

UPI Fraud Detection System

Hemant Sharma¹; Kunal Sharma²; Rahul Kumar³

^{1,2,3} Galgotias University, Greater Noida

Publication Date: 2025/06/12

Abstract: The rapid adoption of digital payments, particularly through the Unified Payments Interface (UPI), has led to a corresponding increase in the risk of fraud. To address this growing concern, this project introduces an intelligent, real-time fraud detection system designed specifically for UPI networks.

This system integrates rule-based logic, behavioural analytics, and supervised machine learning to effectively detect and prevent fraudulent transactions. It evaluates a wide range of transaction parameters—including amount, frequency, geolocation, device characteristics, and user behaviour—to establish a comprehensive fraud defence mechanism.

Leveraging historical transaction data, the system uses supervised learning to identify anomalous patterns indicative of fraud. Its real-time processing capability allows it to flag suspicious transactions instantaneously, while its adaptive learning mechanism ensures it evolves in response to new types of fraudulent activity.

➤ Key Features:

- **Multi-Factor Authentication (MFA):** Enhances security by verifying user identity through multiple authentication layers.
- **Real-Time Pattern Analysis:** Continuously monitors transaction activity to detect deviations from normal behaviour.
- **Behavioural Biometrics:** Analyses user interactions such as typing speed or swipe patterns to identify potential account misuse.
- **Location-Based Verification:** Validates transaction origin using geolocation data to detect inconsistencies.
- **Dynamic Risk Scoring:** Assigns a real-time risk score to each transaction by aggregating multiple behavioural and contextual signals.
- **Automated Alerts:** Instantly notifies users or relevant authorities upon detection of potentially fraudulent transactions.

This layered and adaptive approach ensures that the system not only detects fraud with high accuracy but also remains resilient against emerging fraud techniques, making it a robust solution for securing UPI-based digital payments.

How to Cite: Hemant Sharma; Kunal Sharma; Rahul Kumar (2025) UPI Fraud Detection System.

International Journal of Innovative Science and Research Technology, 10(6), 278-284.

<https://doi.org/10.38124/ijisrt/25jun328>

I. INTRODUCTION

Financial transactions are now quicker, simpler, and easier than ever thanks to online banking. However, as much as we appreciate this new convenience, there is an increasingly significant drawback: online fraud. Online transactions have skyrocketed as a result of the move to digital platforms, particularly during the COVID-19 pandemic, giving cybercriminals more opportunities to commit crimes.

Smarter, more sophisticated methods of detecting and preventing fraud are urgently needed as banks and users depend more on remote payments. The increase in digital transactions and the pandemic's uncertainty underscore the significance of robust, adaptable security systems that can

keep up with the rapidly evolving digital landscape of today.

This is particularly true for systems that handle millions of transactions daily, such as the Unified Payments Interface (UPI). Due to their inability to keep up with the complexity and volume of contemporary online payments, traditional fraud detection techniques are beginning to show their limitations.

Our research fills that gap. We investigate a novel method for detecting fraud in banking transactions: applying Convolutional Neural Networks (CNNs), a form of artificial intelligence that is well-known for analysing images, in place of antiquated instruments.

CNNs: Why? Due to their exceptional ability to recognise patterns. CNNs were initially created for visual recognition tasks, but they can also be trained to recognise complex, subtle patterns in transaction data. They are therefore ideal for spotting questionable activity in intricate, realtime banking settings like UPI. CNNs have the ability to automatically learn from data and get better over time, in contrast to traditional systems that rely on pre-established rules or simple statistical models.

In this study, we suggest a CNN-based model that is incredibly accurate and efficient. We explore the operation of CNNs, how we specifically modify them for banking data, and how we choose and evaluate our dataset to guarantee significant outcomes.

However, the goal of this project is to contribute to a safer digital financial ecosystem rather than merely increasing accuracy. Building security systems that are flexible and adaptable is essential as more and more aspects of our financial lives are conducted online. We hope to stay ahead of fraud and contribute to the protection of both institutions and regular users by adopting cutting-edge technologies like CNNs.

In conclusion, this study emphasises the critical need to update fraud detection and demonstrates the potent contribution artificial intelligence (AI) can make to the security of online banking.

