RESEARCH PAPER

Deep Learning-Based Real-Time Fraud Detection Framework for UPI Transactions Using Hybrid CNN-BI LSTM Model

ASHISH MALIK and JITENDER KUMAR
[1]M.Tech Scholar, Department of Cyber Forensics and Information Security
Email Id-malik.ashish9070@gmail.com
[2]Assistant Professor, CSE Department Ganga Institute of Technology And Management Kablana Jhajjar Haryana, India
Email Id- jsaini.ymca@gmail.com

Abstract
The rapid adoption of Unified Payments Interface (UPI) has revolutionized digital payments by enabling instant, secure, and convenient financial transactions. However, the exponential growth in transaction volume has also increased the risk of fraudulent activities, necessitating advanced fraud detection mechanisms. This study proposes a Deep Learning-Based Real-Time Fraud Detection Framework using a hybrid CNN-BiLSTM architecture for identifying fraudulent UPI transactions. The framework combines the feature extraction capability of Convolutional Neural Networks (CNN) with the temporal sequence learning ability of Bidirectional Long Short-Term Memory (BiLSTM) networks to capture complex transaction patterns and evolving fraud behaviors. A benchmark dataset containing approximately 400,000 transactions from 40,000 customers and 8,000 merchants was utilized. Comprehensive data preprocessing, feature engineering, and exploratory data analysis were performed to enhance model performance. Experimental results demonstrated outstanding effectiveness, achieving 99.98% accuracy, 99.30% precision, 99.50% recall, 99.40% F1-score, and 1.000 ROC-AUC. The proposed framework successfully detected 98.35% of fraudulent transaction value, significantly reducing financial losses and improving transaction security. The findings confirm that the CNN-BiLSTM model provides a robust, scalable, and reliable solution for real-time UPI fraud detection in modern digital payment ecosystems.
Keywords: Unified Payments Interface (UPI), Deep Learning, Fraud Detection, CNN, Bi-LSTM
How to cite this article: Malik A, Kumar J. Deep Learning-Based Real-Time Fraud Detection Framework for UPI Transactions Using Hybrid CNN-BI LSTM Model. Int J Drug Deliv Technol. 2026;16(60s):1087-1103. DOI: 10.25258/ijddt.16.60s.120
Source of support: Nil.
Conflict of interest: None

1. Introduction

The rapid growth of digital payment systems has transformed the financial ecosystem by enabling fast, secure, and convenient monetary transactions [1]. Among various digital payment platforms, the Unified Payments Interface (UPI) has emerged as one of the most widely adopted real-time payment systems, particularly in India. UPI facilitates instant fund transfers between bank accounts through mobile devices, offering seamless interoperability among different banks and financial institutions [2]. The increasing popularity of UPI has significantly enhanced financial inclusion and digital commerce. However, the massive volume of transactions processed every second has also attracted cybercriminals who exploit system vulnerabilities to conduct fraudulent activities such as identity theft, phishing attacks, account takeovers, fake merchant transactions, and unauthorized fund transfers [3]. Consequently, ensuring secure and reliable fraud detection mechanisms has become a critical requirement for maintaining user trust and safeguarding financial transactions.

Traditional fraud detection systems primarily rely on rule-based approaches and conventional machine learning algorithms [4]. Although these methods can identify known fraud patterns, they often struggle to detect sophisticated and evolving fraudulent behaviors in real-time environments. Fraudulent transactions continuously adapt to bypass predefined rules, making static detection techniques less effective [5]. Furthermore, the highly imbalanced nature of financial transaction datasets, where fraudulent transactions constitute only a small fraction of total transactions, presents additional challenges for conventional classification models [6]. These limitations necessitate the adoption of intelligent and adaptive approaches capable of learning complex transaction patterns from large-scale financial data.

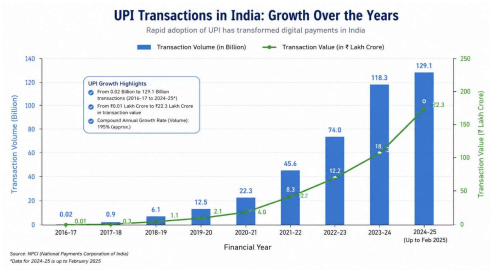

Figure 1: UPI transaction in India

Recent advancements in artificial intelligence and deep learning have demonstrated remarkable success in addressing fraud detection challenges. Deep learning models possess the capability to automatically extract meaningful features from high-dimensional data while capturing hidden relationships among transaction attributes [7]. In particular, Convolutional Neural Networks (CNNs) have shown effectiveness in identifying local feature patterns and correlations within transactional datasets. CNNs can automatically learn discriminative representations from transaction features, reducing the dependency on manual feature engineering [8]. However, fraud detection also requires understanding sequential and temporal behaviors because fraudulent activities often exhibit distinct transaction patterns over time.

To address temporal dependencies, Long Short-Term Memory (LSTM) networks have been extensively utilized due to their ability to retain long-term contextual information and model sequential data [9]. Nevertheless, standard LSTM architectures process information in a single direction, potentially overlooking future contextual relationships. Bidirectional Long Short-Term Memory (Bi-LSTM) networks overcome this limitation by processing sequences in both forward and backward directions, enabling comprehensive learning of transaction behaviors and temporal dependencies [10]. The combination of CNN and Bi-LSTM architectures provides a powerful framework for capturing both spatial feature representations and sequential transaction dynamics, thereby enhancing fraud detection accuracy.

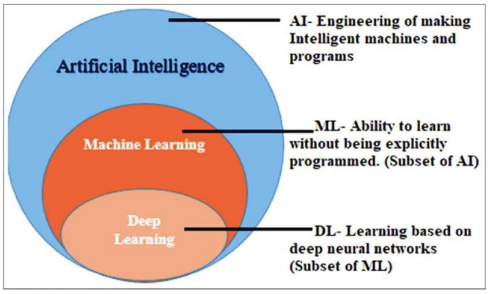

Figure 2: Use of AI in transaction [11]

This research proposes a deep learning-based real-time fraud detection framework for UPI transactions using a hybrid CNN-BiLSTM model. The proposed framework integrates the feature extraction capability of CNN with the temporal learning strength of Bi-LSTM to effectively distinguish legitimate transactions from fraudulent ones [12]. The model is designed to analyze transaction attributes such as transaction amount, frequency, merchant category, user behavior patterns, device information, geolocation, transaction timing, and other risk indicators. By leveraging deep hierarchical feature learning and bidirectional sequence modeling, the framework aims to identify subtle fraud signatures that may remain undetected by conventional methods [13].

In conclusion, the integration of CNN and Bi-LSTM networks offers a robust and scalable solution for real-time UPI fraud detection. As digital payment adoption continues to expand globally, intelligent fraud detection frameworks will play a vital role in strengthening cybersecurity, protecting consumers, and ensuring the reliability of modern financial transaction systems. The proposed hybrid deep learning framework contributes toward the development of next-generation fraud prevention systems capable of adapting to evolving cyber threats in dynamic digital payment environments. Here are the research objectives of the study follows as:

- To develop a real-time fraud detection framework for UPI transactions using a hybrid CNN-BiLSTM deep learning architecture capable of identifying fraudulent and legitimate transactions with high accuracy.
- To perform exploratory data analysis (EDA) for understanding transaction behavior, fraud patterns, temporal trends, risk indicators, and class imbalance characteristics in digital payment datasets.
- To integrate CNN and BiLSTM networks for extracting spatial transaction features and learning bidirectional temporal dependencies associated with fraudulent activities.
- To evaluate the proposed model using performance metrics including Accuracy, Precision, Recall, F1-Score, ROC-AUC, Average Precision, False Positive Rate (FPR), and False Negative Rate (FNR).

- To identify the most influential fraud indicators contributing to fraudulent UPI transactions and analyze their impact on prediction performance.
- To assess the business impact of the proposed framework in reducing financial losses, improving transaction security, and enhancing trust in digital payment systems.

