

Research Paper

UPI Fraud Transaction Detection using Machine Learning

Shoeb Khan¹, Mohammed Abbas Qureshi², Subramanian K.M³

¹PG Scholar, Department of Computer Science and Engineering.

Shadan College of Engineering and Technology, Hyderabad, Telangana, India – 500086

Email: shoeb333m@gmail.com

²Associate Professor, Department of Computer Science and Engineering.

Shadan College of Engineering and Technology, Hyderabad, Telangana, India – 500086

Email: mohammedabbas09@yahoo.com

³Professor, Department of Computer Science and Engineering.

Shadan College of Engineering and Technology, Hyderabad, Telangana, India – 500086

Email: kmsubbu.phd@gmail.com

ABSTRACT

The rapid growth of digital payment systems and Unified Payments Interface (UPI) transactions has significantly transformed financial transactions by providing convenience, speed, and accessibility. However, this widespread adoption has also increased the risk of fraudulent activities, leading to substantial financial losses for users, banks, and payment service providers. To address this challenge, the proposed UPI Fraud Transaction Detection using Machine Learning system introduces an intelligent framework capable of identifying fraudulent transactions with high accuracy

and efficiency. The system collects and preprocesses transaction data by handling missing values, converting categorical attributes into numerical formats, and normalizing the dataset to ensure data quality and consistency.

Keywords

UPI Fraud Detection, Machine Learning, Digital Payments, Fraudulent Transactions, Principal Component Analysis (PCA), Data Preprocessing, Feature Selection, Transaction Security, Financial Fraud, Predictive Analytics.

I. INTRODUCTION

The advancement of digital technology has revolutionized the financial sector by introducing fast, secure, and convenient online payment systems. One of the most widely adopted digital payment platforms in India is the Unified Payments Interface (UPI), which enables users to transfer money instantly between bank accounts using mobile devices. The increasing popularity of UPI has simplified financial transactions, reduced dependency on cash, and promoted a cashless economy [1], [2].

Despite its numerous advantages, the rapid growth of digital payment systems has also increased the risk of fraudulent activities. Cybercriminals continuously develop new techniques such as phishing attacks, identity theft, fake payment requests, social engineering, and unauthorized account access to exploit users and financial institutions. These fraudulent transactions result in substantial financial losses and create security concerns among users and service providers. Therefore, developing an efficient fraud detection mechanism has become essential for ensuring the safety

and reliability of digital payment systems [15], [16], [20], [21].

Traditional fraud detection methods mainly rely on predefined rules and manual monitoring processes. However, these approaches often fail to identify complex and evolving fraud patterns due to the increasing volume and diversity of online transactions. Machine Learning (ML) provides an effective solution by automatically learning transaction behaviours and identifying suspicious activities based on historical data. Machine learning algorithms can analyze large volumes of transaction data, detect hidden patterns, and accurately classify transactions as genuine or fraudulent [6], [11], [12], [18].

II. LITERATURE SURVEY

The rapid growth of Unified Payments Interface (UPI) transactions has increased the need for intelligent fraud detection systems. Researchers have proposed various Machine Learning (ML) techniques to improve fraud detection accuracy, reduce financial losses, and enhance transaction security [1], [2], [20], [21].

Nazmoddin et al. developed a Machine Learning-based UPI fraud detection system that effectively classifies fraudulent transactions using transaction features [1].

Rani et al. proposed a secure ML-driven fraud detection framework that improves transaction security through intelligent classification [2]. Dave and Chudasama demonstrated that Ensemble Learning and Neural Networks improve fraud detection performance compared to conventional methods [3], while Verma and Tyagi applied XGBoost-based outlier detection to identify suspicious UPI transactions with improved accuracy [4]. Ramya et al. further enhanced fraud detection by integrating Explainable AI using SHAP with XGBoost, improving both prediction accuracy and model interpretability [5], [8].

XGBoost, introduced by Chen and Guestrin, has become one of the most effective ensemble learning algorithms due to its high accuracy and computational efficiency [6]. Géron and Raschka et al. discussed practical Machine Learning techniques, including data preprocessing, feature engineering, and model evaluation using Scikit-learn, which are widely adopted in fraud detection applications [7], [18]. Dal Pozzolo et al. highlighted the importance of handling imbalanced datasets, while Aggarwal emphasized anomaly detection techniques for identifying abnormal transaction patterns [9], [10].

Several researchers have focused on improving fraud detection using advanced Machine Learning and Deep Learning approaches. Chen et al. presented a survey on Machine Learning techniques for Big Data analytics [11], while Sethi et al. proposed a Random Forest-based UPI fraud detection model with high classification accuracy [12]. Chollet, LeCun et al., and Chakka and Saheb demonstrated that Deep Learning and hybrid Autoencoder-XGBoost models further improve fraud detection performance for complex financial datasets [13], [14], [24].