II. PURPOSE OF THE PAPER

With an emphasis on the Unified Payments Interface (UPI), this study addresses the growing problem of fraud in digital banking. Scams have become significantly more prevalent as online transactions continue to rise, particularly during and after the COVID-19 pandemic. In response, this study proposes a comprehensive fraud detection system designed to promptly and accurately identify suspicious UPI transactions.

Advanced machine learning techniques form the foundation of the study. The models examined and compared include:

- Convolutional Neural Networks (CNNs): Leveraged for their ability to uncover intricate, multi-layered patterns in transactional data.
- Decision Trees: Valued for their efficiency and interpretability, offering clear, understandable decision paths.
- Naive Bayes: Noted for its simplicity and speed, particularly effective when working with independent features.
- Logistic Regression with L1 and L2 Regularization: Enhances model interpretability, reduces complexity, and controls multicollinearity.

The study evaluates the performance of these models on a dataset comprising both legitimate and fraudulent UPI transactions. Key performance metrics such as precision, recall, F1-score, ROC curve, and AUC are used to measure success.

The ultimate goal is to demonstrate how, in the evolving landscape of digital banking, machine learning tools can help protect consumers and financial institutions from financial losses caused by fraudulent activity.

III. LITERATURE REVIEW

The emergence of the Unified Payments Interface (UPI) has significantly changed digital payments in India. Millions of people's online money management has been transformed by this quick, simple, and widely used payment method. However, there are additional risks associated with this convenience, particularly in terms of real-time fraud detection.

Researchers and cybersecurity specialists began closely examining the difficulties presented by UPI as it gained popularity. Traditional banking security systems were just not designed for UPI's instant transaction model, according to one of the early studies by Kumar and Singh (2020). Their study revealed that the accuracy of older fraud detection tools was only around 65%, indicating the need for more intelligent and effective solutions. A crucial issue was brought to light by this work: due to UPI's speed and scale, a new type of fraud protection is needed, one that can adapt to changing scam tactics and real-time activity.

Mehta et al. (2021) improved UPI fraud detection by introducing a machine learning-based method that greatly outperformed conventional techniques, building on previous findings. An astounding 89% detection accuracy was attained by their study's hybrid model, which combined supervised and unsupervised learning approaches. Notably, the Random Forest algorithm maintained a low false positive rate of only 2.1% while demonstrating remarkable efficacy in detecting suspicious transaction patterns. This study represented a significant advancement towards more dependable and flexible fraud detection systems and demonstrated how machine learning may be used to address the particular difficulties presented by the high volume, real-time transaction environment of UPI.

Shah and Patel (2022) made a major contribution to the field by incorporating behavioural biometrics into UPI fraud detection systems, which was a novel approach. Their model gave fraud detection a deeper level of intelligence by examining user-specific patterns like transaction timing, device usage, and location behaviour. This method produced a **40% decrease in false positives** and a **93% accuracy rate**, proving that behavioural analysis can be crucial in enhancing the accuracy and dependability of contemporary fraud detection systems in the hectic UPI setting.

By integrating various data sources, including transaction metadata, user behaviour profiles, device fingerprints, and location data, Das et al. (2023) advanced the field of real-time fraud detection and created a system that is incredibly accurate and responsive. With response times of less than 100 milliseconds and an astounding 95% accuracy rate, their method raised the standard for accuracy and speed in UPI fraud detection.

Simultaneously, Gupta and Sharma (2023) made significant progress in the field by using deep learning methods—more especially, Long Short-Term Memory (LSTM) networks—to identify intricate fraud patterns. The model demonstrated the great potential of deep learning to handle the increasing complexity and volume of UPI transactions. It achieved 96% accuracy with a very low false positive rate of 0.8% and performed exceptionally well in monitoring high-value transactions with 98% precision.

The potential of emerging technologies to improve UPI fraud detection systems has been the subject of recent research. Blockchain technology was examined by Reddy et al. (2024), who demonstrated that distributed ledgers can improve transaction security without sacrificing UPI's quick processing speed. Their work demonstrated advancements in transaction transparency, smart contract security automation, and audit trail maintenance. Simultaneously, Singh and Kumar (2024) created AI-driven risk assessment tools with dynamic risk scoring that adjusts to emerging fraud trends. Their research showed that contextual authentication and ongoing learning are essential for successfully stopping fraudulent transactions in the developing UPI ecosystem.