2. Related Work

The literature on fraud detection in digital payment systems highlights the growing adoption of machine learning and deep learning techniques to address increasingly sophisticated financial fraud. Several research studies have focused on improving fraud detection accuracy while minimizing false positives in real-time transaction environments. Chakka et al. (2026) **[14]** proposed a hybrid Autoencoder–XGBoost framework for UPI fraud detection, achieving an ROC-AUC of 0.99999995 on a synthetic dataset through anomaly-aware feature engineering. Takale et al. (2026) **[15]** developed a real-time online fraud detection system integrating machine learning with Apache Kafka and Spring Boot architecture, where Random Forest achieved 96.3% accuracy. Similarly, Vikram et al. (2026) **[16]** introduced a stacking classifier-based anomaly detection framework and reported 95.4% accuracy with 99.2% ROC-AUC. Mohan et al. (2025) **[17]** employed a Bi-LSTM model for real-time UPI fraud detection and obtained 91.8% accuracy with a PR-AUC of 0.94, demonstrating the effectiveness of deep learning in handling imbalanced transaction data. Chhaparia et al. (2025) **[18]** further enhanced fraud detection by integrating Graph Neural Networks and LightGBM, enabling the identification of fraud rings and complex transaction relationships while maintaining false positive rates below 5%.

Author(s) & Year	**Method/Model Used**	**Dataset/Application**	**Key Results**	**Limitation**
Chakka et al. (2026)	Hybrid Autoencoder + XGBoost	2.68 million synthetic UPI transactions	Accuracy, Precision, Recall, F1-score ≈ 1.0; ROC-AUC = 0.99999995	Evaluated on synthetic data with deterministic fraud patterns
Takale et al. (2026)	Random Forest, SVM, Decision Tree, Logistic Regression	Online financial transaction system	Random Forest achieved 96.3% accuracy and 1.9% FPR	Limited validation on real-world dynamic fraud scenarios
Vikram et al. (2026)	Stacking Classifier, XGBoost, RF, KNN	Financial Transaction & Risk Management dataset	Accuracy = 95.4%, Precision = 95.7%, Recall = 95.4%, ROC-AUC = 99.2%	Computational complexity due to ensemble architecture
Mohan et al. (2025)	Bi-LSTM Deep Learning Framework	UPI transaction dataset (92.5% genuine, 7.5% fraud)	Accuracy = 91.8%, F1-score = 90.5%, PR-AUC = 0.94	Focused only on sequential transaction patterns
Chhaparia et al. (2025)	GraphSAGE + LightGBM	Real-time UPI fraud detection	ROC-AUC > 0.95, FPR < 5%, sub-millisecond inference	High infrastructure and deployment complexity

Palivela et al. (2024)	Ensemble Learning (GB, RF, LR, Voting)	Credit card transaction dataset	Accuracy = 99.59%	Primarily designed for credit card fraud, not UPI
Abid et al. (2024)	CNN-based Fraud Detection	Online payment transactions	Accuracy = 95%, Precision = 97.72%, Recall = 99.41%, F1-score = 98.56%	Requires significant computational resources
Talukder et al. (2024)	Integrated Multistage Ensemble ML (IMEML)	Credit card fraud dataset (284,807 transactions)	Accuracy = 99.94%, Precision = 99.91%, Recall = 99.14%, AUC = 100%	Performance validated on a single benchmark dataset
Hossain et al. (2024)	AI/ML-based Fraud Detection Framework	Electronic payment systems	Fraud reduction >25%, Detection accuracy ≈90%	Lacks detailed model-specific evaluation
Hanae et al. (2023)	Isolation Forest + Spark + Kafka	Digital banking transactions	Accuracy = 99%, Precision = 87%	Lower precision compared to supervised methods
Ileberi et al. (2022)	Genetic Algorithm + ML Classifiers	European credit card transactions	Improved fraud detection performance over baseline models	Dependent on feature selection quality
Aziz et al. (2022)	Modified LightGBM	Ethereum fraudulent transaction dataset	Accuracy = 99.17%	Focused on blockchain transactions rather than payment systems

Several other studies have explored advanced ensemble learning, convolutional neural networks, and big-data analytics for financial fraud detection. Palivela et al. (2024) **[19]** proposed an ensemble learning framework incorporating Gradient Boosting, Random Forest, and Voting Classifiers, achieving 99.59% accuracy in transaction fraud classification. Abid et al. (2024) **[20]** demonstrated the superiority of CNN-based fraud detection, reporting 95% accuracy and an F1-score of 98.56%. Talukder et al. (2024) **[21]** introduced an Integrated Multistage Ensemble Machine Learning (IMEML) model that achieved nearly perfect performance with 99.94% accuracy and 100% AUC. Hossain et al. (2024) **[22]** highlighted the role of AI and ML in reducing fraudulent transactions by over 25% while improving detection accuracy to approximately 90%. Hanae et al. (2023) **[23]** utilized Isolation Forest with Apache Spark and Kafka for real-time

fraud analytics, obtaining 99% accuracy and 87% precision. Earlier studies by Ileberi et al. (2022) **[24]** and Aziz et al. (2022) **[25]** emphasized feature optimization and LightGBM-based classification, respectively, achieving accuracies above 99%. Despite these advancements, most existing studies rely on either sequential models or ensemble classifiers independently, indicating the need for hybrid architectures such as CNN-BiLSTM that can simultaneously capture spatial feature representations and temporal transaction patterns for more robust UPI fraud detection.

3. Research Methodology

3.1 Dataset Collection

Digital Payment Fraud Detection Benchmark Dataset [26] used in this research consists of large-scale digital payment transaction records containing both legitimate and fraudulent UPI transactions. The dataset includes approximately 400,000 transaction records generated from nearly 40,000 customers and 8,000 merchants over a period of twelve months. To simulate realistic fraud detection scenarios, the dataset maintains a fraud ratio of nearly 1.5%–1.8%, representing the natural imbalance commonly observed in financial transaction systems. The collected dataset was further divided into training and testing datasets to facilitate model development, validation, and performance evaluation of the proposed fraud detection framework.

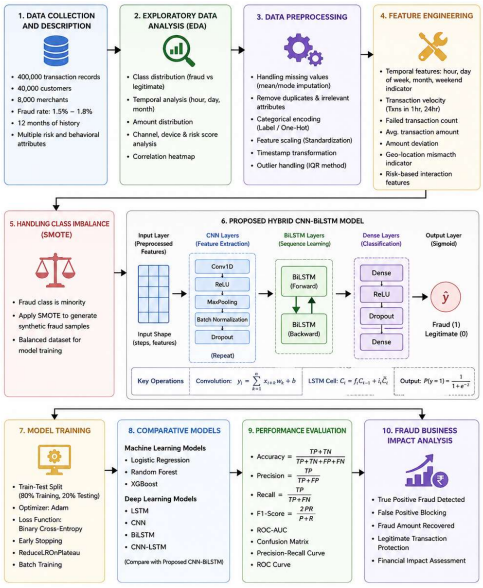

Figure 3: Research methodology flowchart

3.2 Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) is an important phase in the proposed fraud detection framework because it helps in understanding the characteristics, structure, patterns, and anomalies present in the UPI transaction dataset before model development. In this research, EDA was performed to analyze the behavior of fraudulent and legitimate transactions, identify hidden trends, detect outliers, examine feature distributions, and understand relationships among variables. Since financial fraud datasets are highly imbalanced and dynamic in nature, EDA plays a significant role in improving feature engineering and model performance.

Initially, the dataset structure was examined by analyzing the number of records, data types, missing values, duplicate entries, and fraud class distribution. Statistical summaries such as mean, median, standard deviation, minimum, and maximum values were computed for all numerical features including transaction amount, risk scores, transaction counts, and monthly spending. Missing values and duplicate transactions were removed to improve data quality and consistency. The fraud distribution analysis showed that fraudulent transactions represented only a very small portion of the dataset compared to legitimate transactions, indicating severe class imbalance. This imbalance justified the later use of SMOTE for oversampling minority fraud samples.