Recent studies have emphasized Explainable AI, cloud computing, and real-time fraud detection for secure financial systems. Almalki and Masud, Fajar et al., and Uddin and Aziz proposed explainable fraud detection models using ensemble learning and SHAP-based techniques [15], [16], [17]. Zaharia et al. introduced Apache Spark for large-scale data processing [19], while Wang et al. and Singh and Sharma presented real-time and cloud-based fraud detection frameworks [20], [21]. Molnar and Brownlee further discussed interpretable Machine Learning and practical Python implementations for developing reliable fraud detection systems [22], [23]. Recent research has also explored hybrid Machine Learning models, explainable risk scoring, and AI-based

fraud analytics to further enhance the security and scalability of UPI transaction systems [25]–[30].

III. EXISTING SYSTEM

Existing UPI fraud detection systems primarily rely on rule-based techniques and conventional Machine Learning algorithms to identify suspicious transactions. These systems monitor transaction behaviour using predefined rules such as transaction limits, unusual login locations, multiple transaction attempts, and account activity patterns. Although these methods are simple to implement, they often fail to detect newly emerging fraud techniques and complex transaction patterns [1], [2], [11].

Several existing approaches employ algorithms such as Naïve Bayes, Decision Trees, Logistic Regression, and Random Forest for fraud classification. While these models provide satisfactory performance on historical datasets, their effectiveness decreases when handling highly imbalanced data and evolving fraud patterns, leading to increased false positives and reduced prediction accuracy [9], [10], [12], [18].

Furthermore, many traditional fraud detection systems lack real-time monitoring, explainability, and scalability required for modern digital payment platforms. These limitations highlight the need for intelligent, adaptive, and interpretable Machine Learning models capable of improving fraud detection performance while reducing computational complexity [15], [20], [21], [22].

IV. PROPOSED SYSTEM

The proposed system introduces an intelligent UPI Fraud Transaction Detection framework using Machine Learning techniques to accurately identify fraudulent transactions. It overcomes the limitations of traditional rule-based systems by employing advanced data preprocessing, feature selection, and classification methods to improve fraud detection accuracy and reduce false predictions [1], [2], [6], [25], [26].

Initially, the transaction dataset is collected and preprocessed by handling missing values, encoding categorical attributes, and normalizing the data to improve model performance. Principal Component Analysis (PCA) is then applied to select the most relevant features, reducing dimensionality and computational

complexity while preserving important transaction information [7], [9], [18], [23].

The processed dataset is divided into training and testing datasets for model development. The proposed system employs the XGBoost algorithm, an efficient ensemble learning technique capable of handling large datasets and detecting complex fraud patterns with high accuracy. The trained model classifies transactions as either genuine or fraudulent based on learned transaction behaviour [6], [14], [27], [28].

To evaluate the effectiveness of the proposed model, performance metrics such as Accuracy, Precision, Recall, F1-Score, and Confusion Matrix are used. The system is designed to support accurate fraud detection, improved interpretability, and future integration with real-time payment applications and intelligent financial security solutions [5], [8], [15], [16], [17], [29], [30].

TABLE 1: TABLE OF COMPARISON OF METHODS AND DATASETS

S.No	Author / Year	Method / Algorithm Used	Dataset Used	Advantages	Limitations
1	Nazmoddin et al. (2024)	Machine Learning Classification	UPI Transaction Dataset	Effective fraud classification and reduced manual monitoring	Performance depends on quality of training data
2	Rani et al. (2024)	Machine Learning-Based UPI Fraud Detection	UPI Transaction Dataset	Improved transaction security and fraud detection accuracy	Requires continuous model updates
3	Dave and Chudasama (2025)	Ensemble Learning and Neural Networks	UPI Fraud Dataset	Higher prediction accuracy and robustness	Increased computational complexity
4	Verma and Tyagi (2025)	XGBoost Outlier Detection	UPI Transaction Dataset	Efficient anomaly detection and fraud identification	Sensitive to parameter tuning
5	Ramya et al. (2026)	XGBoost with SHAP Explainability	UPI Transaction Dataset	High accuracy with model interpretability	Additional computational overhead for explanations

TABLE 2: DATASET COMPARISON TABLE

S.No	Dataset Name	Number of Records	Detailed Fraud Type	Advantages	Limitations
1	Credit Card Fraud Dataset	284,807	Unauthorized credit card transactions, stolen card usage, card-not-present (CNP) fraud, counterfeit card fraud, transaction amount manipulation	Widely used benchmark dataset	Highly imbalanced data
2	Bank Transaction Dataset	150,000+	Account takeover fraud, unauthorized fund transfers, money laundering transactions, identity theft, suspicious banking activities	Real banking transaction patterns	Limited public availability
3	PayPal Transaction Dataset	100,000+	Fake merchant transactions, account hijacking, refund fraud, payment reversal fraud, phishing-based payment fraud	Useful for e-commerce fraud detection	Proprietary dataset
4	UPI Fraud Transaction Dase	63,636	UPI fraud, QR code scams, fake payment confirmation fraud, mobile wallet fraud, SIM swap attacks, phishing attacks	Supports mobile transaction analysis	Limited fraud samples
5	Financial Transaction Dataset	200,000+	Transaction laundering, synthetic identity fraud, fraudulent transfers, money mule activities, suspicious financial transactions	Large transaction volume	High preprocessing requirements

V. METHODOLOGY

The methodology of the UPI Fraud Transaction Detection using Machine Learning project consists of a sequence of processes designed to accurately identify fraudulent UPI transactions. The methodology begins with collecting the UPI fraud transaction dataset, which contains information about both genuine and fraudulent transactions. The collected dataset serves as the foundation for training and testing machine learning models. Since raw data often contains missing values, inconsistencies, and non-numeric attributes, preprocessing is performed to improve data quality and prepare the dataset for further analysis.

