

UPI Fraud Detection Using Ensemble Approach

Rebeca K. Yeptho¹, Bendangmenla², Monita Wahengbam³, L Lanuwabang^{4 1-4}

National Institute of Electronics & Information Technology (NIELIT), Kohima

rebecayeptho@nielit.ac.in1, bendang@nielit.ac.in2, monita@nielit.gov.in3, lanu@nielit.gov.in4

Abstract: Unified Payment Interface (UPI) has paved the way for convenient digital transaction and has enhanced the speed of digital transaction. Users can simply use the technology to make online payments at an accurate speed through mobile number, Virtual Payment Address (VPA) or QR code. But with the evolving technologies and advanced system, there has been cases where fraudulent activities are encountered leading to financial risk and losses. In order to tackle this problem, this research develops an ensemble RNN and XGBoost system for detecting UPI fraud by examining the patterns to recognize and alert on potentially fraudulent activities. It also encompasses data cleaning and preparation, and training of predictive models. Dataset from open source has been trained and tested against the Ensemble RNNBoost (Ensemble RNN and XGBoost) model and has been compared with various ML models. The model's performance was assessed using various evaluation metrics, including F1-score, precision, recall, and accuracy. The main reason to combine XGBoost and RNN is to exploit sequential patterns and tabular feature interactions simultaneously, leading to overall performance and more robust predictions. As a result, RNNBoost model produce the maximum accuracy rate of 96.44% in comparison with other traditional models like DT (95.00%), RF (95.88%), LR (80.25%), KNN (83.00%), SVM (83.00%), NB (83.00%) and RNN (81.25%).

Keywords: Ensemble RNNBoost, Fraud detection, RNN, Transaction, UPI

(Article history: Selected papers from NIELIT's International Conference on Communication, Electronics and Digital Technologies (NICE-DT'26), 29-30Jan, 2026, Gorakhpur, India)

I. INTRODUCTION

Unified Payment Interface (UPI), developed in India, is a real-time payment platform that enables users to perform online transactions at an ease and faster rate. Users utilize various apps like Phonepay, Paytm, Gpay etc. to make digital payments through mobile number, QR code or Virtual Payment Address (VPA) enabling smoother transaction process without the need for bank details every time. But with evolving technologies and advanced system, there has been a rise in fraudulent activities leading to financial losses. Both the consumers and financial institutions fall prey to such risks and financial losses. Traditional models (DT, RF, LR, KNN, SVM, NB) are simple, fast and easier to implement but may not give the desired accuracy or higher efficiency. Traditional models are limited and some may underfit complex fraud patterns and give higher false positive rates.

The motive of this research is to reduce the risk of financial losses by developing an ensemble RNNBoost model that examines the transactional pattern to recognize and alert on potentially fraudulent activities. RNN, which is best for examining sequential or time-series based data extracts features like transaction timestamps, transaction frequency and user behavior patterns. XGBoost which is best for non-linear structured or tabular based data extracts features like transaction amount, merchant type, categorical and numerical data like DOB, zip code, state. The features extracted from both RNN and XGBoost are stacked and combined. In this model, RNN is used to examine the embedding from sequential transaction data, while XGBoost is used to leverage these embedding along with the tabular, structured data for final fraud prediction. Fraud datasets are highly imbalanced, but the ensemble model handles this imbalance better than the traditional model by using weighted loss functions (XGBoost) and capturing sequential anomalies (RNN). The ensemble model combines the strength of RNN's deep learning adaptability and XGBoost's interpretability making the detection of fraud patterns efficient. The ensemble model is useful in detecting temporal fraud patterns, handling user, merchant or tabular data with high accuracy in UPI transactions where fraudulent activities occur rapidly and the model is robust to mixed datatypes (text, numbers, categories). The model also detects complex and new fraud patterns with higher accuracy and efficiency as compared with other traditional models. Various performance metrics such as F1-score, precision, recall, and accuracy were computed from the confusion matrix to evaluate and obtain the desired results. By enhancing security and building user trust, this approach supports the long-term resilience and growth of digital financial ecosystems.