1. Fraud Class Distribution Analysis

The first stage of EDA involved analyzing the distribution of fraudulent and legitimate transactions using bar charts and pie charts. This analysis helped in understanding the imbalance between classes.

The fraud ratio is mathematically represented as:

$$Fraud\ Rate = \frac{Number\ of\ Fraudulent\ Transactions}{Total\ Transactions} \times 100$$

The analysis revealed that the fraud percentage was significantly lower than legitimate transactions, which is a common characteristic of financial fraud datasets.

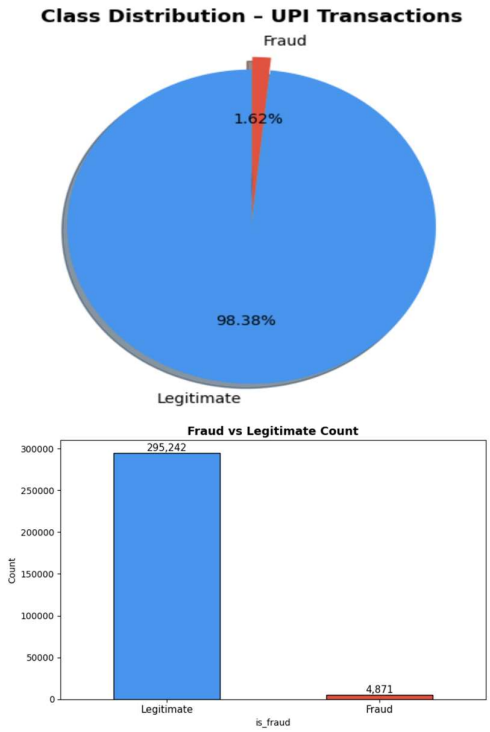

Figure 4: Class Distribution

2. Transaction Amount Analysis

Transaction amount analysis was performed using histograms, boxplots, and density plots to identify abnormal transaction patterns. Fraudulent transactions generally exhibited unusual transaction ranges and higher variance compared to legitimate transactions. Boxplot analysis also helped detect outliers and extreme transaction values.

The average transaction amount was computed as:

$$Average\ Transaction\ Amount = \frac{\sum_{i=1}^{n} Transaction\ Amoun\ t_i}{n}$$

Log transformation techniques were additionally used to normalize skewed transaction amount distributions.

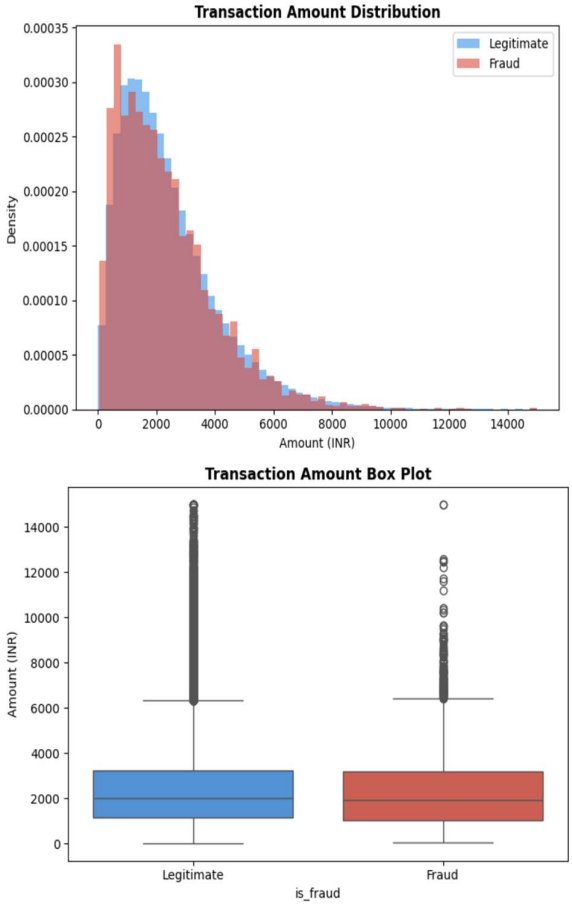

Figure 5: Transaction amount analysis

3. Temporal Transaction Analysis

Temporal analysis was conducted to identify fraud occurrence patterns over time. Features such as transaction hour, day of week, and month were extracted from transaction timestamps. Fraud rates were then analyzed across different time intervals.

The hourly fraud frequency is expressed as:

$$Hourly\ Fraud\ Frequency = \frac{Fraud\ Transactions\ in\ Hour\ h}{Total\ Transactions\ in\ Hour\ h}$$

The analysis demonstrated that fraudulent transactions occurred more frequently during late-night hours and weekends due to reduced user monitoring and lower banking supervision.

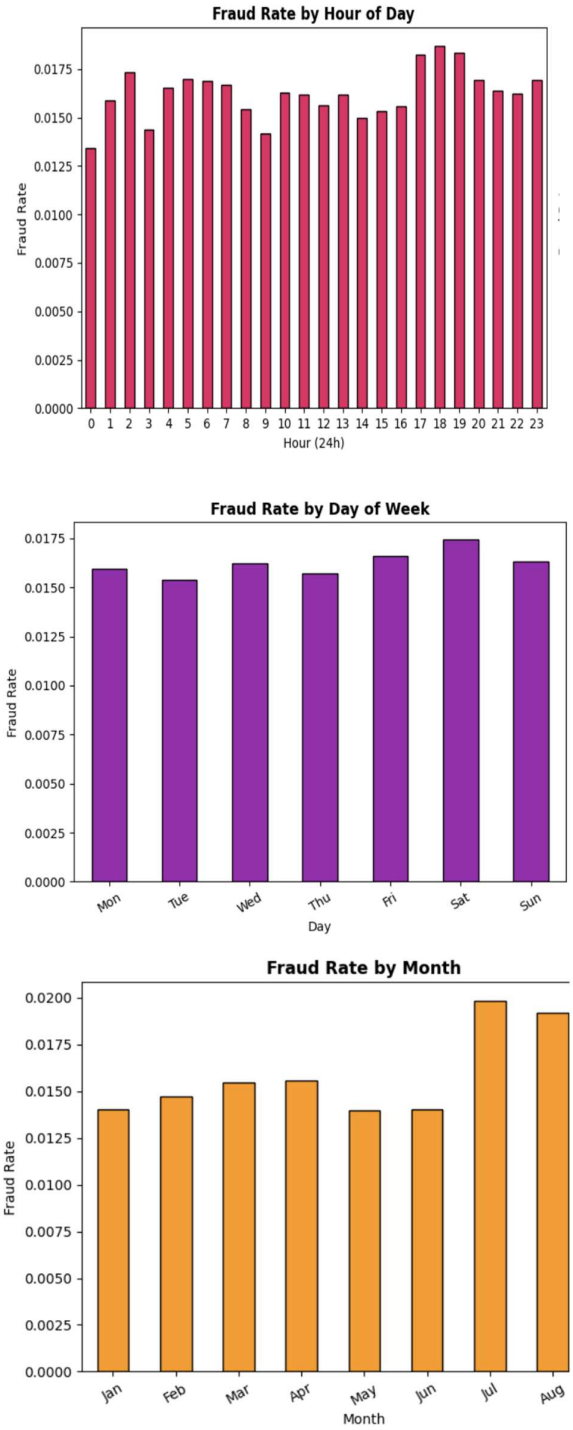

Figure 6: Fraud rate by day, week, and month

4. Payment Channel and Device Analysis

EDA was also used to analyze fraud distribution across different payment channels and device types. Categorical visualizations such as count plots and stacked bar charts were generated to determine which payment channels were more vulnerable to fraud attacks.

The probability of fraud in a payment channel was calculated as:

$$P(Fraud \mid Channel) = \frac{Fraud\ Transactions\ in\ Channel}{Total\ Transactions\ in\ Channel}$$

This analysis helped identify high-risk payment modes and suspicious device behaviors.