During the preprocessing stage, missing values are handled appropriately, and categorical attributes are converted into numerical values using label encoding techniques. The dataset is then normalized and shuffled to ensure consistency and eliminate bias during model training. Data preprocessing is an essential step because the quality of input data directly affects the performance and accuracy of machine learning algorithms. A clean and structured dataset enables the models to learn

transaction patterns more effectively and generate reliable predictions.

After preprocessing, Principal Component Analysis (PCA) is applied for feature selection. PCA is a dimensionality reduction technique that identifies the most important features within the dataset while eliminating redundant and irrelevant attributes. By reducing the number of features, PCA decreases computational complexity and improves the efficiency of machine learning models. The selected features retain the most significant information required for fraud detection, resulting in faster training and improved prediction performance.

VI. IMPLEMENTATION

The implementation phase of the UPI Fraud Transaction Detection using Machine Learning project focuses on developing a complete system capable of detecting fraudulent UPI transactions with high accuracy. The implementation integrates data preprocessing techniques, feature selection methods, machine learning algorithms, performance evaluation mechanisms, and a graphical user interface into a single application. The primary objective of the implementation is to

provide an automated solution that can efficiently identify fraudulent transactions and assist financial institutions in preventing financial losses.

After loading the dataset, the preprocessing module is executed. In this stage, missing values are identified and replaced with appropriate values to ensure data consistency. Since machine learning algorithms require numerical inputs, categorical attributes are converted into numerical values using Label Encoding techniques. The dataset is then normalized and shuffled to remove bias and improve model performance. These preprocessing operations enhance data quality and prepare the dataset for effective machine learning analysis.

VII. SYSTEM ARCHITECTURE

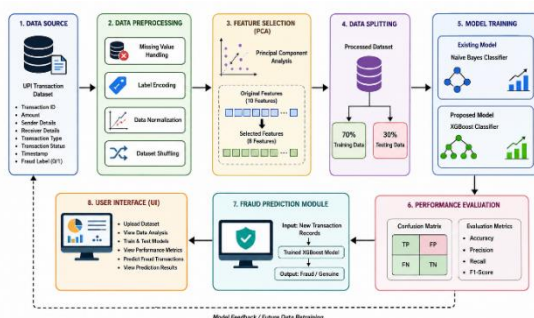


Fig. 7.1: System Architecture

VIII. RESULTS AND DISCUSSION

8.1: NEW USER REGISTRATION PAGE

The New User Registration page enables first-time users to create an account in the proposed system. Users are required to enter essential details such as Username, Password, Contact Number, Email ID, and Address. This interface ensures that all mandatory fields are completed before the registration process is initiated. The design is simple and user-friendly, allowing users to register quickly and securely.

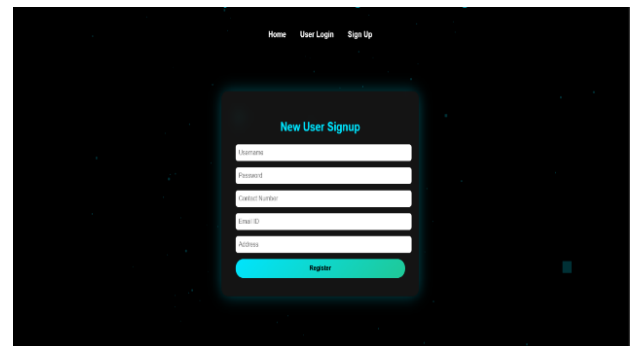


Fig. 8.1: New User Registration (Sign-Up) Interface

8.2: USER REGISTRATION FORM

The above figure illustrates the registration form used for collecting user information. After entering valid details and clicking the Register button, the system validates the input data and stores the user information securely in the database. Once the registration is successful, the user can log in using the registered credentials and access the system's services.

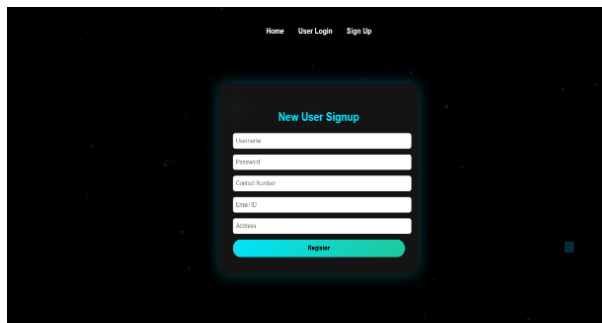


Fig. 8.2: User Registration Form with Input Fields

8.3: ADMIN DASHBOARD INTERFACE

The Admin Dashboard serves as the main interface of the Online Fraud Payment Detection System. It provides navigation to key functionalities, including loading the dataset, balancing the data, training machine learning models, fraud detection, viewing manuals, and accessing the prediction history. This dashboard enables administrators to efficiently manage the complete fraud detection workflow through a centralized interface.