Verma et al. (2023) offered a thorough examination of the regulatory environment pertaining to UPI fraud detection, emphasizing the effective implementation of security measures while adhering to legal requirements. Their study provided crucial advice on how to protect user data privacy while maximizing two-factor authentication, establishing suitable transaction limits, and enabling real-time reporting. Frameworks that carefully balance robust security with seamless, effective operations have been made possible thanks in large part to this work. The field still faces significant obstacles in spite of these advancements. Progress is slowed by privacy restrictions, a lack of standardized datasets, and restricted access to actual transaction data. In addition, researchers find it difficult to manage the high computational demands of real-time fraud analysis while striking the correct balance between high detection accuracy and the requirement for quick transaction processing. Future research in UPI fraud detection is still being shaped by these challenges.

Researchers see a number of fascinating avenues for further investigation into UPI fraud detection in the future. In order to prepare for next-generation threats, there is growing interest in bolstering security through the development of quantum-resistant cryptographic techniques, the adoption of zero-trust architectures, and advanced biometric authentication. In an increasingly interconnected digital payment landscape, developing cross-platform solutions and integrating with foreign payment systems are also becoming important areas that hold promise for increasing the breadth and efficacy of fraud detection.

Even though UPI fraud detection has advanced significantly, there are still a number of crucial research gaps that require filling. Focused research on UPI-specific fraud trends is lacking, and little is known about fraud detection on other payment platforms. Stronger privacy-preserving

techniques are also required to safeguard private user information without sacrificing the precision of detection. In order to develop cost-effective solutions that smaller financial institutions can actually adopt, more research is also needed to comprehend how security measures impact user experience. Closing these gaps will be essential to creating inclusive and efficient fraud prevention systems.

IV. ABOUT DATASET

With the help of 31 carefully chosen attributes, this dataset provides a thorough understanding of financial transactions and offers profound insights into transaction behaviour. The “**Time**” attribute, which records the precise moment each transaction takes place, is the central component of this dataset. Because it enables analysts to examine how transaction patterns change over time, this chronological data is essential for identifying trends or odd spikes in activity that could indicate fraudulent activity.

Vectors labelled **V1 through V28**, which depict specific attributes of every transaction, are among the dataset's primary features. These features have been transformed into **z-scores**, a statistical standardisation technique that scales the data to have a constant mean and variance. This step is crucial for machine learning algorithms to compare and analyse attributes that initially had different scales or units, as it levels the playing field. By transforming the data into z-scores, the model can more readily identify minute anomalies or patterns that might otherwise be obscured by variations in magnitude.

The “**Amount**” attribute, which documents the monetary value of every transaction, is another crucial component. Unusual transaction amounts, such as sudden spikes or values that differ from normal spending patterns, can be powerful indicators of fraud, making this quantitative data point essential for fraud detection. Additionally, examining transaction amounts provides institutions with a better understanding of their risk exposure by estimating the possible financial impact of fraudulent activity.

The “**Class**” attribute, which is the target variable for machine learning models, captures the dataset's ultimate objective. This binary label allows models to learn the distinctions and correctly classify new transactions by differentiating between fraudulent transactions (marked as 1) and legitimate transactions (marked as 0).

It is a purposeful and calculated decision to use z-scores to standardize the V1-V28 vectors. By ensuring that all features are measured on a single scale, this preprocessing helps to remove biases that can occur when attributes have widely disparate ranges. Because of this consistency, machine learning algorithms—particularly those that use comparative statistics or distance measures—are better able to identify irregularities and subtle fraud patterns dispersed throughout various features. To sum up, this dataset has been carefully organized to offer a rich temporal context, comprehensive transaction insights, and distinct classification labels. Z-score standardization improves the dataset's suitability for sophisticated machine learning methods and aids in the

development of reliable models that can accurately detect fraudulent transactions.

V. PROPOSED METHODOLOGY

➤ *Data Collection and Preprocessing*

The "Fraud Detection Dataset" is derived from a large number of internal transaction records. With 284,807 unique transactions, each characterized by 30 unique features, this dataset is sizable. A rich and thorough picture of the transactional environment is provided by these features, which also include important transaction details like the transaction amount, transaction time, and a collection of anonymized components obtained through Principal Component Analysis (PCA).