II. LITERATURE SURVEY

Recent research in UPI fraud detection, as outlined in this section, highlights the previous studies largely emphasizing the use of ML and DT approaches for detecting complex fraud patterns. Chen Y. et al. [1] conducted a study on the application of machine learning in financial anomaly detection, exploring supervised, unsupervised, and deep learning approaches. Their findings indicated that ensemble learning techniques excel at maintaining a balance between detection accuracy and model interpretability. The study also recommended future research focused on enhancing data sharing and model collaboration while ensuring the protection of personal privacy and proprietary information.

Abdulalem Ali et al [2] presented that SVM and ANN are popular fraud detection, particularly in credit card transactions, is one of the most common applications of machine learning algorithms. This study proposes that advanced text-mining and word-embedding methods, including Word2Vec, Doc2Vec, and BERT, can be employed to convert financial text data into feature vectors, which can then serve as input for building machine learning models. Palak Gupta et al [3] proposed that XGBoost classifier produced the best possible result among the following algorithm i.e LR, DT Classifier, Artificial Neural Network, XGBoost Classifier that were used for modelling for balancing the data.

Vaishali Gupta et al [4] Deployed an RNN-based LSTM model to a large dataset of UPI transactions which demonstrated the reliability and resilience in identifying fraudulent transaction in real-time. The proposed approach was applied to the test data on confusion matrix and got the TPR as 87.5% and the FPR as 13.4%. Vitthal B Kamble et al [5] used the combination of RF and SVM for complex Fraud Detection System. The use of feature augmentation and ensemble learning techniques helps address challenges such as overfitting and class imbalance, making this method well-suited for practical fraud detection applications.

Mr.C.Lakshminath Reddy1 et al [6] UPI Fraud Detecting Using Machine Learning examined the historical transaction data. The system aims to detect subtle indicators of fraudulent activity, thereby enhancing the security of the UPI transactions. The approach involves data collection and analysis, model development, system integration, and ongoing performance evaluation. Ashish Das et al [7] showcased a practical implementation of a smart fraud detection system for online payments. From design through deployment, it demonstrates strong fraud identification capabilities along with improved user engagement. By integrating ML for decision-making, Optical Character Recognition (OCR), and location tracking, the systems is particularly well-suited for UPI-driven platforms. Features like Real-time fraud alerts, geolocation tagging, and secure password enforcement enhance both the trustworthiness and digital transactions security.

Mrs K Amulya et al [8] proposed that the system effectively minimizes false positives and improves the fraud detection accuracy by leveraging advance machine learning models such as (DT), (RF), (GB), AdaBoost. It also incorporates unsupervised learning technique to detect anomalies and utilizes Deep Learning architecture like LSTMs and transformers to examine sequential transaction data. Additionally, an adaptive fraud detection framework is developed to continuously evolved and respond to new and emerging fraudulent behaviors. Dr C Nadhamuni Reddy et al [9] developed an effective system for fraud detection UPI transactions using a range of ML models. Among these, RF achieved the highest accuracy at 90.75%. Gradient Boosting, LightGBM, and XGBoost also demonstrated strong performance, making them feasible alternatives for detecting fraudulent activity. The results outline the significance of maintaining a balance between recall and precision to reduce false negatives and accurately identify fraudulent transactions.

Mr Khadir Kumar.N et al [10] highlighted the importance of Machine Learning methods in enhancing the effectiveness of detecting and preventing credit card fraud, as well as responding to fraudulent activities. The study leverages frameworks including KNN, AdaBoost, Naïve Bayes, and SVM-DT to evaluate long-term security datasets, facilitating fast and efficient handling of critical security alerts.

III. PROPOSED METHODOLOGY

A full-fledged fraud detection system in UPI transaction contains dataset that captures features required for anomaly detection. With the help of this dataset, the proposed model identifies fraudulent patterns originating from transactional data and user behavior, thus improving accuracy, as well as adaptability to emerging fraud methodologies. This dataset consists of attributes such as Transaction ID, Transaction hours, Transaction month, Transaction year, category, UPI number, age, transaction amount, state, zip, sender and receiver details, status and fraud risks. For training and evaluating the fraud detection system this dataset act as a base. With the samples data collected, to enhance their standard, the sample undergoes preprocessing. In the data preprocessing, missing or null values are handled through imputation or removal. Categorical variables (e.g. Sender and receiver IDs, transaction types) are encoded into numerical formats suitable for ML algorithms. Numerical features such as transaction amounts are normalized to standard scales, which helps in reducing bias towards larger values.