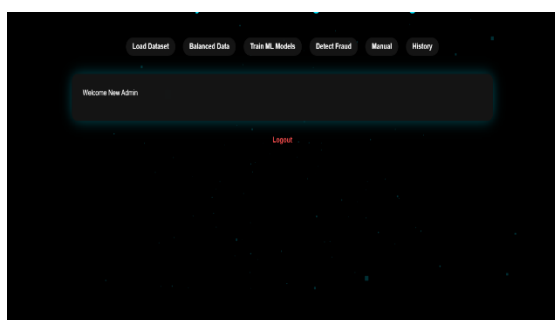


Fig. 8.3: Admin Dashboard of the Online Fraud Payment Detection System.

8.4: DATASET LOADING AND CLASS DISTRIBUTION ANALYSIS



Fig. 8.4: Dataset Loading and Class Distribution Visualization.

This interface displays the successful loading of the Online Fraud Payment Detection dataset. After loading the dataset, the system generates a class distribution graph illustrating the imbalance between fraudulent and non-fraudulent transactions.

8.5: BALANCED DATASET AFTER APPLYING SMOTE

This module presents the dataset after applying the Synthetic Minority Over-sampling Technique (SMOTE) to address the class imbalance problem. The balancing process increases the number of fraudulent transaction samples, resulting in an equal distribution of fraud and non-fraud records. It balances the dataset for improvement.

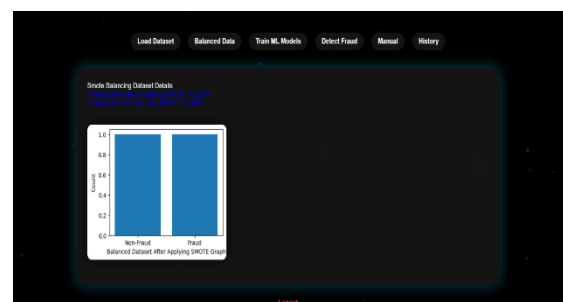
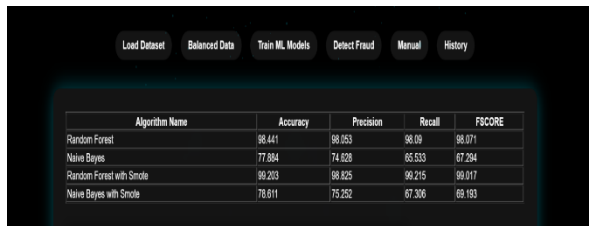


Fig. 8.5: Balanced Dataset Visualization After Applying SMOTE

8.6: MACHINE LEARNING MODEL EVALUATION RESULTS

This module displays the performance metrics of all trained machine learning algorithms used for fraud detection. The evaluation includes Accuracy, Precision



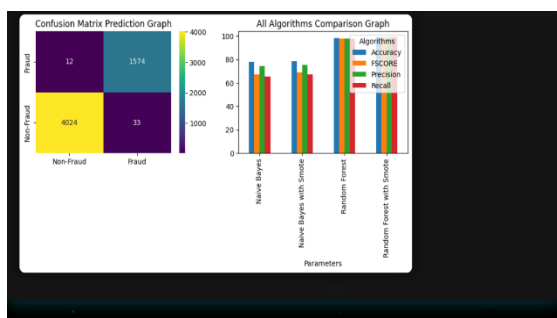
Algorithm Name	Accuracy	Precision	Recall	FSCORE
Random Forest	98.441	98.053	98.09	98.071
Naive Bayes	77.884	74.628	65.533	67.294
Random Forest with Smote	99.203	98.825	99.215	99.017
Naive Bayes with Smote	78.611	75.252	67.306	69.193

Recall, and F1-Score for Random Forest, Naïve Bayes, Random Forest with SMOTE, and Naïve Bayes with SMOTE.

Fig. 8.6: Performance Evaluation of Machine Learning Models.

8.7: MODEL PERFORMANCE ANALYSIS

This interface displays the performance evaluation of the trained machine learning models using a confusion matrix and comparison graph. The confusion matrix illustrates the classification results by showing correctly and incorrectly predicted fraud and non-fraud transactions. The comparison graph evaluates all trained algorithms based on performance metrics such as Accuracy, Precision, Recall, and F1-



Score, enabling the selection of the best-performing model for fraud detection.

Fig. 8.7: Machine Learning Model Performance Comparison and Confusion.

8.8: FRAUD DETECTION FROM TEST DATA

The user enters a transaction index and clicks the Detect Fraud button. The system analyzes the selected transaction and predicts whether it is fraudulent or non-fraudulent using the trained machine learning model.

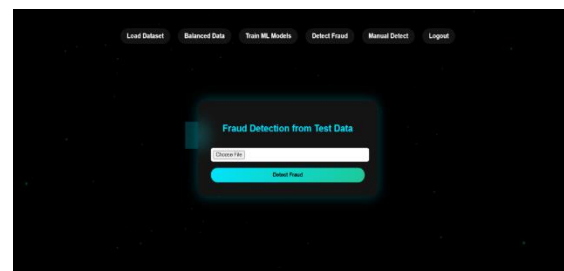


Fig. 8.8: Fraud Detection Interface Using the Trained Machine Learning Model.

