Grenze International Journal of Engineering and Technology, June Issue

Pay Shield: Deep Learning Approach for UPI Fraud Prevention

K. Akhila[1], K. Tejaswani[2], Manas Kumar Maharana[3], R.G.P. Annapurna[4] and S. Neeraja[5]
[1-4]Dept. of Computer Science and Systems Engineering, Lendi Institute of Engineering and Technology, Vizianagaram, India
Email: akhila948k@gmail.com, tejaswanikarnapu@gmail.com, maharanamanas1203@gmail.com, prasunaravinutala@gmail.com
[5]Assistant Professor, Dept. of Computer Science and Systems Engineering, Lendi Institute of Engineering and Technology, Vizianagaram, India
Email: neeraja.s@lendi.edu.in

***Abstract*— UPI Fraud Prevention System is a deep learning–based solution designed to detect and prevent fraudulent digital payment transactions in real time. The system processes transaction data such as amount, channel, type, and temporal features, followed by preprocessing techniques including feature extraction, encoding, and normalization. Deep learning models such as an Autoencoder for anomaly detection and an Artificial Neural Network (ANN) for fraud classification are implemented to identify suspicious transactions accurately. A hybrid security mechanism integrates OTP-based verification for high-value transactions to provide an additional layer of protection. The system also includes a web-based interface and dashboard for monitoring transactions and fraud detection results. This approach enhances transaction security, reduces financial risks, and improves trust in UPI-based digital payment systems.**

***Index Terms*— UPI Fraud Prevention, Deep Learning, Autoencoder, Artificial Neural Network (ANN), OTP Authentication, Anomaly Detection, Digital Payment Security.**

I. INTRODUCTION

Over the past decade, digital payment systems have significantly transformed the global financial ecosystem. In India, the Unified Payments Interface (UPI) has emerged as one of the most widely adopted real-time payment systems, enabling instant, low-cost, and seamless transactions across banking platforms. UPI has played a crucial role in enhancing financial inclusion, reducing dependency on physical cash, and promoting a sustainable digital economy. With billions of transactions processed monthly, UPI has become a vital component of India's digital infrastructure.

However, the rapid growth of UPI transactions has also led to a substantial increase in fraudulent activities. Common fraud techniques include phishing attacks, malicious QR codes, fake payment requests, and social engineering tactics. These threats result in significant financial losses for users and financial institutions, while also reducing trust in digital payment systems. Existing fraud detection mechanisms are largely reactive, relying on post-transaction monitoring to identify suspicious activities after the transaction has already been completed. This limitation reduces the effectiveness of preventing fraud in real time.

Grenze ID: 01.GIJET.12.2.182_20

Recent advancements in deep learning provide promising solutions for proactive fraud detection and prevention. Deep learning models are capable of automatically learning complex patterns and behavioral anomalies from large-scale transaction data without extensive manual feature engineering. Techniques such as Autoencoders and Artificial Neural Networks (ANN) have shown strong performance in anomaly detection and classification tasks within financial security domains. However, integrating these models into a real-time fraud prevention framework with additional security mechanisms remains a challenging task.

In this work, we propose an end-to-end deep learning-based UPI fraud prevention system that combines anomaly detection and supervised classification for enhanced security. The system utilizes an Autoencoder model to learn normal transaction behavior and identify anomalous patterns. In addition, an ANN model is employed to classify transactions as fraudulent or legitimate based on extracted features such as transaction amount, channel, transaction type, and temporal attributes including transaction hour and month. Furthermore, a threshold-based One-Time Password (OTP) authentication mechanism is incorporated to provide an additional layer of security for high-value transactions.

The primary objective of this research is to develop a reliable and scalable framework for real-time UPI fraud detection that can minimize potential financial losses. The proposed system performs real-time transaction analysis and integrates deep learning models with rule-based authentication to enhance overall system security. The performance of the system is evaluated using standard metrics such as accuracy, precision, recall, and F1-score to ensure effectiveness and reliability.

The major contributions of this study include the development of a deep learning-based framework for real-time UPI fraud detection using Autoencoder and ANN models, the implementation of a hybrid fraud prevention mechanism that combines deep learning predictions with OTP-based authentication, and the design of a scalable and user-friendly system for monitoring transactions and enhancing digital payment security. The proposed system integrates intelligent fraud detection with practical implementation, providing an effective solution to improve the security, reliability, and trustworthiness of UPI-based digital payment systems.