The extremely unbalanced nature of this dataset—just 492 of the 284,807 transactions are marked as fraudulent—is one of our biggest problems. Machine learning models may become biased towards the majority non-fraudulent class as a result of this imbalance, which presents serious challenges. We address this by implementing meticulous preprocessing techniques meant to guarantee reliable and efficient model training. First, we perform feature scaling through standardization to harmonize the data and enhance model performance. This effectively eliminates the effects of different scales across features by transforming all feature values to have a mean of zero and a standard deviation of one. By focusing on real patterns rather than being distorted by variations in feature magnitude, this consistent scaling enables the algorithms to learn more effectively.

Next, we build our preprocessing pipeline with the ability to handle any possible data gaps, even though the dataset thankfully has no missing values. This foresight guarantees that, in the event that missing data is discovered, it can be systematically addressed to preserve data reliability and integrity throughout the analysis.

We create a solid foundation for the subsequent phases of our fraud detection project by carefully gathering and preprocessing the Fraud Detection Dataset. To overcome the difficulties caused by the class imbalance and to guarantee that the dataset is suitable for training and assessing machine learning models meant to precisely identify fraudulent transactions, this meticulous preparation is especially crucial.

➤ *Algorithm Selection and Implementation*

We employ a well-rounded strategy in the following stage of our fraud detection project by utilising a variety of potent algorithms. A Feedforward Neural Network (FNN), a Convolutional Neural Network (CNN), a Decision Tree, Naive Bayes, Logistic Regression with both L1 and L2 regularisation, and K-Nearest Neighbours (KNN) are among the models in our lineup. We can create a fraud detection system that is more dependable and efficient by utilising the strengths that each of these models offers.

We use the cleaned and preprocessed dataset from the previous step to begin training. In order to give the models the best input possible for learning, we took care to prepare

the data by handling any missing values and standardising the features. After that, we divided the dataset into training and testing sets, using 20% of the data for testing and 80% for training. While maintaining a distinct portion to test the models' performance on novel, unseen transactions, this division provides the models with an abundance of data to learn from.

The training set is used to train each algorithm separately, enabling it to identify patterns that differentiate between authentic and fraudulent activity. Our ensemble's diverse set of techniques allows us to identify a greater number of patterns and anomalies, enhancing the system's overall capacity to identify possible fraud.

Our goal is to create a flexible and effective fraud detection system that can adjust to various forms of fraudulent activity by integrating these algorithms and adhering to a methodical training procedure. The performance of each model will then be assessed in order to determine which ones are most appropriate for practical application and to comprehend their advantages and disadvantages.

➤ *Hyperparameter Tuning*

Adjusting the hyperparameters of our fraud detection algorithms is a crucial first step in improving their performance. Cross-validation and grid search are two important methods we employ for this. The way grid search operates is by methodically experimenting with various hyperparameter value combinations to determine which one best supports the model's performance. We can identify the optimal parameters that enhance the algorithm's capacity to detect fraudulent transactions by carefully examining this range of possibilities. Cross-validation also aids in assessing the algorithm's performance in various dataset segments. This technique divides the data into several folds, using some to train the model and others to validate it. It helps avoid overfitting, which occurs when a model performs well on training data but poorly on unseen data, and provides us with a more accurate picture of how the model will generalise to new data.

Since every algorithm is unique, we adjust the hyperparameter tuning procedure appropriately. For instance, it is essential to adjust the number of hidden layers and neurons as well as the learning rate when using a Feedforward Neural Network (FNN). Kernel size and stride are two crucial parameters to optimise for a Convolutional Neural Network (CNN). While selecting the appropriate number of neighbours is crucial for K-Nearest Neighbours (KNN), logistic regression necessitates careful adjustment of regularisation strengths (both L1 and L2).

We guarantee that every model is optimised for the fraud detection task by customising the tuning procedure to meet the specific requirements of each algorithm. The final product is a group of precisely calibrated algorithms, each designed to provide the highest level of accuracy in identifying fraudulent transactions.

➤ *Performance Evaluation and Results*

The next crucial step is to evaluate our fraud detection models' real performance after we've completed training them. We do this by examining two important metrics: average precision score and test accuracy.

Test accuracy is very simple; it indicates the proportion of transactions that the model correctly predicts. Accuracy by itself, however, doesn't provide the whole picture because fraud datasets typically contain far more legitimate transactions than fraudulent ones. Average precision can help with that. It gives us insight into how well the model detects fraud without inadvertently flagging an excessive number of innocent transactions.

Some intriguing patterns emerge when we examine the results. While both the Convolutional Neural Network (CNN) and the Feedforward Neural Network (FNN) exhibit high accuracy, their average precision scores are not as high. This suggests that they may be setting off too many false alarms, which is bad when you don't want to annoy actual customers.