8.9: MANUAL FRAUD DETECTION USING TRANSACTION FILE

This module enables users to upload a transaction file for manual fraud detection. After selecting the file, the trained machine learning model analyzes the transaction data and prepares it for fraud prediction.

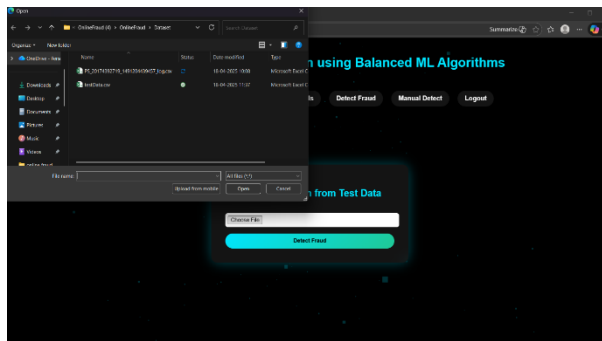


Fig. 8.9: Manual Fraud Detection Using Transaction File Upload.

8.10: FRAUD DETECTION RESULTS

This module displays the fraud detection results for the uploaded transaction data. Each transaction is classified as either Fraud Transaction or Normal Transaction, allowing users to quickly identify suspicious

Load Status	Related Data	Train ML Models	Detect Fraud	Manual	History
Transaction Data					Detection Rate
(step: 00, type: 'TRANSFER', amount: 43807.51, nameCmp: 'C36687787', walletBalance: 43807.51, walletBalance: 0.0, hashCmp: 'C44932926', walletBalance: 0.0, walletBalance: 0.0)					Final Transaction
(step: 00, type: 'CASH_OUT', amount: 43807.51, nameCmp: 'C36687787', walletBalance: 43807.51, walletBalance: 0.0, hashCmp: 'C36687787', walletBalance: 26784.58, walletBalance: 0.0, walletBalance: 0.0)					Final Transaction
(step: 03, type: 'CASH_IN', amount: 309433.53, nameCmp: 'C38911500', walletBalance: 311318.51, walletBalance: 0.0, hashCmp: 'C38911500', walletBalance: 45047.58, walletBalance: 0.0, walletBalance: 0.0)					Final Transaction
(step: 03, type: 'CASH_IN', amount: 99250.4, nameCmp: 'C37664299', walletBalance: 351692.14, walletBalance: 0.0, hashCmp: 'C37664299', walletBalance: 47578.58, walletBalance: 0.0, walletBalance: 0.0)					Final Transaction
(step: 03, type: 'TRANSFER', amount: 217.7, nameCmp: 'C37664299', walletBalance: 40181.96, walletBalance: 0.0, walletBalance: 0.0)					Final Transaction
(step: 03, type: 'TRANSFER', amount: 0.0, walletBalance: 0.0, walletBalance: 0.0)					Final Transaction
(step: 05, type: 'CASH_OUT', amount: 25407.58, nameCmp: 'C38911500', walletBalance: 47483.1, walletBalance: 25218.58, walletBalance: 0.0, hashCmp: 'C38922611', walletBalance: 0.0, walletBalance: 0.0)					Final Transaction
(step: 05, type: 'TRANSFER', amount: 42429.53, nameCmp: 'C31202137', walletBalance: 42679.53, walletBalance: 0.0, hashCmp: 'C31202137', walletBalance: 0.0, walletBalance: 0.0)					Final Transaction
(step: 06, type: 'CASH_OUT', amount: 42679.53, nameCmp: 'C34782669', walletBalance: 42679.53, walletBalance: 0.0, hashCmp: 'C36716111', walletBalance: 0.0, walletBalance: 42679.53)					Final Transaction
(step: 06, type: 'CASH_OUT', amount: 56142.81, nameCmp: 'C31202137', walletBalance: 56142.81, walletBalance: 0.0, hashCmp: 'C36726373', walletBalance: 0.0, walletBalance: 0.0)					Final Transaction
(step: 06, type: 'CASH_OUT', amount: 56142.81, nameCmp: 'C31202137', walletBalance: 56142.81, walletBalance: 0.0, hashCmp: 'C36726373', walletBalance: 0.0, walletBalance: 0.0)					Final Transaction
(step: 06, type: 'CASH_OUT', amount: 25044.58, nameCmp: 'C34159676', walletBalance: 0.0, walletBalance: 0.0, hashCmp: 'C34159676', walletBalance: 0.0, walletBalance: 0.0)					Final Transaction

activities.

Fig. 8.10: Fraud Detection Results for
Uploaded Transactions.

8.11: MANUAL TRANSACTION FRAUD PREDICTION INTERFACE

This module provides a form for manually entering transaction details, including transaction type, amount, account balance, and account ID. The entered information is

used by the trained machine learning model to predict fraudulent transactions.