To enhance the model performance, the properties are extracted from the dataset. These features include transaction frequency i.e., number of transaction conducted by a user over a specific period, amount deviation that is the difference between a transaction amount and the users' typical transaction amounts and user behavioral patterns like temporal patterns, sender-receiver interactions and any anomalies in transaction timings. To refine the model to distinguish between the legitimate and fraudulent transactions, these features are used.

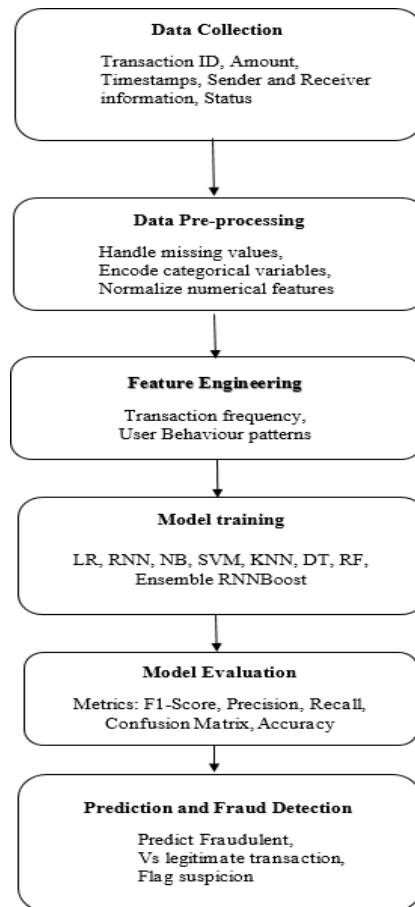


Fig.1. Flow diagram of the proposed methodology.

The preprocessed data is used to train ensemble RNNBoost. For training a total of 1813 dataset transactions is used and for testing 453 dataset transactions are being used. This hybrid model examines all the patterns from the labelled data in order to sort the transactions. After the data is trained and tested, the model predicts whether a new transaction ID is legitimate or fraudulent.

Once the model is trained, it detects any potential fraud transaction and immediately generates an alert. The end user or the system administrators receives the alert and it enables them to prevent from any financial loss by taking rapid action against such fraudulent activities. Hence different performance metrics are calculated using the confusion matrix to evaluate the model.

Decision tree algorithms are used widely in fraud detection because of its simplicity, interpretability, and ability to handle wide-ranging data types. Decision tree is a supervised learning model which is utilized for categorization as well as estimation of task making them flexible in assessing the financial transaction.

They are organized around the hierarchy of “if-else” conditional rules which makes the decision making understandable and user friendly for non-technical as well. Their ability to uncover masked relationship among the data is one of the key advantages of Decision tree which allows them to differentiate the fraudulent and genuine operations. They are often good at managing irregularities such as anomalies values and outliers which are often exist in real-world transaction dataset. Decision trees unlike other models does not presume a distinct data distribution and do not require quality scaling or standardization, which minimize normalization effort substantially.

Random Forest is a model which gives us a better prediction by integrating the output of many decision tree to produce better robustness and accuracy, thereby making it an ensemble learning technique. This algorithm creates many decision tree by taking random parts of data. Each tree will make its own prediction from part of data. Once prediction is done then it combines the prediction and choose the average one from all the tree as the final answer. Since each tree uses random parts of data and the

feature, it gives us better accuracy and even avoid overfitting. Models like Random Forest is often useful in UPI Fraud Detection as it focuses on factors like geographical location, transaction timing, transaction amount, device identifier which collectively contribute on differentiating legitimate ones from fraudulent one.

KNN also called as supervised learning model that is utilize for segmentation as well as continuous output prediction including the fraud detection. KNN compares the new data points to stored historical example instead of building a structured model during the training. For segmentation, the KNN assigned the most common label among the K-Closest neighbors, as for the prediction, the KNN takes the average value of those neighbors. By exploring their resemblance to the previously observed behaviors. This learning approach makes it useful for identifying the fraudulent transaction ID. One of the advantage of KNN is its resilience in dealing with varieties of variables like numerical and unambiguous without inferring certain principal of distribution in the data. Due to this reason it is compatible for financial dataset which has various properties like location, devices identifiers, the transaction values, frequency. One of the strength of KNN is that it can easily adjust from the stored example. Therefore, it does not need to retrained the model and can easily adapt to the recent fraud patterns.

