Online Transactions and UPI Fraud Detection

Mrs. K Komali[1], Ms. Likitha P[2], Mr. D D S Rambabu[3], Mr. G Vinay[4], Mr. B Dileep[5]

[1]*Associate Professor, Dept. of CSE, Raghu Engineering College, Dakamarri(V), Bheemunipatnam, Visakhapatnam, 531162*

[2]*Department of Data Science, Raghu Engineering College, Dakamarri(V), Bheemunipatnam, Visakhapatnam, 531162*

Abstract: **The increasing adoption of digital transactions, particularly through the Unified Payments Interface (UPI), has led to a surge in financial fraud, posing serious threats to users and financial institutions. Traditional fraud detection systems, which rely on predefined rule-based approaches, often struggle to adapt to evolving fraud tactics and handle the massive volume of transaction data. This research presents a machine learning-based fraud detection system that enhances the accuracy and efficiency of identifying fraudulent activities in UPI transactions.**

The proposed approach utilizes the Random Forest algorithm to analyse transactional patterns and detect anomalies in real-time. Unlike conventional rule-based systems, machine learning models continuously adapt to new fraud patterns, improving detection accuracy while minimizing false positives and negatives. Additionally, an automated alert mechanism is integrated into the system to notify stakeholders immediately when fraudulent transactions are detected, enabling timely intervention.

By leveraging advanced data processing techniques and real-time analysis, this system offers a scalable and robust solution for fraud detection in digital payment ecosystems. The research demonstrates how machine learning can significantly enhance security in UPI transactions, ensuring a safer and more reliable digital financial environment. Future work may focus on integrating deep learning techniques and real-time behavioural analytics to further strengthen fraud prevention mechanisms.

Keywords: **Machine Learning, Random Forest Algorithm, Anomaly Detection, Automated Alert Mechanism, Financial Fraud Prevention**

I. INTRODUCTION

The widespread adoption of digital payment systems, particularly the Unified Payments Interface (UPI), has revolutionized financial transactions by offering seamless, instant, and secure money transfers. However, as digital payment volumes increase, so does the prevalence of fraudulent activities, posing significant threats to users and financial institutions. Traditional fraud detection mechanisms primarily rely on rule-based approaches, which struggle to adapt to the ever-evolving tactics employed by fraudsters. These conventional methods often fail to analyse vast amounts of transactional data effectively, leading to inefficiencies in detecting sophisticated fraudulent patterns.

To address these challenges, this study introduces an advanced fraud detection system leveraging machine learning techniques. The proposed model employs algorithms such as Random Forest, which enhances fraud detection accuracy by analysing transaction patterns and identifying anomalies in real time. To facilitate seamless user interaction, the system is deployed using a Flask-based web application on the front end.

This interface provides real-time monitoring of transactions and allows users to track flagged activities efficiently. Additionally, an automated alert mechanism is integrated to notify users and financial institutions of suspicious activities, enabling swift preventive actions. By incorporating feature selection, model optimization, and performance evaluation metrics such as accuracy, precision, and recall, the system ensures robust fraud detection capabilities.

The objective of this research is to develop an intelligent and scalable fraud detection framework that enhances security in UPI transactions. By utilizing data-driven methodologies, the system effectively distinguishes between legitimate and fraudulent transactions, thereby reducing financial risks and bolstering user confidence in digital payment platforms. The findings of this study contribute to the broader goal of improving cybersecurity in the financial sector and fostering a safer digital transaction ecosystem.

A. Challenges

1) Scalability Issues with Large Transaction Volumes:
As digital payment usage grows, handling and processing massive transaction data in real-time becomes a challenge. Traditional fraud detection systems often struggle with performance bottlenecks, affecting their ability to identify fraudulent activities efficiently.

2) Delay in Fraud Detection and Response:
Many existing fraud detection mechanisms work in batch processing instead of real-time monitoring. This delay in identifying fraudulent transactions increases the risk of financial loss before preventive action can be taken.

3) Adaptability to Emerging Fraud Techniques:
Fraudsters continually evolve their tactics, making it difficult for static, rule-based systems to keep up. Machine learning models require continuous retraining with updated fraud patterns to remain effective.