II. Literature Survey

The rapid growth of digital payment systems, particularly the Unified Payments Interface (UPI), has significantly enhanced financial accessibility and transaction efficiency. However, the large-scale adoption of UPI has also increased exposure to cyber threats and financial fraud. Researchers have extensively explored machine learning and deep learning techniques to strengthen fraud detection mechanisms in digital payment ecosystems.

Traditional machine learning approaches have been widely used for transaction-based fraud detection. AP CA et al. [1] proposed a Support Vector Machine (SVM)-based anomaly detection framework to identify suspicious transaction patterns such as phishing attempts and unauthorized access. Their model achieved high precision and recall, demonstrating the effectiveness of supervised learning techniques in structured financial datasets. Similarly, BR Haleritti et al. [2] conducted a comparative study evaluating multiple algorithms, including Logistic Regression, K-Nearest Neighbors (KNN), Decision Tree, Random Forest, and Convolutional Neural Networks (CNN). Their experimental results indicated that Random Forest outperformed other models due to its robustness in handling high-dimensional features and noisy transaction data. While these approaches provide reliable classification performance, they are largely dependent on labeled datasets and predefined fraud patterns.

Phishing attacks and malicious QR codes represent another major source of fraud in UPI transactions. GR Charan and KD Thilak [3] applied machine learning techniques to detect fraudulent URLs and QR code manipulations, significantly improving transaction security within digital payment systems. Their study emphasized feature extraction from URL structures and image-based QR analysis to classify malicious inputs. However, these solutions primarily focus on identifying known attack signatures rather than detecting novel fraud strategies.

Recent advancements in deep learning have enabled more sophisticated fraud detection frameworks capable of capturing complex behavioral patterns. M Tamilselvi et al. [4] introduced a hybrid model integrating ShuffleNet for feature extraction with SVM for classification, achieving improved computational efficiency and detection accuracy. Likewise, MN Raju et al. [5] implemented Long Short-Term Memory (LSTM) networks to analyze sequential transaction behavior, effectively identifying temporal dependencies and evolving fraud patterns. Sequence-based models such as LSTM are particularly beneficial in detecting irregular transaction frequency and abnormal behavioral deviations over time.

Data imbalance remains a significant challenge in financial fraud detection due to the relatively low proportion of fraudulent transactions compared to legitimate ones. NRS Kumar et al. [6] demonstrated that neural network

models combined with appropriate data balancing techniques and cross-validation strategies can significantly improve fraud detection accuracy and reduce false-positive rates. Their findings highlight the importance of model optimization and validation in enhancing real-world deployment feasibility.

In addition to these approaches, recent studies have explored deep learning-based anomaly detection techniques for financial fraud prevention. S. Roy et al. [7] proposed an Autoencoder-based fraud detection model that learns normal transaction patterns and identifies anomalies using reconstruction error. Their approach demonstrated improved detection of previously unseen fraud cases compared to traditional classification methods. Similarly, K. Veeramachaneni et al. [8] introduced a real-time fraud detection system using machine learning pipelines capable of handling large-scale streaming transaction data, emphasizing scalability and low-latency decision-making.

Furthermore, Y. Jurgovsky et al. [9] developed a deep learning framework using Recurrent Neural Networks (RNNs) for credit card fraud detection, highlighting the importance of sequential modeling in identifying temporal fraud patterns. Their results showed improved fraud detection rates compared to traditional machine learning methods. Additionally, P. Mishra et al. [10] proposed a hybrid fraud detection model combining supervised and unsupervised learning techniques, demonstrating enhanced performance in detecting both known and unknown fraud patterns in financial transactions.

Despite these advancements, most existing studies focus primarily on post-transaction fraud detection, where suspicious activity is identified after transaction initiation or completion. Such reactive mechanisms may reduce future risks but cannot fully prevent immediate financial losses. Furthermore, many models rely solely on classification techniques without integrating anomaly detection mechanisms capable of identifying previously unseen fraud patterns. Limited research has explored hybrid architectures that combine unsupervised anomaly detection with supervised classification for real-time prevention.