The Support Vector machine (SVM) algorithm is utilize with the goal of segmentation as well as continues output prediction. It has also shown promising result in fraud detection tasks. The main concept of this supervised learning algorithm is to draw dividing line between the various classes so that the nearest point of all the classes are as far away as possible from the line so as to give us more authentic predictions.

In UPI Fraud Detection, we used this concept to actually differentiate the legitimate transaction ID from the fraudulent transaction ID by drawing a line and creating an edge in the feature space. SVM effectiveness is its robustness in handling the non-linear and linear data, because of which an input data can also be project into multivariate space where even a sophisticated relationship becomes easier to separate. Therefore, capturing the concealed pattern and delicate pattern that are there in fraudulent transaction. SVM is suited for UPI fraud detection because even if there are irrelevant and noise data present, the model works efficiently as it focuses on the data points which is of closest to the decision boundary.

The algorithm called Naïve Bayes (NB) model is mainly for classifying the tasks. This model is originated on Bayes' Theorem that states that all the properties are disassociated of each other despite the class label. As NB has the ability to handle corrupted data, it often delivers potential result. Since it is based on Bayes' Theorem the NB observed the feature value thereby calculating the probability of all the classes and classifiers the transaction as legitimate or fraud transaction. The NB model is fast and simple, therefore even with little training it can work well with large dataset making fast decision and flagging the suspicious transaction in real time. RNN is a recurrent neural network which works on sequence data where data order matters. In UPI Fraud Detection the RNN looks at the sequence of transaction and detect any unusual pattern which may mark as fraud. One of its advantage is its "memory" which remembers the finding of previous steps. The RNN then make prediction based on previous finding and current information

The Ensemble RNNBoost model combines two different machine learning approaches to leverage their complementary strengths for better prediction performance: During ensemble steps, the model produces accurate probabilities of each sample comparing with its independent model. This model then uses weighted average which then combine the prediction to generate the final conclusion. The output is assessed using measures like accuracy, F1 score, Recall and Precision. Since the model often capitalizes multiple perspective on the data, it is used to improve the detection performance. This algorithm works well with even complex or imbalance datasheet.

IV. RESULT

This portion consists of the performance measures that is used to evaluate the accuracy, precision, recall and f1 score respectively using Eq. 1, Eq. 2, Eq. 3 and Eq. 4. This performance metric accuracy is used to calculate how the model makes a correct prediction.

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)} \dots \dots \dots (\text{Eq.1})$$

Wherein TP = True Positive (Correctly predicted positive cases), TN = True Negative (Correctly predicted negative cases), FP= False Positive (incorrectly predicted positive cases), FN = False Negative (incorrectly predicted negative cases). The precision metrics measures the truly positive from all the predictive positive of the model.

$$\text{Precision} = \frac{(TP)}{(TP+FP)} \dots \dots \dots (\text{Eq.2})$$

Wherein TP = True Positive (Correctly predicted positive cases), FP = False Positive (incorrectly predicted positive cases)

The recall metrics evaluates the number of correct positives being identified by the model.