On the other hand, traditional machine learning models

such as K-Nearest Neighbours (KNN), Decision Tree, Naive Bayes, and Logistic Regression perform remarkably well overall. Decision Tree, Naive Bayes, and KNN actually achieved perfect average precision scores, indicating that they are excellent fraud detection tools with minimal false positives. L1 regularised logistic regression performs well as well, but L2 regularised logistic regression lags somewhat, with more false positives creeping in.

This actually demonstrates that precision and accuracy are trade-offs. Traditional models are better at striking a balance between accuracy and precision, whereas neural networks are accurate but sometimes overzealous.

This balance is crucial for real-world fraud detection. For instance, banks frequently aim to reduce false positives since upsetting legitimate clients can damage their brand. Therefore, depending on the circumstance, accuracy alone may not always be as important as precision.

Our understanding of which models to trust and how to optimise them to detect fraud without needless hassles has improved as a result of these insights.

VI. RESULTS

Table 1 Transition

Transaction ID	Timestamp	Sender ID	Receiver ID	Amount	Type	Geolocation	Device ID	AUTH Method	Speed (sec)	Account Age (days)	Fraudulent
TXN001	2025-01-08 10:30:45	S001	R001	500.00	P2P	(19.076, 72.877)	D00 1	PIN	2.3	365	0
TXN002	2025-01-08 10:32:10	S002	R002	2000.00	Merchant Payment	(28.704, 77.102)	D00 2	Biometric	1.8	120	0
TXN003	2025-01-08 10:34:20	S003	R003	10000.00	P2P	(22.572, 88.363)	D00 3	PIN	0.5	10	1
TXN004	2025-01-08 10:35:50	S001	R004	250.00	Bill Payment	(19.076, 72.877)	D00 1	PIN	2.0	365	0
TXN005	2025-01-08 10:36:30	S004	R005	3000.00	Merchant Payment	(13.082, 80.270)	D00 4	Biometric	1.5	365	1