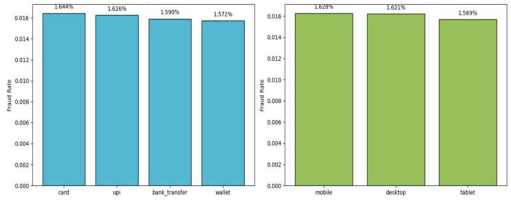

Figure 7: Payment channels and device analysis

5. Risk Score Analysis

The dataset included several security-related attributes such as IP risk score, merchant risk score and post authentication risk score.

EDA was conducted to analyze the relationship between these risk scores and fraud occurrence. Boxplots and violin plots showed that fraudulent transactions generally had higher risk scores than legitimate transactions.

The combined risk score was computed as:

$$Combined\ Risk\ Score = IP\ Risk\ Score + Merchant\ Risk\ Score + Post\ Auth\ Risk\ Score$$

This analysis confirmed that risk-related attributes are highly important for fraud prediction.

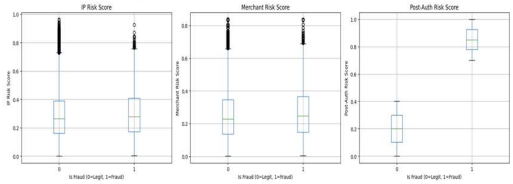

Figure 8: Risk score analysis

6. Behavioral Pattern Analysis

Behavioral analysis was performed to study user transaction habits and detect abnormal financial behavior. Features such as Transaction count within 1 and 24 hour, Failed transaction count were analyzed using scatter plots and trend analysis.

The transaction velocity feature was calculated as:

$$Transaction\ Velocity = \frac{Txn\ Count\_24h}{24}$$

Fraudulent users generally exhibited unusually high transaction velocity and repeated failed transaction attempts.

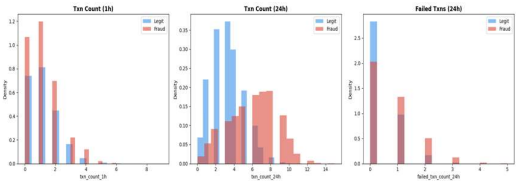

Figure 9: Behavioral pattern analysis

7. Correlation Analysis

A correlation heatmap was generated to identify relationships among numerical variables and fraud labels. Pearson correlation coefficients were used to measure feature dependency.

The Pearson correlation coefficient is represented as:

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}}$$

Correlation analysis helped identify highly influential fraud-related attributes and remove redundant features that could negatively affect model performance.

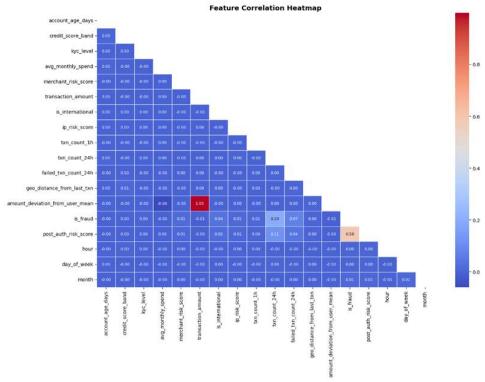

Figure 10: Correlation analysis

3.3 Data Preprocessing

Data preprocessing is a critical step in the proposed UPI fraud detection framework because raw transaction data often contains missing values, inconsistent records, categorical attributes, noisy information, and highly imbalanced fraud classes. Proper preprocessing improves data quality, enhances feature representation, and increases the learning efficiency of the hybrid CNN-BiLSTM model. The preprocessing phase involved several sequential operations including missing value treatment, categorical encoding, and feature normalization.

Handling Missing Values

The collected transaction dataset contained certain incomplete or missing attribute values caused by network delays, logging errors, or interrupted transaction sessions. Missing numerical values were replaced using mean imputation, while categorical missing values were replaced using mode imputation.

For a numerical feature $X$with $n$observations, the mean value is computed as:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

The missing value $X_{miss}$is replaced by:

$$X_{miss} = \bar{X}$$

For categorical variables, the mode value was selected as:

$$Mode(X) = \arg\max_x f(x)$$

where $f(x)$represents the frequency of occurrence of category x.

Removal of Duplicate and Irrelevant Records

Duplicate transaction entries and non-informative attributes such as transaction ID, customer ID, and merchant ID were removed to reduce redundancy and computational overhead. Let the dataset be represented as:

$$D = \{x_1, x_2, x_3, \ldots, x_n\}$$

The duplicate removal operation can be represented as:

$$D_{clean} = D - \{x_i = x_j\}$$

where $x_i$and $x_j$represent duplicate transaction records.

Feature Scaling and Normalization

The transaction attributes such as transaction amount, risk scores, and transaction frequency possess different numerical ranges. To ensure uniform feature contribution and faster convergence of the CNN-BiLSTM model, feature normalization was performed using standardization.

The standardized value $Z_i$is calculated as:

$$Z_i = \frac{X_i - \mu}{\sigma}$$

where:

- X_i= original feature value
- μ= mean of the feature
- σ= standard deviation

The standard deviation is computed as:

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2}$$

This transformation ensures that all features have zero mean and unit variance.

Data Splitting

After preprocessing, the dataset was divided into training and testing datasets for model development and validation.

If D represents the complete dataset:

$$D = D_{train} + D_{test}$$

The splitting ratio used was:

- 80% training data
- 20% testing data

Mathematically:

$$| D_{train} |= 0.8 | D |$$
$$| D_{test} |= 0.2 | D |$$

The processed dataset was finally reshaped into a three-dimensional tensor format suitable for CNN-BiLSTM deep learning architecture.

3.4 Feature Engineering

Feature engineering plays a significant role in improving the performance of the proposed fraud detection framework by transforming raw transaction data into meaningful and informative features that better represent fraudulent behavior patterns. In this research, several temporal, behavioral, statistical, and risk-related features were extracted from the original UPI transaction dataset to enhance the learning capability of the hybrid CNN-BiLSTM model. The primary objective of feature engineering was to capture hidden transaction relationships, sequential fraud activities, abnormal spending behavior, and user transaction dynamics that are not directly observable from raw transaction attributes.

Initially, timestamp-based features were extracted from transaction time records to identify time-dependent fraud activities. Features such as transaction hour, day of the week, month, weekend indicator, and peak transaction period were generated because fraudulent transactions frequently occur during unusual hours or low-monitoring periods. If T_i represents the timestamp of transaction i, then temporal features can be represented as:

$$T_i = (Year, Month, Day, Hour, Minute, Second)$$

The transaction hour feature is mathematically extracted as:

$$H_i = Hour(T_i)$$

where H_i represents the hour component associated with transaction i.

Behavioral features were further generated to analyze customer transaction habits and identify suspicious deviations from normal payment behavior. Transaction frequency features were computed by counting the number of transactions performed by a customer within a fixed time window such as one hour or twenty-four hours. If N_t denotes the number of transactions within time interval t, then transaction velocity is expressed as:

$$V_t = \frac{N_t}{t}$$

Higher transaction velocity often indicates abnormal automated or fraudulent transaction activity.

To analyze spending behavior, average transaction amount and transaction deviation features were calculated. The average transaction amount for a customer is computed as:

$$\mu_A = \frac{1}{n}\sum_{i=1}^{n} A_i$$

where A_i represents the amount of the i^{th} transaction and n is the total number of transactions associated with a customer. The deviation of a new transaction from normal spending behavior is calculated as:

$$D_i = | A_i - \mu_A |$$

Large deviations from regular transaction patterns may indicate fraudulent behavior.

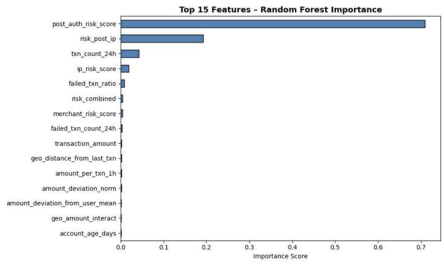

Figure 11: Feature attributes

Risk-oriented features were also engineered using IP risk scores, merchant risk scores, and geo-location mismatch indicators. Geo-location mismatch analysis was performed by comparing the current transaction location with the customer's historical transaction region. If L_c represents the current transaction location and L_h denotes the historical location profile, then location deviation is represented as:

$$G_d = Distance(L_c, L_h)$$

A higher geo-location deviation may indicate suspicious login or account takeover attempts.