[Load Dataset](#) [Balancecard Data](#) [Train ML Models](#) [Detect Fraud](#) [Manual Detect](#) [History](#) [Logout](#)

Manual Transaction Fraud Prediction

Transaction Type

Amount

Old Balance (Sender)

New Balance (Sender)

Account ID (nameOrig)

Predict Fraud

Fig. 8.11: Manual Transaction Fraud Prediction Interface.

8.12: TRANSACTION DETAILS ENTRY FOR FRAUD PREDICTION

This screen displays the completed transaction details entered by the user. After entering the required information, the transaction is submitted to the trained model for fraud analysis and prediction.

[Load Dataset](#) [Train ML Models](#) [Detect Fraud](#) [Manual Detect](#) [History](#) [Logout](#)

Manual Transaction Fraud Prediction

Transaction Type

-- Select --

Amount

Old Balance (Sender)

New Balance (Sender)

Account ID (nameOrig)

Predict Fraud

TRANSACTION BLOCKED (Account Already Blocked)

The account is permanently blocked.

Fig. 8.12: Transaction Details Entered for Fraud Prediction.

8.13: FRAUDULENT TRANSACTION DETECTION

This module displays the prediction result for a fraudulent transaction. When

suspicious transaction details are detected, the system blocks the transaction and notifies the user with an alert message.

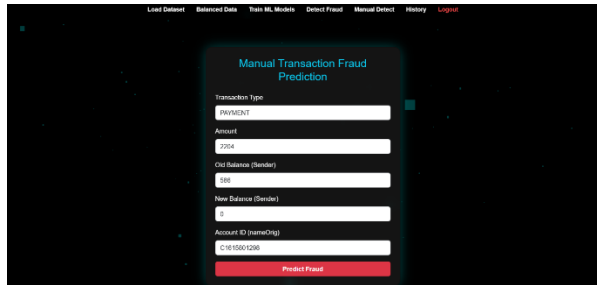


Fig. 8.13: Fraudulent Transaction Detection and Blocking.

8.14: GENUINE TRANSACTION DETECTION

This module displays the prediction result for a genuine transaction. If the transaction is classified as legitimate, the system approves it and confirms the successful transaction.

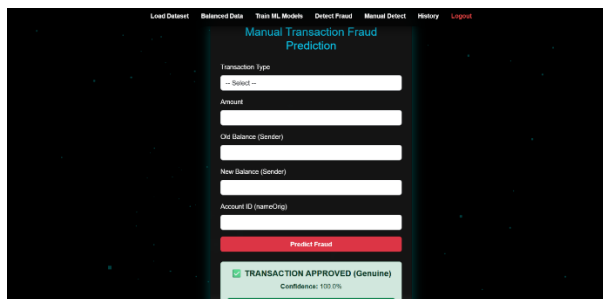


Fig. 8.14: Genuine Transaction Approval After Fraud Analysis.

8.15: MANUAL TRANSACTION DATA ENTRY

This module allows the user to manually enter transaction details such as transaction type, amount, sender's account balance, updated balance, and account ID. The entered information is processed by the

trained machine learning model for fraud prediction.

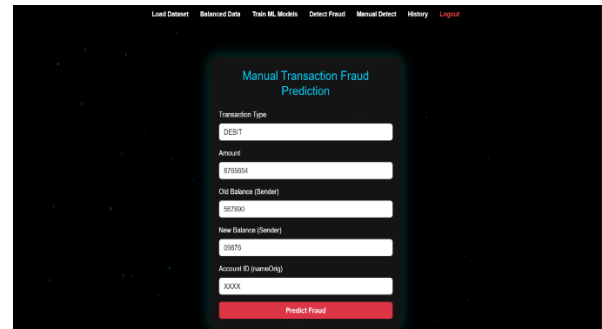


Fig. 8.15: Manual Entry of Transaction Details for Fraud Prediction.

8.16: INVALID ACCOUNT VALIDATION

This module validates the entered account ID before performing fraud prediction. If the account ID is not found in the dataset, the system displays an Invalid Account ID message and prevents further prediction.

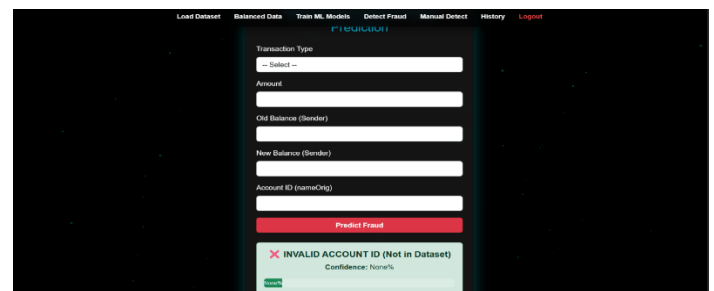
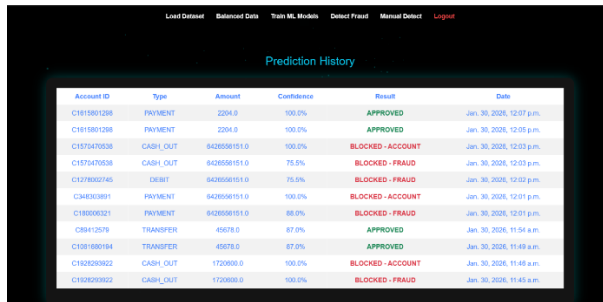


Fig. 8.16: Invalid Account ID Validation During Manual Fraud Detection.