To address these limitations, the present study proposes an integrated framework that combines Autoencoder-based anomaly detection with Artificial Neural Network (ANN)-based classification to enable proactive pre-transaction risk assessment.

Additionally, a threshold-based One-Time Password (OTP) verification mechanism is incorporated to provide multi-layered security for high-risk transactions. By integrating real-time anomaly detection, classification, and user verification, the proposed system aims to enhance financial resilience, reduce economic losses, and strengthen trust in sustainable digital payment ecosystems.

III. METHODOLOGY

The Pay Shield framework aims to detect and prevent fraudulent UPI transactions in real time using a hybrid deep learning approach. This method combines supervised learning to identify known fraud patterns and unsupervised learning to detect unusual or previously unseen fraudulent activities. The integration of these techniques improves detection accuracy while reducing the chances of missing fraudulent transactions.

The Pay Shield workflow consists of key stages including data collection, preprocessing, feature selection, model training, real-time transaction monitoring, decision-making, and continuous improvement.

A. Data Collection

The dataset used in this system consists of structured UPI transaction records containing both legitimate and fraudulent transactions. The dataset includes attributes such as transaction amount, transaction type, channel, date, and time, along with a fraud label.

The data is prepared to ensure consistency and reliability. Duplicate entries and invalid records are removed to maintain data quality. The dataset includes a mix of normal and fraudulent transactions, enabling the model to learn different transaction behaviors effectively.

B. Feature Engineering

Feature engineering in the proposed system focuses on selecting and transforming key attributes that effectively represent transaction behavior. The primary features considered include transaction amount, channel, transaction type, transaction hour, and transaction month.

The original date and time attributes are converted into hour and month to capture important temporal patterns in user transactions. Additionally, numerical values such as transaction amount are normalized to ensure consistent scaling, while categorical features like channel and transaction type are encoded into a suitable format for model processing.

These transformations enable the system to better understand transaction patterns and improve its ability to identify unusual or potentially fraudulent activities.

C. Hybrid Deep Learning Model

The Pay Shield system uses a hybrid deep learning model consisting of:

Artificial Neural Network (ANN)

The ANN is a supervised learning model that classifies transactions as legitimate or fraudulent based on learned patterns from labeled data.

Autoencoder

The Autoencoder is an unsupervised model trained on normal transactions. It detects anomalies by identifying transactions that deviate from learned normal behavior.

This combination enables the system to detect both known and unknown fraud patterns effectively.

D. Training and Validation

The ANN model is trained using labeled transaction data, while the Autoencoder is trained using only non-fraudulent transactions. The models are evaluated using performance metrics such as accuracy, precision, recall, and F1-score. The dataset is split into training and testing sets to ensure proper validation. Threshold values are selected to balance fraud detection and minimize false positives.

E. Real-Time Transaction Monitoring

In a real-time environment, each incoming UPI transaction is processed immediately. The system extracts relevant features and passes them to both the ANN and Autoencoder models.The models analyze the transaction and determine whether it is normal or potentially fraudulent before the transaction is completed.

F. Alert and Response Mechanism

If a transaction is identified as suspicious, the system applies a response mechanism based on predefined rules.

For transactions with an amount greater than ₹1000, an OTP verification is triggered:

- If the OTP is verified successfully, the transaction is approved
- If the OTP verification fails, the transaction is blocked

For lower-value transactions:

- Fraud detected → Transaction Blocked
- Otherwise → Transaction Approved

This mechanism ensures secure and user-verified transaction processing.

G. Model Adaption and Updates

The system supports periodic retraining of models using updated transaction data to adapt to evolving fraud patterns. Continuous updates improve the system's ability to detect new types of fraud and maintain long-term effectiveness

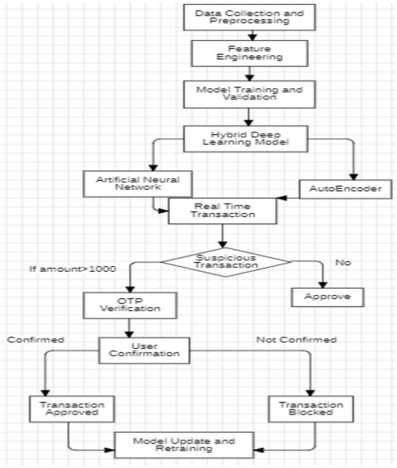

Fig 1. Flowchart illustrating the methodology of the proposed Pay Shield framework.