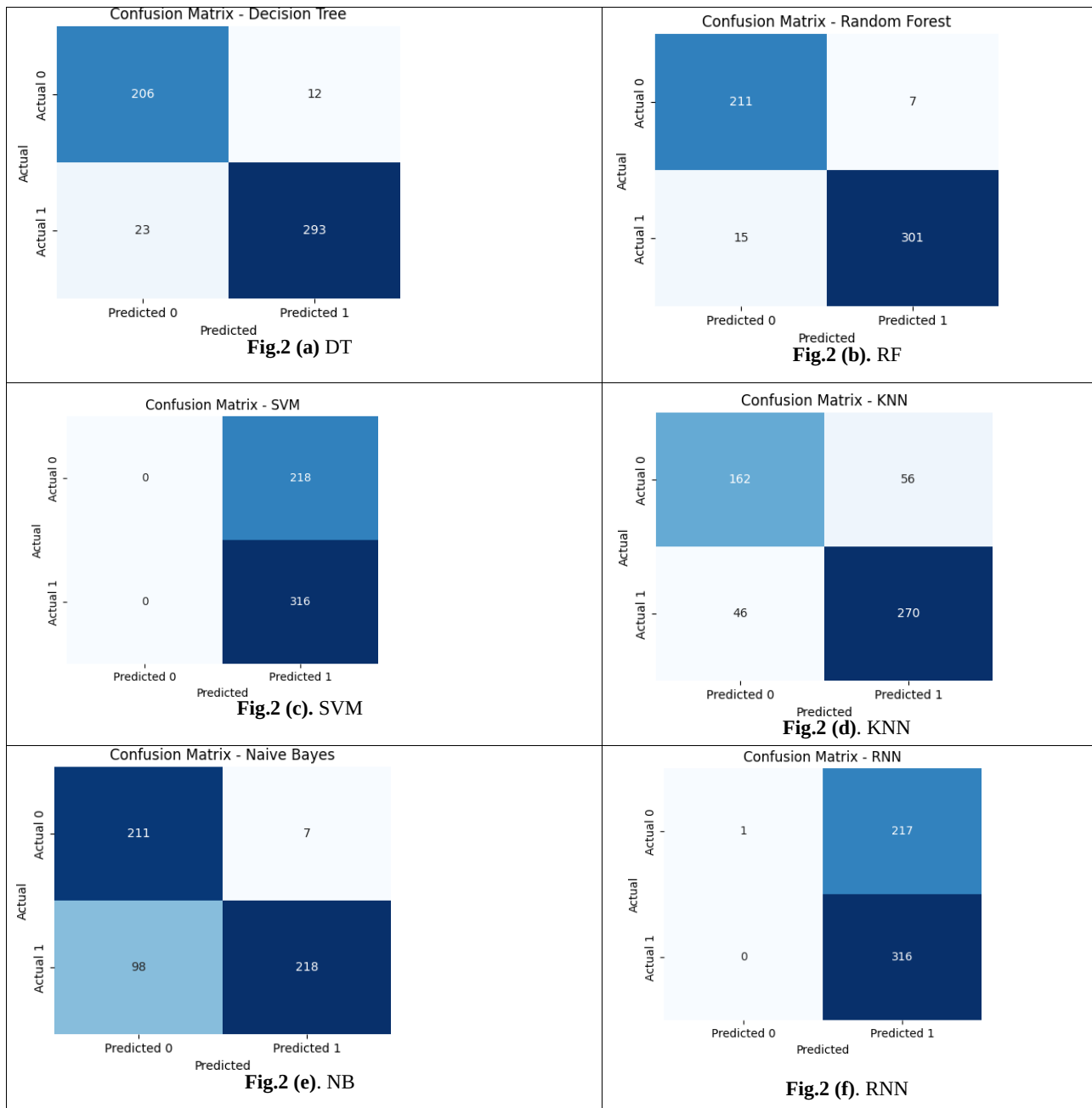
$$\text{Recall} = \frac{(TP)}{(TP+FN)} \dots \dots \dots (\text{Eq.3})$$

Wherein TP = True Positive (Correctly predicted positive cases), FN = False Negative (Positive cases the model missed)

$$F1 \text{ Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \dots \dots \text{(Eq.4)}$$

The recall as well as precision values are balances by the F1 score measure

To show how the model performs, confusion matrix is used often which is in the form of table, Since the dataset in our case is an imbalanced dataset, from fig.2c we can see that SVM model has the poorest performance as compared to the rest of the model. While models like Ensemble RNNBoost, RF and DT has performed well. Ensemble RNNBoost has outperformed from the rest of the model by achieving higher TN of 211 and lower FN of 12 and higher TP of 304 and lower FP of 7 as shown in fig.2h, resulting in a higher accuracy with 96.44 and F1 with 96.97 as mentioned in Table.1, improving this algorithm performance by classifying correctly the fraudulent transaction ID and minimizing the legitimate transaction which are flagged as fraudulent transaction.