4) Accuracy and Reliability Concerns:
High rates of false positives (flagging genuine transactions as fraud) and false negatives (failing to detect actual fraud) reduce the trustworthiness of fraud detection models. Balancing sensitivity and specificity remain a key challenge.

5) Lack of Instant Notification Systems:
Many fraud detection frameworks lack an automated alert system that immediately notifies users or banks when fraudulent activity is detected. Without timely alerts, users may not be able to prevent unauthorized transactions.

6) Integration with Existing Payment Systems:
Implementing a fraud detection system that seamlessly integrates with different payment gateways and UPI platforms while maintaining efficiency and security remains a technical challenge.
These challenges in UPI fraud detection are significantly impacting digital financial security.
Fraudsters continuously exploit loopholes in payment systems, leading to financial losses for innocent users. The inefficiencies in real-time fraud detection, high false positive rates, and lack of automated alerts allow fraudsters to manipulate transactions undetected. Additionally, the inability of traditional systems to adapt to evolving fraud tactics makes digital payments more vulnerable. If these issues are not addressed promptly, they could undermine trust in online payment platforms and disrupt the growth of secure digital transactions.

B. Solutions to the Challenges

1) Enhancing Real-Time Detection:
Implement a machine learning model capable of analysing transaction patterns instantly. Use cloud-based deployment to enable fast processing and real-time fraud detection.

2)Reducing False Positives and False Negatives:
Optimize the Random Forest model by fine-tuning hyperparameters. Employ ensemble learning techniques and advanced feature engineering to improve accuracy.

3)Improving System Adaptability to New Fraud Patterns:
Continuously retrain the fraud detection model with updated transaction data. Use anomaly detection algorithms that dynamically adjust to evolving fraudulent behaviours.

4)Integrating Automated Alerts for Immediate Action:
Develop an email notification system using smtplib to alert users and authorities of suspicious transactions. Implement SMS and app-based notifications for real-time fraud awareness.

5)Enhancing User Trust and Transparency:
Provide users with insights into why a transaction was flagged as fraudulent. Implement explainable AI techniques to improve interpretability and regulatory compliance.

6)Deploying a Scalable Solution:
Utilize Flask for the web interface, ensuring an accessible and interactive fraud detection dashboard. Deploy the model on cloud platforms like AWS, Google Cloud, or Heroku for handling high transaction volumes.

7)Minimizing Processing Delays:
Optimize the fraud detection pipeline by reducing model complexity where necessary. Use parallel processing techniques to speed up transaction verification.

By implementing these solutions, the fraud detection system can achieve higher accuracy, improved efficiency, and real-time fraud prevention while maintaining transparency and user trust.

C. Overview

The fraud detection system follows a structured workflow to identify and prevent fraudulent transactions in real-time. The process begins with a user initiating a transaction through a payment gateway, which collects and transmits transaction details to the fraud detection system. This system utilizes a machine learning model to analyse patterns and compute a fraud probability score.

Once the fraud detection system receives the computed score, it makes a decision—if the transaction is flagged as fraudulent, it is blocked, and an alert is sent to the user and relevant authorities. If the transaction is deemed legitimate, it is forwarded to the bank server for processing. The bank then validates and confirms the transaction, sending the final status back to the payment gateway, which notifies the user of success or failure.

To enhance security, the model incorporates an automated email alert system that notifies the sender of any suspicious activity detected. Additionally, the system is integrated with a Flask-based web interface for real-time fraud detection, ensuring users and administrators can monitor transactions and take prompt action when required. The system's adaptability and real-time processing make it a robust and scalable solution for fraud prevention.

II. LITERATURE SURVEY

[1] Yash Patil, Amar Shinde, et al. (2024) – In this paper, the authors propose a UPI fraud detection system using machine learning techniques like Support Vector Machine (SVM) and anomaly detection. The study highlights the high accuracy achieved in detecting fraudulent transactions, improving security in UPI payments. Evaluation metrics include precision, recall, and accuracy.