8.17: PREDICTION HISTORY

This module displays the history of all fraud prediction results, including account ID, transaction type, amount, confidence score, prediction result, and date. It enables users

to review previous predictions and monitor transaction activities.



Account ID	Type	Amount	Confidence	Result	Date
C1815801298	PAYMENT	2204.0	100.0%	APPROVED	Jan. 30, 2026, 12:07 p.m.
C1815801298	PAYMENT	2204.0	100.0%	APPROVED	Jan. 30, 2026, 12:08 p.m.
C187470838	CASH_OUT	842659151.0	100.0%	BLOCKED - ACCOUNT	Jan. 30, 2026, 12:08 p.m.
C187470838	CASH_OUT	842659151.0	75.0%	BLOCKED - FRAUD	Jan. 30, 2026, 12:08 p.m.
C187470838	CASH_OUT	842659151.0	75.0%	BLOCKED - FRAUD	Jan. 30, 2026, 12:08 p.m.
C34830891	PAYMENT	842659151.0	100.0%	BLOCKED - ACCOUNT	Jan. 30, 2026, 12:08 p.m.
C18808321	PAYMENT	842659151.0	88.0%	BLOCKED - FRAUD	Jan. 30, 2026, 12:08 p.m.
C88412079	TRANSFER	45078.0	87.0%	APPROVED	Jan. 30, 2026, 11:54 a.m.
C1815801298	TRANSFER	45078.0	87.0%	APPROVED	Jan. 30, 2026, 11:54 a.m.
C182829822	CASH_OUT	1720000.0	100.0%	BLOCKED - ACCOUNT	Jan. 30, 2026, 11:49 a.m.
C182829822	CASH_OUT	1720000.0	100.0%	BLOCKED - FRAUD	Jan. 30, 2026, 11:49 a.m.

Fig. 8.17: Prediction History of Fraud Detection Results.

VIII. CONCLUSION

The UPI Fraud Transaction Detection using Machine Learning project successfully demonstrates the application of advanced machine learning techniques for identifying fraudulent digital payment transactions. With the increasing adoption of UPI-based payment systems, ensuring transaction security has become a major concern for financial institutions and users. The proposed system addresses this challenge by providing an intelligent and automated fraud detection mechanism capable of distinguishing genuine transactions from fraudulent ones with high accuracy.

The project utilizes effective data preprocessing techniques to improve data quality and applies Principal Component Analysis (PCA) for selecting the most relevant transaction features. By reducing

dimensionality and eliminating redundant information, PCA enhances computational efficiency and improves the overall performance of the machine learning models. The processed dataset is then used to train and evaluate both the existing Naïve Bayes algorithm and the proposed XG Boost algorithm.

IX. FUTURE WORK

The UPI Fraud Transaction Detection using Machine Learning system provides an effective solution for identifying fraudulent transactions with high accuracy. However, there are several opportunities to further enhance the system and improve its performance, scalability, and security. Future developments can focus on integrating advanced technologies and intelligent mechanisms to make the fraud detection process more robust and adaptable to emerging cyber threats.

One of the major future enhancements is the implementation of real-time fraud detection. The current system analyzes transactions from uploaded datasets, whereas future versions can be integrated directly with banking and UPI payment platforms to monitor transactions in real time. This will enable instant identification and prevention of fraudulent activities before financial losses occur.

Another important area for future improvement is the incorporation of deep learning techniques such as Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), and Recurrent Neural Networks (RNN). These advanced models can learn complex transaction patterns and improve fraud detection accuracy by identifying hidden relationships within transaction data.