IV. Results and Discussion

The proposed hybrid deep learning-based UPI fraud prevention system (Pay Shield) was evaluated using a dataset of UPI transactions containing both legitimate and fraudulent records. The dataset was divided into training and testing sets using an 80:20 split. Model performance was evaluated using standard metrics such as Accuracy, Precision, Recall, F1-score, and confusion matrix analysis.

A. Quantitative Results

The hybrid model consisting of Artificial Neural Network (ANN) and Autoencoder achieved the following performance on the test dataset.

Table I: Performance of the Proposed Hybrid Model

Metric	Value
Accuracy	96.08%
Precision (Fraud)	95.00%
Recall (Fraud)	96.00%
F1-Score	95.560%

The proposed model achieved an accuracy of 96%, with fraud precision of 95% and fraud recall of 96%. The F1-score of 95.5% indicates a good balance between precision and recall, ensuring that fraudulent transactions are detected effectively while minimizing false alerts.

The model correctly classified the majority of transactions, with a high number of true positives and true negatives. False positives represent legitimate transactions incorrectly flagged as fraud, while false negatives correspond to missed fraudulent transactions. The relatively low number of false negatives demonstrates that the system is effective in detecting fraudulent activities

B. Confusion Matrix Analysis

Table II: Confusion Matrix of the Proposed Hybrid Model

	Predicted Normal	Predicted Fraud
Actual Normal	185	6
Actual Fraud	7	35

The confusion matrix shows that the model correctly classified 185 legitimate transactions (True Negatives) and 35 fraudulent transactions (True Positives). Only 6 legitimate transactions were incorrectly flagged as fraud (False Positives), while 7 fraudulent transactions were missed (False Negatives).

The low number of misclassifications demonstrates that the model performs effectively in distinguishing between normal and fraudulent transactions, achieving high detection accuracy with minimal errors

C. Qualitative Results

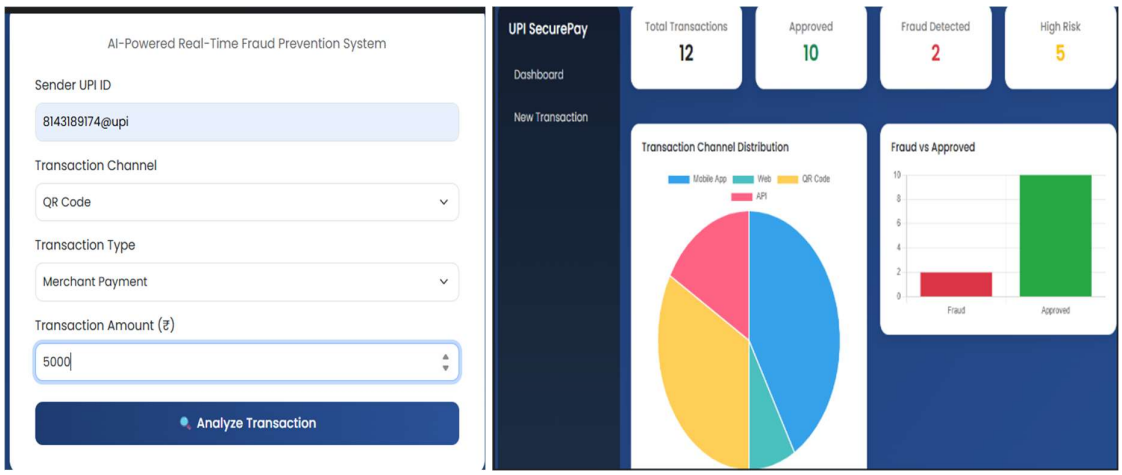

Fig.2. Real-time output of the UPI Fraud Prevention

The proposed system was implemented in a real-time transaction monitoring environment to evaluate its practical applicability. Each UPI transaction is analyzed dynamically using the trained hybrid model. Transactions exceeding ₹1000 trigger an additional OTP-based verification step. If the OTP entered by the user is correct, the transaction is successfully processed. Otherwise, the transaction is flagged as fraudulent and blocked. For low-value transactions, the model directly predicts whether the transaction is legitimate or fraudulent. The system demonstrated efficient real-time performance with minimal latency, ensuring secure and seamless transaction processing. Fraudulent transactions were effectively detected and prevented before completion, reducing the risk of financial loss.