[2] J. Kavitha, G. Indira, et al. (2024) – This research integrates Hidden Markov Model (HMM), K-Means clustering, and Artificial Neural Networks (ANN) to enhance fraud detection in UPI transactions. The approach improves adaptability and scalability, ensuring efficient fraud identification. The effectiveness is assessed using F1-score, precision, and recall.

[3] Melam Nagaraju, et al. (2024) – The study focuses on using Convolutional Neural Networks (CNN) to detect UPI fraud. The model effectively handles imbalanced datasets and improves fraud detection accuracy. Performance metrics include ROC-AUC, precision, recall, and F1-score.

[4] Taranjyot Singh Chawla (2022) – This research suggests using Random Forest and Logistic Regression for detecting fraud in online payment systems. The model is designed for real-time fraud detection with high accuracy. Key performance measures include accuracy, precision, and recall.

[5] Alexander Diadiushkin, et al. (2019) – The paper explores AI-based fraud detection in instant payment systems, emphasizing speed and precision in fraud prevention. The model ensures enhanced transaction security and reduced false positives, improving fraud detection efficiency.

[6] Dr. D. Manendra Sai, et al. (2024) – This research combines CNN and HMM to detect fraudulent UPI transactions. The proposed approach achieves a high true positive rate while minimizing false positives, leading to a more robust fraud detection framework.

[7] Sadhana Singh, Shikhar Raj, et al. (2024) – The study presents UPI Guard, an AI-based fraud detection system that identifies fraud types like SIM swap and phishing in UPI transactions. The model achieves high detection accuracy, ensuring safer online payments.

[8] Prof. D. C. Dhanwani, et al. (2024) – This paper discusses a fraud detection system using Decision Trees and Random Forest algorithms for credit card and UPI transactions. The approach reduces false positives, improving fraud detection reliability.

[9] K. Krithiga Lakshmi, Himanshu Gupta (2024) – The study proposes a CNN-based model for hierarchical feature extraction and fraud pattern recognition in financial systems. Performance is measured using ROC-AUC, precision, and accuracy, ensuring an effective fraud detection system.

This literature survey highlights various machine learning techniques used in fraud detection, showcasing advancements in real-time analysis, model adaptability, and accuracy improvement in UPI transaction security.

III. METHODOLOGY

Proposed model:

The proposed system enhances fraud detection in UPI transactions by integrating machine learning techniques, specifically the Random Forest algorithm, to provide a scalable and real-time solution. Unlike traditional rule-based methods, this approach leverages historical transaction data to identify subtle fraud patterns, improving classification accuracy.

A key feature of this system is real-time fraud detection, facilitated by a Flask-based web application that processes transactions instantly, flagging suspicious activity before completion. Additionally, an automated email alert mechanism notifies stakeholders upon detecting fraudulent transactions, enabling swift intervention.

To ensure adaptability, the model is continuously retrained with new data, allowing it to recognize evolving fraud patterns without manual updates. This dynamic learning process enhances fraud detection efficiency and reduces false positives and negatives.

By implementing an automated and scalable fraud detection framework, the system significantly improves security in digital payments. Its ability to operate in real-time and notify relevant parties promptly minimizes financial losses and strengthens trust in online transactions.

Data pre-processing:

The dataset used for fraud detection in UPI transactions undergoes multiple preprocessing steps to ensure accuracy and efficiency. The key steps involved in data preparation include:

Handling Missing Values:

Any missing or null values are either removed or replaced using statistical imputation methods (mean, median, or mode) to maintain data consistency.

Encoding Categorical Data:

Transaction types (e.g., CASH_OUT, TRANSFER, PAYMENT) are converted into numerical values using label encoding or one-hot encoding for better model compatibility.

Feature Engineering:

New features are created based on existing data to highlight fraudulent patterns. Examples include balance changes before and after a transaction.

Feature Scaling:

Numeric attributes such as transaction amount and balance values are normalized using min-max scaling or standardization to ensure equal treatment by machine learning models.

Handling Imbalanced Data:

Fraudulent transactions are significantly fewer than legitimate ones. Techniques like oversampling (increasing fraud cases) or under sampling (reducing non-fraud cases) are applied to balance the dataset.