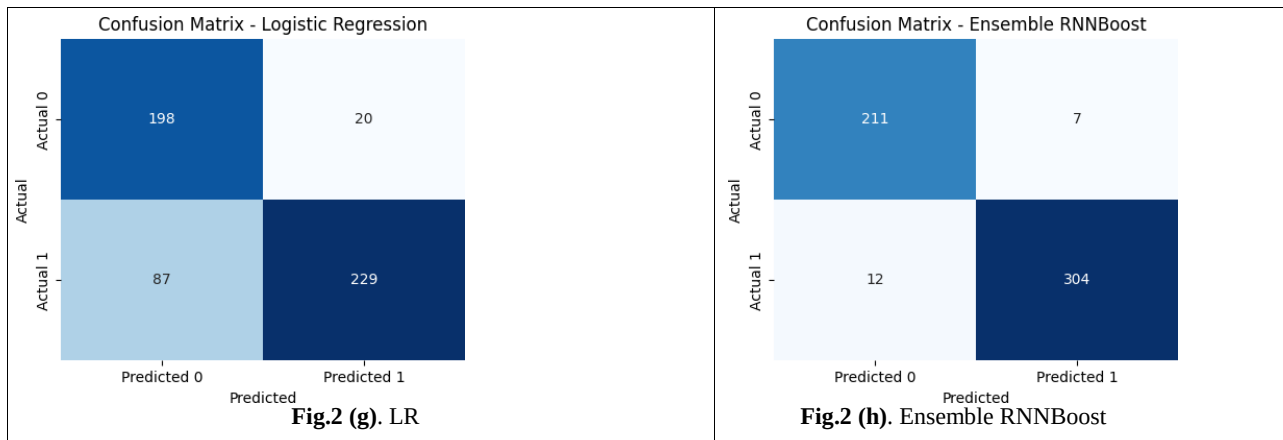


Fig.2. Confusion Matrix of various models

In Fig. 3 Ensemble RNNBoost has achieved highest accuracy of 96.44 which indicate that the model has correctly identified the legitimate transaction as well as the fraudulent transaction correctly. But, since the accuracy value alone can be misleading sometimes therefore, in this research we have calculated the F1-Score, Precision and recall to give us accurate performance which is shown in the Table.1

	Algorithm Name	Accuracy Score (%)
7	Ensemble RNNBoost	96.441948
5	Random Forest	95.880150
4	Decision Tree	95.000000
1	K-Nearest Neighbors	83.000000
2	Support Vector Machine	83.000000
3	Naive Bayes	83.000000
6	RNN Model	81.250000
0	Logistic Regression	80.250000

Fig. 3. Performance summary of models along with Accuracy Score.

Table.1. Comparison of performance metrics

Metrics	Ensemble RNNBoost	RF	DT	KNN	SVM	RNN	NB	LR
F1 score (%)	96.97	96.47	94.36	84.11	74.35	74.44	80.59	81.06
Precision (%)	97.75	97.72	96.07	82.82	59.18	59.29	96.89	91.97
Recall (%)	96.20	95.25	92.72	85.44	100	100	68.99	72.47
False Positive Rate (%)	0.32	3.21	5.5	2.57	10	9.95	0.32	0.92
False Negative Rate (%)	0.38	4.75	7.28	1.46	0	0	3.1	2.75
Accuracy (%)	96.44	95.88	95	83	83	81.25	83	80.25

In UPI Fraud Detection, we focused on the F1-score because it highly ensures that the model performs well in detecting actual fraudulent transaction and avoiding the false positive. Therefore, as mentioned above in Table.1 and Fig.4, although other models like SVM and RNN shows slightly higher Recall, the Ensemble RNNBoost achieve the highest F1-score, Precision and Accuracy among all the other models which indicate that it balanced the performance making it the most robust model for Fraud Detection. Also the ensemble RNNBoost results in very low FPR and FNR, which shows that the model is most efficient overall.