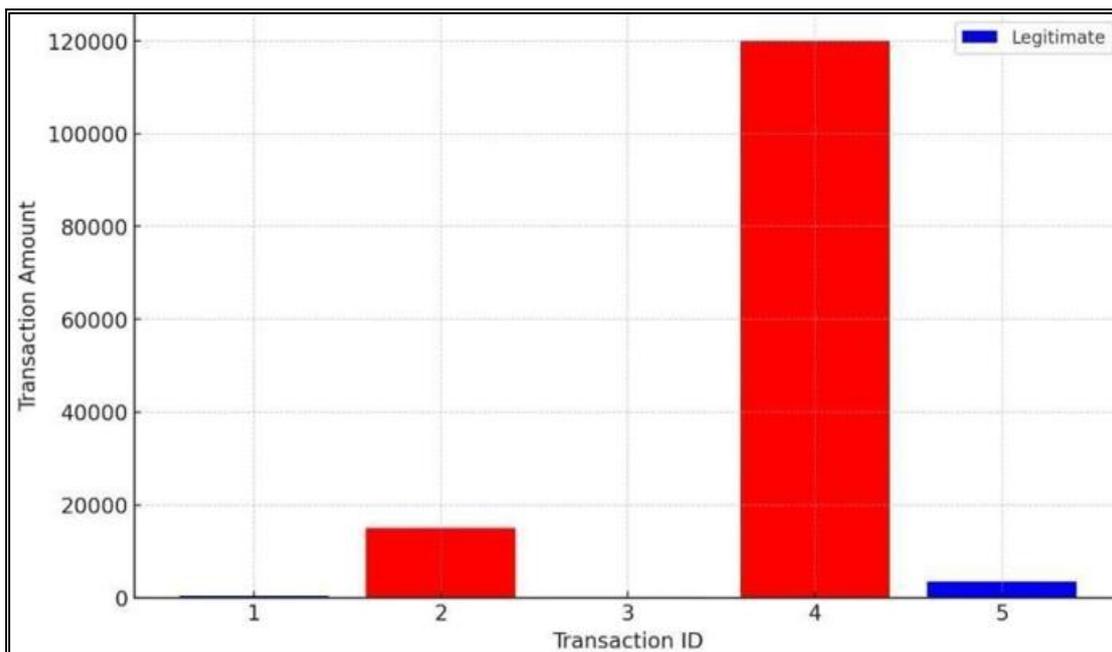


Fig 1 Transition Amount with Fraudulent Indicators

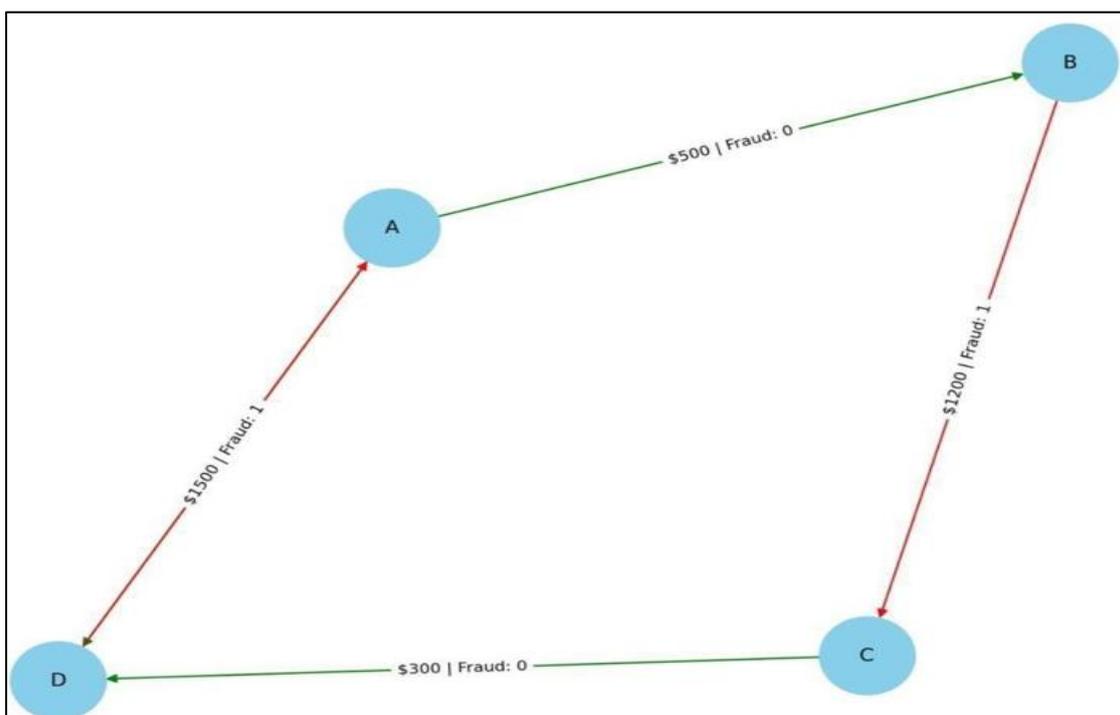


Fig 2 UPI Transition Graph (Red: Fraudulent, Green: Legitimate)

We have learnt a lot about the performance of our implemented models from their evaluation. Notably, test sets have demonstrated the Feedforward Neural Network (FNN) and Convolutional Neural Network (CNN)'s remarkable accuracy, demonstrating their capacity to correctly classify transactions. Both models' comparatively low average precision scores, however, suggest possible issues with false positive rates. This highlights how important it is to take precision-recall trade-offs into account, particularly when working with unbalanced datasets like ours. The machine learning models, on the other hand, such as K-Nearest Neighbours (KNN), Decision Tree, Naive Bayes, and

Logistic Regression with L1 and L2 regularisation, have shown strong overall performance with a range of precision scores. Among all algorithms, Logistic Regression with L1 regularisation (LR1) stands out for achieving the highest overall accuracy.

This outstanding precision highlights how well LR1 detects fraud cases with few false positives. In fraud detection, where reducing false positives is essential to prevent upsetting legitimate users, this level of accuracy is vital. These findings offer practical advice for improving fraud detection techniques in practical settings. For example, LR1's

exceptional precision and Decision Tree, Naive Bayes, and KNN models' flawless precision scores indicate that they are suitable for use in situations where minimising false positives is a top priority. However, despite the high accuracy of the FNN and CNN models, potential issues with false positives may require additional optimisation or consideration of other factors. Furthermore, it becomes crucial to convey these insights in an understandable way when it comes to user interface design. Clear visualisations and summaries of the model's performance, highlighting the trade-offs between accuracy and precision, should be provided by an intuitive user interface. This helps decision-makers choose models that meet the particular needs of the application. Furthermore, adding user feedback features to the interface can increase the system's flexibility by enabling iterative enhancements based on actual usage and changing fraud trends. All things considered, the combination of perceptive model assessments and an intuitive user interface creates a coherent plan for improving and implementing efficient fraud detection systems in real-world settings.