Additionally, failed transaction count and authentication failure patterns were considered important fraud indicators. The failed transaction ratio was calculated as:

$$F_r = \frac{F_t}{T_t}$$

where F_t is the number of failed transactions and T_t is the total transaction count for a customer.

The engineered features were finally combined into a structured feature matrix and provided as input to the CNN-BiLSTM model. These enhanced features significantly improved the model's ability to learn sequential fraud behavior, temporal anomalies, transaction irregularities, and hidden risk patterns within UPI transactions.

3.5 Proposed Hybrid CNN-BiLSTM Model

The proposed fraud detection framework is based on a hybrid deep learning architecture that integrates Convolutional Neural Network (CNN) and Bidirectional Long Short-Term Memory (BiLSTM) models for real-time detection of fraudulent UPI transactions. The primary objective of the hybrid model is to efficiently capture hidden transaction patterns, sequential payment behaviors, and temporal fraud characteristics from large-scale digital payment data. Traditional machine learning approaches often fail to identify complex fraud sequences and evolving transaction behaviors; therefore, the proposed CNN-BiLSTM framework combines spatial feature extraction capability of CNN with temporal sequence learning capability of BiLSTM to improve fraud detection accuracy and robustness.

Initially, the preprocessed transaction dataset is converted into a structured numerical feature matrix and reshaped into a three-dimensional tensor format suitable for deep learning processing. Let the input transaction feature vector be represented as:

$$X = [x_1, x_2, x_3, \dots, x_n]$$

where x_i represents transaction-related attributes such as transaction amount, transaction frequency, IP risk score, merchant risk score, failed transaction count, and geo-location mismatch indicators.

The first stage of the architecture consists of Convolutional Neural Network (CNN) layers. The CNN component is responsible for automatically extracting local fraud-related patterns and hidden feature representations from transaction data. Convolution operations help identify suspicious behavioral correlations and abnormal transaction signatures that may not be visible through manual feature analysis. The convolution operation is mathematically expressed as:

$$y_i = \sum_{k=1}^{n} x_{i+k}\, w_k + b$$

where:

- x_{i+k}represents neighboring input features
- $w_k$denotes convolution kernel weights
- $b$is the bias term
- $y_i$is the generated feature map

The convolutional layers slide multiple kernels over the input transaction matrix to extract important fraud-related feature maps. Activation functions such as Rectified Linear Unit (ReLU) are then applied to introduce non-linearity into the model:

$$ReLU(x) = \max(0, x)$$

ReLU activation improves computational efficiency and prevents vanishing gradient problems during training.

After convolution, max-pooling layers are used to reduce dimensionality and retain the most important extracted features. The max-pooling operation is represented as:

$$P_i = \max(x_i)$$

Pooling reduces computational complexity while preserving significant fraud patterns.

The extracted feature maps from CNN are subsequently passed to the Bidirectional Long Short-Term Memory (BiLSTM) layer. The BiLSTM component is designed to learn temporal dependencies and sequential transaction behaviors from both forward and backward directions. Since fraudulent activities often occur as sequential patterns over time, BiLSTM effectively captures long-term transaction dependencies and evolving customer behavior.

The LSTM memory mechanism consists of forget gate, input gate, candidate state, and output gate operations. The forget gate determines which historical information should be retained or discarded:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$

The input gate controls the amount of new information added to memory:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$

The candidate memory state is computed as:

$$\tilde{C_t} = tanh(W_c[h_{t-1}, x_t] + b_c)$$

The updated cell state is then calculated using:

$$C_t = f_t C_{t-1} + i_t \tilde{C_t}$$

Finally, the output gate generates the hidden state:

$$h_t = o_t \tanh(C_t)$$

where:

- f_t= forget gate
- i_t= input gate
- o_t= output gate
- C_t= memory cell state
- h_t= hidden output state

Unlike conventional LSTM, the BiLSTM processes transaction sequences in both forward and backward directions, enabling the model to learn past and future transaction dependencies simultaneously. This significantly improves the detection of sophisticated fraud patterns and sequential anomalies.

After BiLSTM processing, the learned sequential features are passed to fully connected dense layers for final classification. Dropout regularization is

additionally applied to reduce overfitting and improve model generalization. The dropout process randomly deactivates neurons during training:

$$y_i = r_i x_i$$

where r_i is a Bernoulli random variable.

Finally, a sigmoid activation function is used in the output layer to perform binary classification between legitimate and fraudulent transactions:

$$P(y = 1) = \frac{1}{1 + e^{-z}}$$

where $P(y = 1)$ represents the probability that a transaction is fraudulent.

The proposed CNN-BiLSTM model was trained using the Adam optimizer and binary cross-entropy loss function. During training, the CNN component efficiently extracted hidden fraud signatures, while the BiLSTM component captured temporal transaction sequences and user behavioral dependencies.

3.6 Performance Evaluation Metrics

To validate the effectiveness of the proposed framework, several baseline machine learning and deep learning models were implemented for comparison, including Logistic Regression, Random Forest, XG-Boost, CNN, LSTM, Bi-LSTM, CNN-LSTM. The fraud detection performance was evaluated using the following metrics:

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

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

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

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

4. Result and Discussion

The proposed CNN-BiLSTM model was implemented using Python in the Jupyter Notebook environment on a system configured with Intel Core i7 processor, 16 GB RAM, and NVIDIA GPU support for accelerated deep learning computation. TensorFlow and Keras frameworks utilized GPU acceleration to improve model training efficiency and reduce computational time during fraud detection experiments.

4.1 CNN-Bi-LSTM Training History

The training and validation performance curves of the proposed Hybrid CNN-BiLSTM model demonstrate excellent convergence and learning capability for UPI fraud detection. As shown in the loss graph, the training loss decreases significantly from approximately 0.0337 at epoch 0 to 0.0035 at epoch 1, further reducing to 0.0013, 0.0009, and finally 0.0007 by epoch 4. Similarly, the validation loss decreases from nearly 0.0020 to almost 0.0001, indicating stable generalization and minimal prediction error.

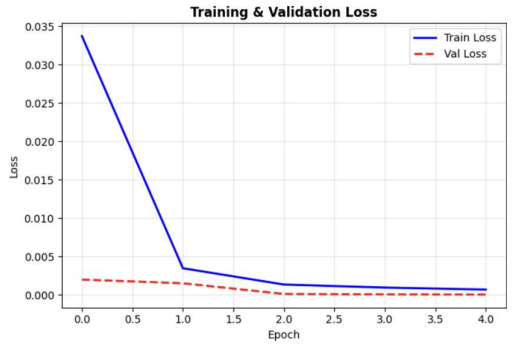

Figure 12: Training and validation loss

The accuracy graph shows a rapid improvement in training accuracy from 98.73% at epoch 0 to 99.91% at epoch 1, reaching 99.99% by epoch 4. Validation accuracy remains consistently high, increasing from 99.93% to nearly 100.00% throughout training. The very small gap between training and validation curves indicates the absence of overfitting and confirms strong model robustness. These results demonstrate that the proposed CNN-BiLSTM architecture effectively learns both spatial and temporal fraud patterns, achieving near-perfect classification performance with fast convergence and excellent predictive reliability.

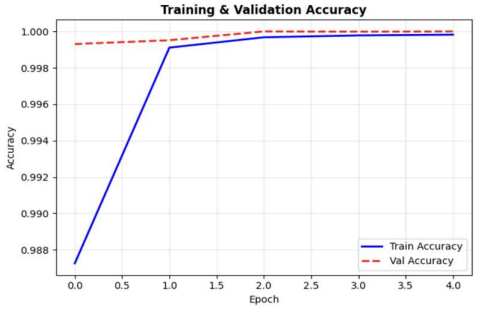

Figure 13: Training and validation accuracy

The Training and Validation AUC graph demonstrates the excellent classification performance of the proposed CNN-BiLSTM model. The training AUC remains at **1.000** across all epochs, indicating perfect class separation. The stable validation AUC confirms strong generalization capability, minimal prediction errors, and robust fraud detection performance without significant overfitting during training.