X. REFERENCES

- [1] M. D. Nazmoddin, M. Swetha, G. Yashwanthi, and Y. Divyasree, "UPI Fraud Detection Using Machine Learning," *Journal of Computational Analysis and Applications*, vol. 33, no. 6, pp. 1704–1707, 2024.
- [2] R. Rani, A. Alam, and A. Javed, "Secure UPI: Machine Learning-Driven Fraud Detection System for UPI Transactions," in *2024 2nd International Conference on Disruptive Technologies (ICDT)*, IEEE, pp. 924–929, 2024.
- [3] V. Dave and D. Chudasama, "Enhancing Fraud Detection in UPI Transactions using Ensemble Learning and Neural Networks," *Journal of Information Systems Engineering and Management*, vol. 10, no. 45s, 2025.
- [4] N. Verma and N. Tyagi, "Fraud Detection Mechanism for UPI Payments Using XGBoost Outlier Detection," *Lecture Notes in Electrical Engineering*, vol. 1433, pp. 329–338, 2025.
- [5] G. Ramya, A. N. V. Gayathri, H. Altaf, and M. I. K. Shariff, "Explainable UPI Fraud Detection Using XGBoost and SHAP-Based Interpretable AI," *International Journal of Engineering Research and Science & Technology*, vol. 22, no. 1(2), pp. 289–295, 2026.
- [6] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, 2016.
- [7] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 2nd ed., Sebastopol, CA, USA: O'Reilly Media, 2019.
- [8] S. M. Lundberg and S. I. Lee, "A Unified Approach to Interpreting Model Predictions," *Advances in Neural Information Processing Systems*, vol. 30, pp. 4765–4774, 2017.
- [9] A. Dal Pozzolo, O. Caen, R. A. Johnson, and G. Bontempi, "Calibrating Probability with Under sampling for Unbalanced Classification," *IEEE Symposium Series on Computational Intelligence*, pp. 159–166, 2015.
- [10] C. C. Aggarwal, *Outlier Analysis*, 2nd ed., Cham, Switzerland: Springer, 2017.
- [11] J. Chen, K. Li, K. Rong, H. Bilal, and K. Li, "A Survey of Machine Learning for Big Data Processing," *EURASIP Journal on Advancing Signal Processing*, vol. 2018, no. 1, pp. 1–16, 2018.
- [12] B. Sethi, S. Mhatre, S. Yadav, S. Das, and V. Jadhav, "Machine Learning-Based UPI Fraud Detection: A Comprehensive

Approach Using Random Forest,” in Proceedings of ICAIEHS 2025, Atlantis Press, 2025.

[13] F. Chollet, Deep Learning with Python, 2nd ed., Manning Publications, 2021.

[14] N. B. Chakka and S. S. Saheb, “A Hybrid Autoencoder-XGBoost Framework for High- Performance UPI Fraud Detection,” Proceedings of Engineering and Technology Innovation, vol. 33, pp. 93–105, 2026.

[15] F. Almalki and M. Masud, “Financial Fraud Detection Using Explainable AI and Stacking Ensemble Methods,” arXiv preprint arXiv:2505.10050, 2025.

[16] A. Fajar, S. Yazid, and I. Budi, “Enhancing Phishing Detection through Feature Importance Analysis and Explainable AI: A Comparative Study of CatBoost, XGBoost, and EBM Models,” arXiv preprint arXiv:2411.06860, 2024.

[17] M. N. Uddin and M. M. Aziz, “Shapley Value-Guided Adaptive Ensemble Learning for Explainable Financial Fraud Detection with U.S. Regulatory Compliance Validation,” arXiv preprint arXiv:2604.14231, 2026.

[18] S. Raschka, H. Liu, and V. Mirjalili, Machine Learning with PyTorch and Scikit-Learn, Packt Publishing, 2022.

[19] M. Zaharia et al., “Apache Spark: A Unified Engine for Big Data Processing,” Communications of the ACM, vol. 59, no. 11, pp. 56–65, 2016.

[20] S. Wang, Y. Wang, and X. Liu, “Real-Time Financial Fraud Detection Using Machine Learning Techniques,” International Journal of Information

Management Data Insights, vol. 5, no. 1, pp. 100–112, 2025.

[21] A. Singh and P. Sharma, “Cloud-Based Fraud Detection Framework for Digital Payment Systems,” Journal of Cloud Computing, vol. 14, no. 2, pp. 45–58, 2025.

[22] C. Molnar, Interpretable Machine Learning, 2nd ed., Lulu.com, 2022.

[23] J. Brownlee, Machine Learning Mastery with Python, Machine Learning Mastery, 2020.

[24] Y. LeCun, Y. Bengio, and G. Hinton, “Deep Learning,” Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[25] D. Sreenivasulu, S. Chakri, U. Karthik, and U. Aravind, “Fraud Detection in UPI Transactions Using Machine Learning,” *International Journal for Research in Applied Science and Engineering Technology (IJRASET)*, vol. 14, no. 3, 2026, doi: 10.22214/ijraset.2026.78762.

[26] T. Sekhar and G. Vijaya Kumar, “UPI Fraud Detection Using Hybrid Machine Learning Models with Explainable Risk Scoring and Real-Time Monitoring,” *International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE)*, 2026, doi: 10.17148/IJARCCE.2026.15419.

[27] S. Sahu, P. Chaudhury, S. P. Senapati, S. Pradhan, and S. K. Nahak, “Detection of Mule Accounts and Fraudsters in UPI Transactions Using AI and Machine Learning Techniques,” *Iconic Research and Engineering Journals*, vol. 9, no. 10, pp. 904–913, 2026.

[28] J. Kumar and N. Rani, “Optimized Machine Learning and Deep Learning

Approaches for Effective Detection of Fraud in Unified Payments Interface (UPI) Transactions,” *International Journal of Scientific and Advanced Technology (IJSAT)*, vol. 16, no. 4, 2025.

[29] J. Arjun, D. Karthikeyan, and D. Deepika, “UPI Fraud Detection Using Machine Learning,” in *Challenges in Information Technology*, CRC Press, Taylor & Francis, 2024.

[30] M. Sorinola and K. Hamed, “Machine Learning for Digital Payment Fraud Detection in Financial Institutions: A Review of Methods, Explainability, and Performance,” *Cureus Journal of Computer Science*, vol. 3, 2026.