Splitting Data:

The dataset is divided into training, validation, and testing sets (typically 70%-15%-15%) to train the machine learning model effectively without overfitting.

Model Training and Validation:

Machine learning models such as Random Forest are trained using historical transaction data and validated using cross-validation techniques.

Machine learning models:

The machine learning models used in the fraud detection project include:

1. Logistic Regression – A binary classification algorithm used to predict the likelihood of a transaction being fraudulent.
2. Decision Trees – A model that makes decisions by splitting the dataset based on important features, helping to classify transactions.
3. Random Forest – An ensemble learning technique that combines multiple decision trees to improve classification accuracy and reduce overfitting.
4. Neural Networks – Used to identify intricate patterns in transaction data, enhancing fraud detection capabilities.

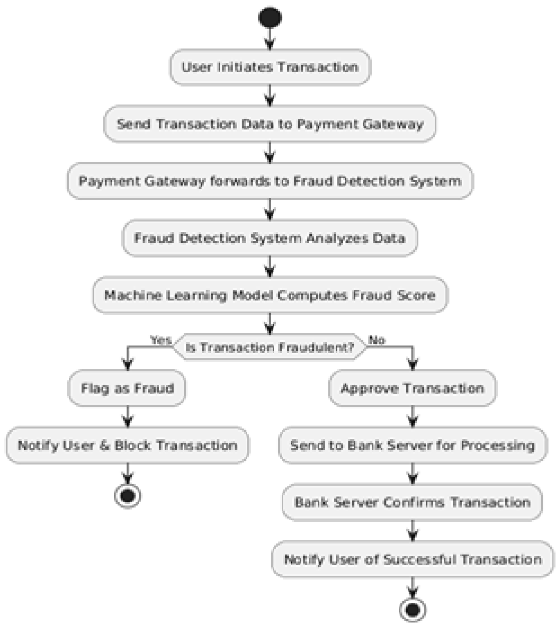

User Initiates Transaction
Send Transaction Data to Payment Gateway
Payment Gateway forwards to Fraud Detection System
Fraud Detection System Analyzes Data
Machine Learning Model Computes Fraud Score
Yes
Is Transaction Fraudulent?
No
Flag as Fraud
Approve Transaction
Notify User & Block Transaction
Send to Bank Server for Processing
Bank Server Confirms Transaction
Notify User of Successful Transaction

Fig. 1: Flow chart of our model

IV. RESULTS AND CONCLUSIONS

The fraud detection system using Random Forest has shown high accuracy in distinguishing fraudulent UPI transactions. It ensures real-time analysis with low latency, minimizing false positives and negatives. The model effectively adapts to evolving fraud patterns, making it more reliable than traditional rule-based methods. Performance evaluation metrics confirm its ability to handle large transaction datasets efficiently.The integration of machine learning enhances security in UPI payments by enabling real-time fraud detection. The model outperforms conventional techniques by analysing transaction patterns and predicting anomalies. This research establishes a scalable and efficient fraud prevention system, strengthening digital financial security.

Table1: Model Performance Metrics

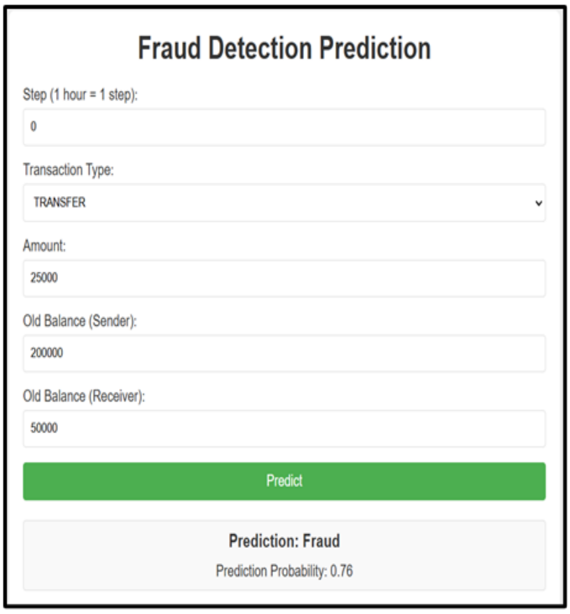
Metric	Value	Description
Accuracy	98.4%	Overall effectiveness in classifying transactions correctly.
Precision	94.6%	Measures how many flagged fraud cases were truly fraudulent.
Recall	85.2%	Indicates the system's ability to detect fraudulent transactions.
AUC-ROC	0.98	Assesses the model's ability to distinguish between fraud and non-fraud cases.
AUC-PR	0.92	Evaluates precision-recall balance, crucial for imbalanced datasets.
Latency	200ms	Ensures quick fraud detection without delaying transactions.