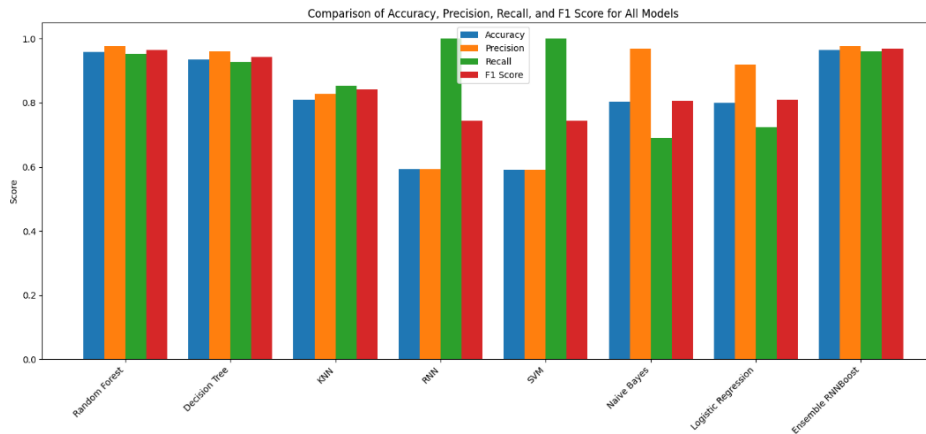


Fig.4. Model accuracy comparison

V. CONCLUSION

This research highlights the successful implementation of ensemble RNNBoost model for detecting fraudulent activities in UPI transactions. By leveraging transaction patterns and user behavior data, the system effectively differentiates between genuine and suspicious transactions. The ensemble RNNBoost got the highest accuracy which is the desired output for this research. This solution offers a valuable step toward enhancing the safety and trustworthiness of digital payment systems. Future enhancements can be done using adaptive learning techniques and advanced AI to address shifting fraud behaviors more efficiently and ensure model interpretability. Also technologies like block chain can be incorporated to the system to strengthen data privacy, security and system reliability.

REFERENCES

- [1] Chen Y the Application of Machine learning in Financial Anomaly Detection: Expert systems with applications, 154, Article113443(2020) <https://www.scirp.org/journal/paperinformation?paperid=137740>
- [2] Abdulalem Ali, Shukar Abd Razak, Siti Hajar Othman, Taiseer Abdalla Elfadil Eisa, Arafat Al-Dhaqm, Maged Nasser, Tusneem Elhassan, Hashim Elshafie and Abdu Saif . Financial Fraud detection based on Machine Learning: A Systematic Literature Review (2022 Sept) <https://www.mdpi.com/2076-3417/12/19/9637>
- [3] Palak Gupta, Anmol Varshney, Mohammad Rafeek Khan, Rafeeq Ahmed, Mohammad Shuaib and Shadab Alam. Unbalanced Credit Card Fraud detection Data: A Machine learning-oriented comparative study of balancing techniques. (2023) <https://www.sciencedirect.com/science/article/pii/S1877050923002314>
- [4] Vaishali Gupta, Sahil Sharma, Suhani NImkar, Suhani Pathak. UPI Based Financial Fraud Detection Using Deep Learning Approach (2024) <https://ieeexplore.ieee.org/document/10743663>
- [5] Vitthal B Kamble Enhancing UPI Fraud Detection (2025 Jan - March) <https://www.ijmsm.org/volume2-issue1/IJMSM-V2I1P108.pdf>
- [6] Mr.C.Lakshminath Reddy, Bindela.Akhila, Putluru Bhavana, Chekka Anitha and Bellary Jayasimha Raju. UPI Fraud Detecting Using Machine Learning (2025 March-April) <https://www.ijfmr.com/papers/2025/2/42152.pdf>
- [7] Ashish Das, Roshni Chaubey, Sristita Paul, Rahul kumar Singh, Siddhartha Majumder. Online Payment Fraud Detection Using Machine Learning (2025 April) <https://ijnrd.org/papers/IJNRD2504501.pdf>
- [8] Mrs K Amulya, Muliki Bhumika, Teki Srinivasa Murari, Potti Sriramulu Chalavadi Mallikarjuna, Saketh komaravalli and Vudatha Venu Gopal UPI Fraud Detection Using Machine Learning (2025 April) <https://ijsdr.org/papers/IJSDR2504164.pdf>
- [9] Dr C Nadhamuni Reddy, Mr E D Pavan Kumar, S Karishma, Y Prashanthi, M Varshini and D Shafi Qur Rehaman. UPI fraud detection using Machine Learning (2025 April) <https://www.jetir.org/view?paper=JETIR2504567>
- [10] Mr Khadir Kumar.N, Ramesh N, Muralidharan P, Sarathi M, Yuvaraj. UPI fraud detection Using Particle Swarm Optimization Algorithm Based On the Machine Learning (2025 May) <https://ijirt.org/article?manuscript=178536>