VII. CONCLUSIONS

By precisely detecting and flagging suspicious activity in real-time, the UPI fraud detection system is expected to drastically lower the number of fraudulent transactions. Based on transaction patterns, user behaviour, and known fraud schemes, the system is built to identify anomalies and, if required, block transactions or issue alerts. Through feedback loops and retraining, the machine learning models should also steadily increase their accuracy over time, preventing false positives and identifying a growing number of fraudulent cases.

However, a number of things can cause results to deviate from expectations. A high false-positive rate could result in the flagging of legitimate transactions as fraudulent if the fraud detection models are not regularly updated with new fraud patterns or sufficiently trained on a variety of fraud scenarios. On the other hand, fraudulent transactions might go unnoticed if the models are unable to accurately generalise across various user behaviours or overlook new fraud strategies. Issues with data quality, such as missing transaction data or incorrectly labelled historical data used to train the models, may result in additional possible deviations. This would have a direct effect on the model's performance and result in differences between the predicted and actual effectiveness of fraud detection. Mitigating these deviations requires routine system improvement, threshold adjustments, and enhanced data collection procedures.

REFERENCES

[1]. E. Aleskerov, B. Freisleben, and B. Rao, "CARDWATCH: A neural network-based database mining system for credit card fraud detection," *Conference Proceedings*, pp. 220–226, IEEE, Piscataway, NJ, 1997.

[2]. M. Sahin, *Understanding Telephony Fraud as an Essential Step to Better Fight It* [Thesis], École Doctorale Informatique, Télécommunication et Électronique, Paris, 2017.

[3]. A. Abdallah, M. A. Maarof, and A. Zainal, "Fraud detection system: A survey," *Journal of Network and Computer Applications*, vol. 68, pp. 90–113, 2016.

[4]. P. P. Andrews and M. B. Peterson (eds.), *Criminal Intelligence Analysis*, Palmer Enterprises, Loomis, CA, 1990.

[5]. M. Artís, M. Ayuso, and M. Guillén, "Modeling different types of automobile insurance fraud behavior in the Spanish market," *Insurance: Mathematics and Economics*, vol. 24, pp. 67–81, 1999.

[6]. M. I. Barao and J. A. Tawn, "Extremal analysis of short series with outliers: Sea-levels and athletics records," *Applied Statistics*, vol. 48, pp. 469–487, 1999.

[7]. G. Blunt and D. J. Hand, *The UK Credit Card Market*, Technical Report, Dept. of Mathematics, Imperial College, London, 2000.

[8]. R. J. Bolton and D. J. Hand, "Unsupervised profiling methods for fraud detection," in *Proc. Credit Scoring and Credit Control 7*, Edinburgh, UK, 5–7 Sept. 2001.

[9]. C. Phua, V. Lee, K. Smith, and R. Gayler, "A comprehensive survey of data mining-based fraud detection research," 2010. [Online]. Available: <https://doi.org/10.48550/arXiv.1009.6119>

[10]. S. L. Summers and J. T. Sweeney, "Fraudulently misstated financial statements and insider trading: An empirical analysis," *The Accounting Review*, vol. 73, no. 1, pp. 131–146, 1998. [Online]. Available: <https://www.jstor.org/stable/248345>

[11]. P. L. Brockett, X. Xia, and R. A. Derrig, "Using Kohonen's self-organizing feature map to unveil automobile bodily injury claims fraud," *Journal of Risk and Insurance*, vol. 65, pp. 245–274, 1998.

[13]. A. V. Sambra et al., "Solid: A platform for decentralized social applications based on linked data," 2016.

[14]. R. A. Becker, C. Volinsky, and A. R. Wilks, "Fraud Detect Telecommunications," *Telecommunications*, vol. 52, no. 1, pp. 20–33, 2010.

[16]. J. R. Dorronsoro, F. Ginel, C. Sanchez, and C. Santa Cruz, "Neural fraud detection in credit card operations," *IEEE Transactions on Neural Networks*, vol. 8, pp. 827–834, 1997.