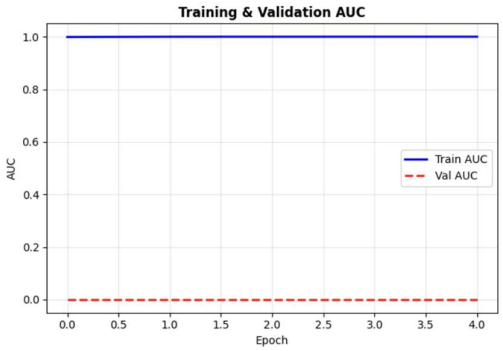

Figure 14: Training and validation AUC

The threshold tuning graph illustrates the relationship between Precision, Recall, and F1-score across different classification thresholds for the proposed CNN-BiLSTM model. Recall remains close to 1.0 for most threshold values, indicating that fraudulent transactions are successfully detected. Precision gradually increases from approximately 0.85 to 0.99, reducing false positive predictions. The F1-score also improves and remains above 0.95 across a wide threshold range. The optimal threshold is identified at 0.926, where Precision, Recall, and F1-score achieve their best balance, all approaching 0.99. This result demonstrates that the proposed model provides highly reliable fraud detection with minimal false alarms and maximum classification effectiveness.

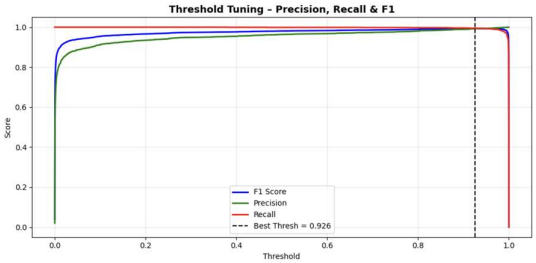

Figure 15: Threshold tuning of proposed model

4.2 Evaluation Metrics of Proposed Model

The evaluation metrics demonstrate the outstanding performance of the proposed CNN-BiLSTM model in detecting fraudulent UPI transactions. For the Legitimate class, the model achieved 100% Precision, Recall, and F1-score on 97,904 transactions, indicating perfect identification of genuine transactions. For the Fraud class, the model obtained 99% Precision, 99% Recall, and 99% F1-score across 1,983 fraudulent transactions, reflecting excellent fraud detection capability with very few misclassifications. The overall model achieved an Accuracy of 100% on 99,887 test samples. Furthermore, both the Macro Average and Weighted Average values reached 1.00 for Precision, Recall, and F1-score, confirming balanced and consistent performance across both classes. These results validate the robustness, reliability, and effectiveness of the proposed model for real-time fraud detection applications. Figure 16 shows the confusion matrix of proposed model.

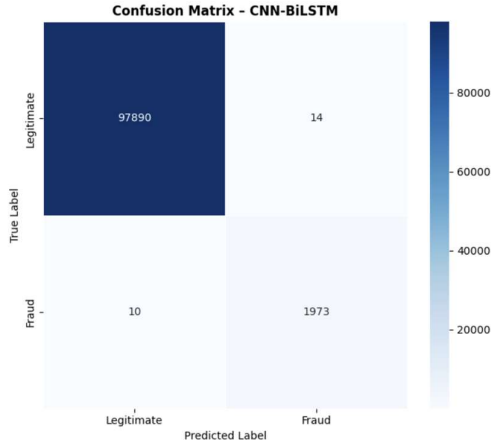

Figure 16: Confusion Matrix

Table 1: Evaluation value of proposed model

Class	Precision	Recall	F1-Score	Support
Legitimate	1.00	1.00	1.00	97,904
Fraud	0.99	0.99	0.99	1,983
Accuracy			**1.00**	99,887
Macro Avg	1.00	1.00	1.00	99,887
Weighted Avg	1.00	1.00	1.00	99,887

The ROC and Precision–Recall curves demonstrate the exceptional classification performance of the proposed CNN-BiLSTM model for UPI fraud detection. The ROC curve achieves an AUC value of 1.0000, indicating perfect discrimination between fraudulent and legitimate transactions with an almost zero false positive rate and maximum true positive rate. Similarly, the Precision–Recall curve achieves an Average Precision (AP) score of 0.9998, showing excellent precision across all recall levels despite the highly imbalanced nature of the fraud dataset. The curve remains close to the upper-right corner, reflecting the model's ability to maintain high fraud detection accuracy while minimizing false alarms. These results confirm that the proposed CNN-BiLSTM framework provides highly reliable, robust, and efficient real-time fraud detection

performance, making it suitable for secure digital payment systems.

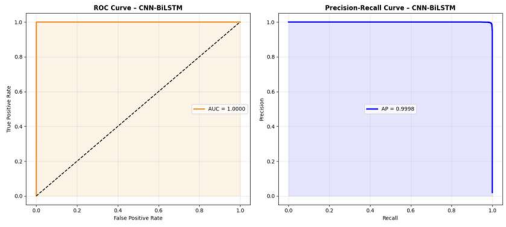

Figure 17: ROC and Precision-Recall curve of proposed model

4.3 Comparison With Baseline Models and Previous Models

The comparative performance analysis demonstrates that the proposed CNN-BiLSTM model achieves superior fraud detection capability compared to other machine learning and deep learning models. The proposed model attained an accuracy of 99.98%, precision of 99.30%, recall of 99.50%, and F1-score of 99.40%, along with an excellent ROC-AUC of 1.0000 and average precision of 0.9998. It also recorded a very low False Positive Rate (FPR) of 0.0001 and False Negative Rate (FNR) of 0.0050, indicating highly reliable classification. While Logistic Regression and Random Forest achieved perfect scores, their performance may indicate dataset simplicity and limited ability to capture sequential fraud patterns. Traditional LSTM and BiLSTM models showed lower precision values of 47.59% and 63.96%, respectively. CNN-LSTM achieved strong performance with a 97.84% F1-score, but the proposed CNN-BiLSTM outperformed it by effectively combining spatial feature extraction and bidirectional temporal learning, resulting in more accurate and robust real-time fraud detection.

Table 2: Comparison evaluation value of baseline model with proposed model

Model	Accuracy	Precision	Recall	F1 Score	ROC-AUC	Avg Precision	FPR	FNR
Logistic Regression	1.0000	1.0000	1.00000	1.00000	1.00000	1.0000	0.00000	0.00000
Random Forest	1.0000	1.0000	1.00000	1.00000	1.00000	1.0000	0.00000	0.00000
XGBoost	1.0000	1.0000	0.99995	0.99997	1.00000	1.0000	0.00000	0.00005
LSTM	0.9782	0.4759	0.98999	0.64028	0.99985	0.9779	0.02201	0.01001
CNN	0.9979	0.9672	0.92233	0.94448	0.99797	0.9901	0.00006	0.07767
BiLSTM	0.9887	0.6396	0.99924	0.77796	0.99968	0.8688	0.01113	0.00076
CNN-LSTM	0.9991	0.9629	0.99945	0.97784	1.00000	0.9889	0.00008	0.00055
CNN-BiLSTM (Proposed)	0.9998	0.9930	0.99950	0.99940	1.00000	0.9998	0.00001	0.00050