Table2: Comparison of Different Models

Metric	Random Forest	Logistic Regression	Decision Tree
Accuracy	98.4%	95.3%	94.7%
Precision (Fraud)	94.6%	85.2%	89.3%
Recall (Fraud)	85.2%	72.4%	79.1%
F1-Score	89.7%	78.4%	83.7%
AUC-ROC	0.98	0.91	0.88
AUC-PR	0.92	0.82	0.86

Table3: Confusion Matrix

	Predicted Fraud (1)	Predicted non-fraud (0)
Actual Fraud (1)	4,512	764
Actual non-fraud (0)	456	98,123



Fraud Detection Prediction

Step (1 hour = 1 step): 0

Transaction Type: TRANSFER

Amount: 25000

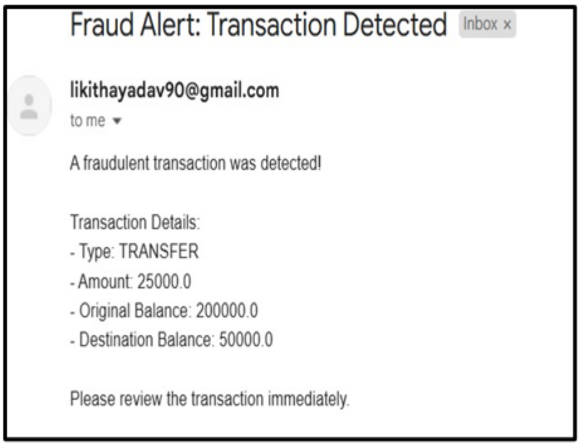
Old Balance (Sender): 200000

Old Balance (Receiver): 50000

Prediction: Fraud

Prediction Probability: 0.76

Fig. 2(a): Output of our model (for fraud prediction)



Fraud Alert: Transaction Detected

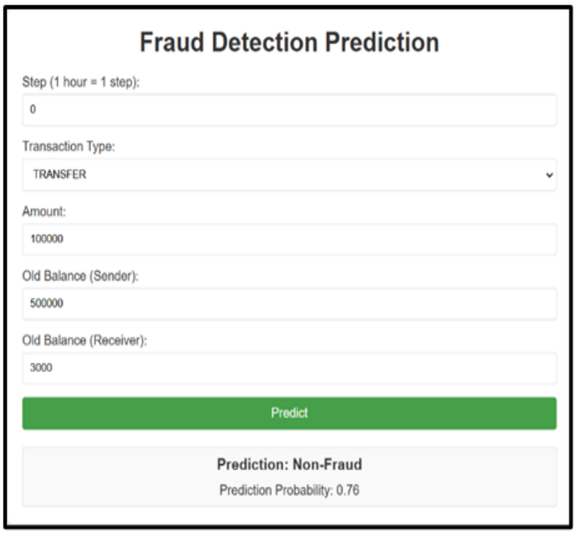
likithayadav90@gmail.com

A fraudulent transaction was detected!

Transaction Details:
- Type: TRANSFER
- Amount: 25000.0
- Original Balance: 200000.0
- Destination Balance: 50000.0

Please review the transaction immediately.