The experimental results demonstrate that the proposed hybrid model provides an effective solution for UPI fraud detection by combining ANN and Autoencoder to detect both known patterns and anomalies. The model significantly improves fraud detection performance, particularly in terms of recall, while maintaining low false positives and false negatives. The integration of OTP-based verification further enhances system security by adding an additional authentication layer. However, the system may face challenges in detecting highly sophisticated fraud patterns and its performance depends on the quality of the dataset.

V. Conclusion

The study presented a hybrid deep learning-based UPI fraud prevention system (Pay Shield) for real-time detection of fraudulent transactions. The system integrates an Artificial Neural Network (ANN) with an Autoencoder model to effectively capture both known fraud patterns and anomalous transaction behaviors. The proposed model achieved high accuracy of 96.08%, along with strong precision and recall, demonstrating its capability to accurately identify fraudulent transactions while minimizing false alarms. Confusion matrix analysis further confirmed the robustness of the model, with low false positives and false negatives, ensuring reliable transaction classification. Compared to individual models such as ANN and Autoencoder, the hybrid approach showed improved performance, highlighting the advantage of combining supervised and unsupervised learning techniques. The integration of OTP-based verification adds an additional security layer for high-value transactions, enhancing the system's practical applicability in real-world financial environments. Real-time evaluation demonstrated efficient performance with minimal latency, ensuring seamless and secure transaction processing. Despite the strong results, certain limitations remain, particularly in detecting highly sophisticated fraud patterns and dependency on dataset quality. Future work may focus on incorporating behavioral analytics, advanced deep learning models, explainable AI techniques, and large-scale real-world deployment. Overall, the findings indicate that hybrid deep learning approaches provide an effective and scalable solution for enhancing the security of digital payment systems and mitigating UPI fraud.

References

[1] A. P. C. A., V. M. Monika, H. R. Pragathi, M. R. Rakshitha, and T. M. Kavya, "Secure Digital Payments: Fraud Detection in UPI Using Machine Learning Techniques," Proc. Int. Conf. Intelligent Systems, 2025.

[2] B. R. Haleritti, M. S. Muneshwara, J. Kania, and M. S. Swetha, "Detecting UPI Fraud Using Machine Learning Techniques in Python," Proc. Int. Conf. Computational Systems, 2025.

[3] G. R. Charan and K. D. Thilak, "Detection of Phishing Link and QR Code of UPI Transaction using Machine Learning," Proc. Int. Conf. Innovative Mechanisms, 2023.

[4] M. Tamilselvi, R. Begum, J. Giri, D. Sheela, and M. O. Sabri, "Experimental Evaluation of UPI Fraud Detection System Using Elevated Deep Learning Methodology," Proc. Int. Conf. Emerging Trends, 2024.

[5] M. N. Raju, Y. C. Reddy, P. N. Babu, V. S. P. Ravipati, and V. Chaitanya, "Detection of Fraudulent Activities in Unified Payments Interface using LSTM Networks," Proc. Int. Conf. Circuit Power and Computing Technologies, 2024.

[6] N. R. S. Kumar, K. Hemavathi, K. Sravani, T. Aparna, and M. Deepthi, "Enhancing UPI Security through Deep-Learning Based Fraud Detection," Proc. OPJU Int. Technology Conf. on Smart Computing, 2025.

[7] S. Roy, J. J. Qian, and V. K. Jain, "Anomaly Detection in Financial Transactions using Autoencoders," Proc. IEEE Int. Conf. Data Mining Workshops, 2018.

[8] K. Veeramachaneni, I. Arnaldo, V. Korrapati, C. Bassias, and K. Li, "AI²: Training a Big Data Machine to Defend," Proc. IEEE Int. Conf. Big Data Security, 2016.

[9] J. Jurgovsky, M. Granitzer, S. Ziegler, et al., "Sequence Classification for Credit Card Fraud Detection," Proc. IEEE Int. Conf. Data Science and Advanced Analytics, 2018.

[10] P. Mishra, S. Sharma, and R. Gupta, "Hybrid Machine Learning Approach for Financial Fraud Detection," Proc. IEEE Int. Conf. Computing, Communication and Automation, 2020.