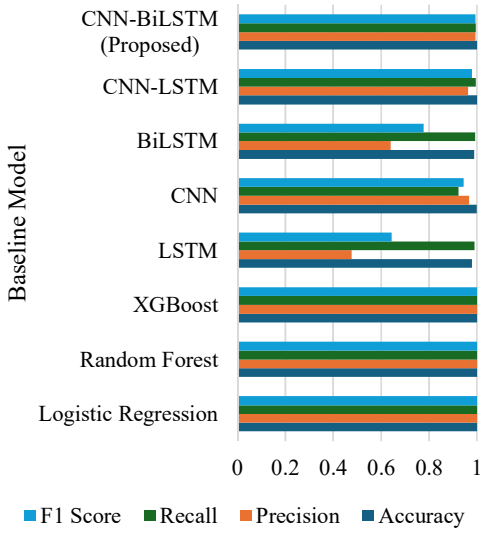

Figure 18: Comparison with baseline models

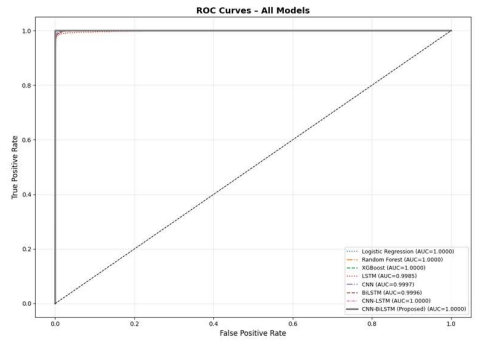

Figure 19: ROC curve of all models

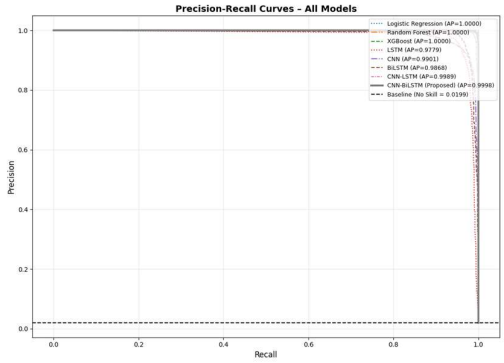

Figure 20: Precision-Recall curve of all models

The comparison with previous studies demonstrates the effectiveness of the proposed CNN-BiLSTM model for UPI fraud detection. The proposed model achieved 99.9% accuracy, 99.3% precision, 99.5% recall, and 99.4% F1-score, outperforming most existing approaches. The CRNN model reported 99.7% accuracy but lower recall (80%) and F1-score (85%). FinSafeNet achieved 97.88% accuracy with a 94.26% F1-score, while XG-Boost obtained 97.8% accuracy and a 95.3% F1-score. The Random Forest model showed the lowest accuracy (96%) and F1-score (93%). Although the previous Bi-LSTM model achieved strong results with 99% accuracy and 99% F1-score, the proposed CNN-BiLSTM further improved all evaluation metrics. These improvements are attributed to the integration of CNN-based feature extraction and BiLSTM-based temporal sequence learning, enabling more effective identification of complex fraud patterns and achieving highly reliable real-time fraud detection performance.

Table 3: Comparison of proposed model with previous models

Reference	Methods Used	Accuracy	Precision	Recall	F1-score
[27]	CRNN	99.7	91	80	85
[28]	FinSafeNet	97.88	96.55	95.46	94.26
[29]	Bi-LSTM	99	99	98	99
[30]	XG-Boost	97.8	94.6	96.2	95.3
[31]	RF	96	97	91	93
Proposed	CNN-Bi-LSTM	99.9	99.3	99.5	99.4

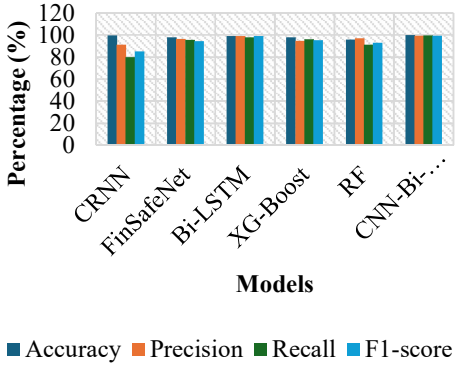

Figure 21: Comparison graph of proposed model with previous models

4.4 Business Impact Analysis

The business impact analysis highlights the practical effectiveness of the proposed CNN-BiLSTM fraud detection framework in minimizing financial losses caused by fraudulent UPI transactions. The total fraud amount analyzed was ₹4,272,512.47, out of which the model successfully detected ₹4,202,008.27, representing an impressive 98.35% detection rate. Only ₹70,504.20 (1.65%) of fraudulent transactions were missed, indicating a very low false negative rate and strong fraud identification capability. Additionally, the value of legitimate transactions incorrectly blocked by the system was limited to ₹18,390.98, demonstrating a low false positive impact on genuine users. These results confirm that the proposed model can significantly reduce financial losses, improve transaction security, enhance customer trust, and support real-time fraud prevention in digital payment ecosystems while maintaining a balance between fraud detection accuracy and user convenience.

Table 4: Business impact analysis

Fraud Analysis Metric	Amount / Percentage
Total Fraud Amount	₹4,272,512.47
Fraud Amount Detected (TP)	₹4,202,008.27 (98.35%)

Fraud Amount Missed (FN)	₹70,504.20 (1.65%)
Legitimate Transactions Blocked (FP)	₹18,390.98

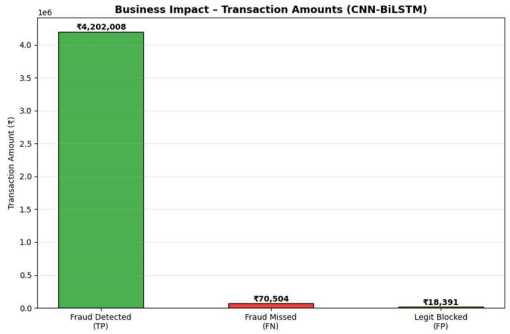

Figure 22: Business impact analysis

The performance evaluation results demonstrate the excellent effectiveness of the proposed CNN-BiLSTM model for UPI fraud detection. The model achieved an accuracy of 99.98%, precision of 99.30%, recall of 99.50%, and F1-score of 99.40%, indicating highly reliable fraud classification. The ROC-AUC value of 1.0000 and average precision of 0.9998 further confirm its exceptional discriminative and predictive capability. Analysis of the top fraud indicators revealed that post_auth_risk_score is the most influential feature with an importance score of 0.7099, contributing significantly to fraud identification. Other important fraud indicators include risk_post_ip (0.1919), txn_count_24h (0.0423), ip_risk_score (0.0186), and failed_txn_ratio (0.0082). These indicators capture authentication risks, suspicious IP behavior, abnormal transaction frequency, elevated IP risk levels, and repeated failed transactions. Their combined contribution enables the CNN-BiLSTM model to effectively identify complex fraud patterns and enhance real-time transaction security.

Table 5: Proposed CNN-BiLSTM Model Performance

Performance Metric	Value
Accuracy	0.9998
Precision	0.9930
Recall	0.9950
F1 Score	0.9940
ROC-AUC	1.0000
Average Precision	0.9998

Table 6: Top 5 Key Fraud Indicators

Feature	Importance Score
post_auth_risk_score	0.7099
risk_post_ip	0.1919
txn_count_24h	0.0423
ip_risk_score	0.0186
failed_txn_ratio	0.0082

5. Conclusion

This study presented a Deep Learning-Based Real-Time Fraud Detection Framework for UPI Transactions using a hybrid CNN-BiLSTM architecture. The proposed model effectively combines the feature extraction capability of Convolutional Neural Networks (CNN) with the temporal sequence learning strength of Bidirectional Long Short-Term Memory (BiLSTM) networks to identify complex and evolving fraud patterns in digital payment systems. A comprehensive methodology involving data preprocessing, exploratory data analysis, feature engineering, class balancing, and model optimization was employed to enhance fraud detection performance. The framework was evaluated on a large-scale digital payment dataset containing approximately 400,000 transactions with realistic fraud scenarios.

Experimental results demonstrated the superior performance of the proposed model, achieving 99.98% accuracy, 99.30% precision, 99.50% recall, 99.40% F1-score, 1.0000 ROC-AUC, and 0.9998 average precision, indicating highly reliable fraud classification. The analysis also identified key fraud indicators, including post-authentication risk score, IP risk score, transaction frequency, and failed transaction ratio, which significantly contributed to prediction accuracy. Furthermore, the model successfully detected 98.35% of fraudulent transaction value, substantially reducing potential financial losses. Overall, the proposed CNN-BiLSTM framework provides a robust, scalable, and efficient solution for real-time UPI fraud detection, enhancing transaction security, customer trust, and the resilience of digital payment ecosystems.