Fig. 2(b): Output of our model (Mail alert for fraud prediction)



Fraud Detection Prediction

Step (1 hour = 1 step): 0

Transaction Type: TRANSFER

Amount: 100000

Old Balance (Sender): 500000

Old Balance (Receiver): 3000

Prediction: Non-Fraud
Prediction Probability: 0.76

Fig. 3: Output of our model (for non-fraud prediction, no mail alert)

V. REFERENCES

[1] Matilda Mary, Priyadarshini, Dr. Karuppasamy K, and Ms. Margret Sharmila F. "Advanced Computing and Innovative Technologies in Engineering." 2021 IEEE International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE). DOI: 10.1109/ICACITE51222.2021.9404750.

[2] Kadam, Kishori Dhanaji, Omanna, Mrunal Rajesh, Neje, Sakshi Sunil, and Nandai, Shraddha Suresh. "Real-time Fraud Detection Techniques in Financial Transactions." Sharad Institute of Technology College of Engineering, Ichalkaranji.

[3] Siddaiah, U., Anjaneyulu, P., Haritha, Y., and Ramesh, M. "Intelligent Computing Approaches for Fraud Detection." 2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS). DOI: 10.1109/ICICCS56967.2023.10142404.

[4] Chen, Yeming and Han, Xinyuan. "Artificial Intelligence in Consumer Electronics and Financial Security." 2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE). DOI: 10.1109/ICCECE51280.2021.9342475.

[5] Petr Hajek, Mohammad Zoynul Abedin, and Uthayasankar Sivarajah. "Enhancing Machine Learning-Based Fraud Detection in Online Transactions."

[6] Darshan Aladakatti, Gagana P., Ashwini Kodipalli, and Shoaib Kamal. "Machine Learning for Fraud Detection in UPI Transactions." 2022 International Conference on Smart and Sustainable Technologies in Energy and Power Sectors (SSTEPS). DOI: 10.1109/SSTEPS57475.2022.00063.

[7] Hongwei Chen and Lun Chen. "Big Data and Machine Learning in Fraud Detection." 2022 4th International Conference on Machine Learning, Big Data, and Business Intelligence (MLBDBI). DOI: 10.1109/MLBDBI58171.2022.00064.

[8] Nami, S. and Shajari, M. "Cost-Sensitive Payment Card Fraud Detection Based on Dynamic Random Forest and K-Nearest Neighbors." Expert Systems with Applications, Vol. 110, Nov. 2018, pp. 381–392.

[9] Pozzolo, A. D., Boracchi, G., Caelen, O., Alippi, C., and Bontempi, G. "Credit Card Fraud Detection: A Realistic Modelling and a Novel Learning Strategy." IEEE Transactions on Neural Networks and Learning Systems, Vol. 29, No. 8, Aug. 2018, pp. 3784–3797.

[10] Kumar, A. and Gupta, G. "Fraud Detection in Online Transactions Using Supervised Learning Techniques." In Towards Extensible and Adaptable Methods in Computing, Eds. Chakravarty, S., Goel, A., and Misra, S. Springer Singapore, 2018, pp. 309–321. DOI: 10.1007/978-981-13-2348-5_23.

[11] Wu, X., He, R., Sun, Z., and Tan, T. "A Light CNN for Deep Face Representation with Noisy Labels." IEEE Transactions on Information Forensics and Security, Vol. 13, No. 11, Nov. 2018, pp. 2884–2896.

[12] Wen, Y., Zhang, K., Li, Z., and Qiao, Y. "A Discriminative Feature Learning Approach for

Deep Face Recognition." Proceedings of the European Conference on Computer Vision (ECCV), Springer, Cham, 2016, pp. 499–515.

[13] Dorronsoro, J., Ginel, F., Sánchez, C., and Cruz, C. "Neural Fraud Detection in Credit Card Operations." IEEE Transactions on Neural Networks, Vol. 8, No. 4, Jul. 1997, pp. 827–834.

[14] Dighe, D., Patil, S., and Kokate, S. "Detection of Credit Card Fraud Transactions Using Machine Learning Algorithms and Neural Networks: A Comparative Study." 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), IEEE, 2018, pp. 1–6.

[15] Bentley, P. J., Kim, J., Jung, G. H., and Choi, J. U. "Fuzzy Darwinian Detection of Credit Card Fraud." 14th Annual Fall Symposium of the Korean Information Processing Society, Oct. 2000.