment systems. This combination of techniques makes this suggested solution efficient for preventing any type of financial fraud.

Further studies might be directed towards incorporating real-time data streams, scaling up models, and utilizing advanced AI algorithms for improving accuracy in detection.

ACKNOWLEDGMENT

The authors sincerely thank M.I.E.T. Institutions, Trichy, for providing the research facilities and academic support that enabled the successful completion of this work. Special appreciation is extended to Ms.V.Bhuvanesvari, Research Guide, Department of Computer Science and Engineering, for her valuable guidance, technical insights, and constructive feedback throughout the project. The authors also acknowledge the encouragement from faculty members and peers, as well as the availability of institutional datasets and analytical tools, which greatly contributed to the effective implementation of the proposed UPI fraud detection system.

REFERENCES

[1] R. Ashok Kumar, Shaik Ishrat, Mallela Durga Prasad, Poona Abubakar Siddiq, and N. C. Hari Shankar, "AI-Driven Detection Mechanism for UPI Fraud and QR Code Tampering," Proc. Int. Conf. Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), IEEE, 2025.

[2] Rupa Rani, Adnan Alam, and Abdul Javed, "Machine Learning Driven Fraud Detection System for UPI Transaction," Proc. Int. Conf. Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), IEEE, 2024.

[3] Aditya S. Mandlik, Maitreya S. Ganeshpure, Chaitanya N. Kaddas, Lakshman Korra, Jayaraj U. Kidav, and Manjiri A. Lavadkar, "AI-Driven Real-Time Threat Detection for UPI Transaction," Proc. Int. Conf. Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), IEEE, 2025.

[4] Ankaj Kumar, Kabli Sethi, and Amit Verma, "A Comprehensive Review of Machine Learning Techniques in Fraud Detection," Bhara University Journal of Computer Science, 2025.

[5] Thadukarla Jaswanthi, Talluri Harshitha, Simhadri Tanya, and Meena Belwal, "Malicious URL Detection: Comparative Study of Machine Learning Algorithms," Proc. Int. Conf. Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), IEEE, 2024.

[6] Siddharth Singhal, Utkarsh Chawla, and Rajeev, "Machine Learning and Concept Drift Based Approach for Malicious Website Detection," Proc. Int. Conf. Cyber Security and Digital Forensics, 2020.

[7] Read Bani Hani, Motsame Ammourah, Yazeed Abu Khalil, and Mohamad Swailm, "Malicious URL Detection Using Machine Learning," Jordan University Journal of Information Technology, 2024.

[8] Delano Oscar Do Rosario Lourenco, M. V. H. Sai Sriraj, and Krotha Kayne Thambi, "Malicious URLs and QR Code Classification Using Machine and Deep Learning Techniques," Proc. Int. Conf. Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), IEEE, 2025.