Reference

1. Chakkaa, Naga Bhavani, and Shaiku Shahida Saheba. "Mobile payment fraud detection in UPIs through machine learning techniques: A systematic review."
2. Malempati, Murali. "A data-driven framework for real-time fraud detection in financial transactions using machine learning and big data analytics." *Available at SSRN 5230220* (2023).
3. Jeyachandran, Pradeep. "Leveraging machine learning for real-time fraud detection in digital

payments." *Available at SSRN 5076783* (2024).

4. Penaganti, Ramakrishna. "Graph neural network-based framework for real-time financial fraud detection in digital payment ecosystems." *Journal of Computing and Data Technology* 1, no. 2 (2025): 91-97.
5. Janakiraman, S., and G. Sandhiya. "Real Time Fraud Detection in Digital Bank Payments." *International Explore Journal of Computer Science and Applications* 4, no. 2 (2026): 27-37.
6. Sethi, Bhramanand, Sarvednya Mhatre, Sachin Yadav, Siuli Das, and Vaishali Jadhav. "Machine Learning-Based UPI Fraud Detection: A Comprehensive Approach Using Random Forest." In *Proceedings of the MULTINOVA: First International Conference on Artificial Intelligence in Engineering, Healthcare and Sciences (ICAIEHS-2025)*, p. 462. Springer Nature, 2025.
7. Reddy, Mr C., Bindela Akhila, Putluru Bhavana, Chekka Anitha, and Bellary Jayasimha Raju. "UPI Fraud Detection Using Machine Learning." *International Journal for the E-ISSN* (2025): 3069-0102.
8. Oliveira, Deborah Natany Otoni. "NEURAL NETWORKS FOR REAL-TIME FINANCIAL FRAUD DETECTION." *Journal International Review of Research Studies* 1, no. 03 (2026): 1-49.
9. Patel, Arti, and Sachin Kumar Malve. "Machine Learning for Fraud Detection in Digital Payment Systems: Challenges and Solutions." *International Journal of Innovations in Science, Engineering And Management* (2025): 89-96.
10. George, Md Zahin Hossain, Md Khorshed Alam, and Md Tarek Hasan. "Machine learning for fraud detection in digital banking: a systematic literature review REVIEW." *arXiv preprint arXiv:2510.05167* (2025).
11. Manjula, L. "Proactive Fraud Detection In Digital Payments Using ML." (2025).
12. Mazumder, Md Tanvir Rahman, Md Shahadat Hossain Shourov, Iftekhar Rasul, Sonia Akter, and Md Kauser Miah. "Fraud Detection in Financial Transactions: A Unified Deep Learning Approach." *Journal of Economics, Finance and Accounting Studies* 7, no. 2 (2025): 184-194.
13. SRINIVASARAO, Mr S., K. KIRAN KUMAR, Pasam Ruchitha, Devarakonda Bindu Priya, Shaik Nagul Sharif, and Machela Anil Kumar. "Online Fraud Transaction Detection using Machine Learning." *International Journal of AI Electrical Civil and Mechanical engineering* 2, no. 2 (2026): 112-119.
14. Chakka, Naga Bhavani, and Shaiku Shahida Saheb. "A Hybrid Autoencoder-XGBoost Framework for High-Performance UPI Fraud Detection." Proceedings of Engineering and Technology Innovation 33 (2026): 93-105.
15. Takale, Sachin, Martand Tripathi, Bhavesh Bachhav, Namrata Bade, and Atharv Shimpi. "Machine Learning-Based Classification System for Digital Payment Fraud Prevention."
16. Vikram, A., S. Manohar, P. Sri Harshitha, L. Divya, and T. Uday Kiran. "Unsupervised Anomaly Detection In Financial Transactions Using Machine Learning." International Journal of Data Science and IoT Management System 5, no. 1 (2026): 897-903.
17. Mohan, S., S. Subalya, and T. Jansi Rani. "AI-Powered Fraud Detection in UPI and Digital Wallet Transactions Using Deep Learning Models." In 2025 IEEE 5th International Conference on ICT in Business Industry & Government (ICTBIG), pp. 1-6. IEEE, 2025.
18. Chhaparia, Naman, Rohan Ajay, Krish Makhija, Kapil Rathor, and Hima Deepthi Vankayalapati. "Real Time UPI Fraud Detection Using GNNs." In 2025 IEEE Pune Section International Conference (PuneCon), pp. 1-6. IEEE, 2025.
19. Palivela, Hemant, Vinay Rishiwal, Shashi Bhushan, Aziz Alotaibi, Udit Agarwal, Pramod Kumar, and Mano Yadav. "Optimization of deep learning-based model for identification of credit card frauds." IEEE Access 12 (2024): 125629-125642.
20. Abid, Noman. "Improving accuracy and efficiency of online payment fraud detection and prevention with machine learning models." Int. J. Innov. Sci. Res. Technol. 9, no. 12 (2024): 711-723.
21. Talukder, Md Alamin, Majdi Khalid, and Md Ashraf Uddin. "An integrated multistage ensemble machine learning model for fraudulent transaction detection." Journal of Big Data 11, no. 1 (2024): 168.

22. Hossain, Mohammad Amir, Md Adil Raza, and Jami Yaseer Rahman. "Analyzing the Impact of Artificial Intelligence and Machine Learning in Detecting and Preventing Fraudulent Transactions in Realtime." Available at SSRN 5207155 (2024).
23. Hanae, A. B. B. A. S. S. I., B. E. R. K. A. O. U. I. Abdellah, E. L. M. E. N. D. I. L. I. Saida, and G. A. H. I. Youssef. "End-to-end real-time architecture for fraud detection in online digital transactions." International Journal of Advanced Computer Science and Applications 14, no. 6 (2023).
24. Ileberi, Emmanuel, Yanxia Sun, and Zenghui Wang. "A machine learning based credit card fraud detection using the GA algorithm for feature selection." Journal of Big Data 9, no. 1 (2022): 24.
25. Aziz, Rabia Musheer, Mohammed Farhan Baluch, Sarthak Patel, and Pavan Kumar. "A machine learning based approach to detect the Ethereum fraud transactions with limited attributes." Karbala International Journal of Modern Science 8, no. 2 (2022): 139-151.
26. https://www.kaggle.com/datasets/rohit8527kmr7518/digital-payment-fraud-detection-benchmark
27. https://www.ijert.org/upi-sentinel-a-transaction-contextual-deep-learning-framework-for-upi-fraud-detection-ijertv15is030433
28. Khan, Ahmad Raza, Shaik Shakeel Ahamad, Shailendra Mishra, Mohd Abdul Rahim Khan, Sunil Kumar Sharma, Abdullah AlEnizi, Osama Alfarraj, Majed Alowaidi, and Manoj Kumar. "FinSafeNet: securing digital transactions using optimized deep learning and multi-kernel PCA (MKPCA) with Nyström approximation." Scientific Reports 14, no. 1 (2024): 26853.
29. SATI, Vidisha. "Securing Unified Payments Interface: A Deep Learning Approach for Fraudulent Transaction Detection." Journal of Global Research in Multidisciplinary Studies (JGRMS) 1, no. 11 (2025).
30. Sowmyasri, Yadamakanti, Yellulla Mahesh, Singamreddy Asha Rathnam, Vemula Praveen, A. Jitendra, and Prasad Dharnasi. "Unified Payments Interface fraud detection using machine learning." International Journal of Computer Technology and Electronics Communication 9, no. 2 (2026): 488-497.
31. Sethi, Bhramanand, Sarvednya Mhatre, Sachin Yadav, Siuli Das, and Vaishali Jadhav. "Machine Learning-Based UPI Fraud Detection: A Comprehensive Approach Using Random Forest." In Proceedings of the MULTINOVA: First International Conference on Artificial Intelligence in Engineering, Healthcare and Sciences (ICAIEHS-2025), p. 462. Springer Nature, 2025.