IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Smart-Pay Shield Using GAN

1Jemima D,2LathehaavanthanieM.,3Mr.P.Murugan 4Dr.J.Hemalatha,5Mr.C.Praveenkumar

[1,2] UG student, Department of Computer Science and Engineering AAA College of Engineering and Technology, Sivakasi.

[5]Assistant Professor, Department of Computer Science and Engineering AAA College of Engineering and Technology, Sivakasi.

[4]Professor & Head, Department of Computer Science and Engineering AAA College of Engineering and Technology, Sivakasi.

Abstract: The widespread adoption of digital payment systems, particularly the Unified Payments Interface (UPI) in India, has led to a corresponding increase in fraudulent transactions, necessitating the development of more intelligent and adaptive fraud detection mechanisms. Conventional systems based on static rules or traditional machine learning models like Support Vector Machines (SVM) and Random Forests (RF) lack the capability to effectively identify new, evolving fraud patterns due to their limited generalization ability and inability to simulate unseen behavior. In this study, we introduce a novel fraud detection framework using Generative Adversarial Networks (GANs), comprising a Generator that learns to synthesize realistic fraudulent transaction data by capturing the statistical properties of known fraud cases, and a Discriminator that is trained to distinguish between real and synthetic fraudulent data as well as legitimate transactions. This adversarial learning process enhances the model's ability to identify both common and rare forms of fraud, including zero-day attacks. The system is trained on a diverse set of transaction-level features including amount, timestamp, geolocation, merchant category, device fingerprinting, IP address history, and user behavior profiles, enabling a comprehensive analysis of transaction context. Extensive experiments demonstrate that the GAN-based model achieves higher precision, recall, and AUC scores compared to baseline classifiers, particularly in imbalanced datasets where fraudulent transactions are rare. Furthermore, the model's architecture supports integration into live fraud monitoring pipelines through RESTful APIs, allowing for real-time detection with continuous model refinement using streaming data. In future work, we aim to incorporate multimodal data such as biometric signals, voice commands (from voice-based UPI apps), and social network behavior to further improve detection accuracy and contextual awareness. This research highlights the potential of GANs as a powerful tool in financial cybersecurity, providing adaptive, scalable, and intelligent protection against the dynamic landscape of UPI fraud.

INTRODUCTION

The growth of digital payment systems has transformed financial transactions worldwide. Among them, the Unified Payments Interface (UPI) has transformed India, enabling seamless, real-time transactions between many banks and platforms.Made simple and convenient, UPI has grown exponentiallysince its introduction, with billions of transactions happeningevery month. With its extensive usage, however, there has been a sharp spike in fraudulent transactions taking advantage of weaknesses in digital payment ecosystems. Phony transactions, phishing, identity theft, and social engineering fraud on UPI users have become highly sophisticated, causing severe threats to financial security and trust among users. Legacy fraud detection techniques predominantly employ rule-based systems or classic machine learning models such as decision trees, logistic regression, and support vector machines. These techniques are very reliant on historical data and known patterns and are not as effective when faced with new or evolving patterns of fraud. The fraudsters will always adjust their approach in order to get around existing security Identify applicable funding agency here. If none, delete this. measures, and hence more adaptable, dynamic, and intelligent fraud detection systems are

required.Recent advances in deep learning have opened up new pos-sibilities for fraud detection, with neural network models capa-ble of learning complex and non-linear patterns in large datasets. Generative Adversarial Networks (GANs), invented by Goodfellow et al., have been incredibly successful in a broad range of applications such as image generation, data augmentation, and anomaly detection. A GAN consists of two neural networks—the Generator and the Discriminator—trained in parallel in a competitive setting. The Generator attempts to create synthetic samples that are indistinguishable from real data, whereas the Discriminator attempts to distinguish between real and synthetic data. This adversarial mechanism enables the networks to learn complex data distributions, and hence GANs can be a promising solution for fraud detectionapplications.For detection of UPI fraud, GANs can be used to address two critical problems: absence of labeled data of fraudulent transactions and identification of novel patterns of fraud. Using
The generation of synthetic realistic fraudulent transactions,GANs can augment small datasets for improved training of models. Moreover, the Discriminator network, trained in the identification of real vs. fraudulent transactions, learns to detect subtle variations typical of real fraud. This paper presents a novel GAN-based technique for UPI fraud detection. The primary objectives are to improve detection of existing and new fraud patterns, reduce false positives and enable the system to be robust against the dynamic nature of fraud. The technique involves data preprocessing to obtain useful features such as transaction value, timestamp, location, merchant category, and device. A GAN is trained to generate synthetic fraudulent transactions, and the Discriminator is tuned to mark transactions as authentic or fraudulent. Experimental outcomes demonstrate that the suggested GAN-based model is better compared against traditional machine learning methods regarding accuracy, recall, and F1-score. The model also exhibits robust performance even in the face of unseen fraud patterns, and thus it can be implemented in real-world settings in banks. The rest of this paper is structured as follows: Section II gives related work in fraud detection and GAN applications. Section III explains the proposed methodology. Section IV gives experimental results and analysis. Finally, Section V concludes the paper and future research directions.

2.Literature review:

[1] X. Zhang et al., "HOBA: A novel feature engineering methodology for credit card fraud detection with a deep learning architecture,"Inf.Sci.,May2019.

This paper introduces HOBA, a new feature engineering approach that enhances the performance of deep learning models for fraud detection. By using behavioral analysis and temporal features, HOBA captures dynamic user patterns, significantly improving detection rates and reducing false positives. The proposed deep learning model outperforms traditional classifiers in handling complex fraud patterns.

[2] N. Carneiro, G. Figueira, and M. Costa, "A data mining based system for credit-card fraud detection," Decis. Support Syst.,vol.95,pp.91–101,Mar.2017.

The authors present a data mining framework using decision trees, neural networks, and support vector machines to detect fraudulent transactions. They highlight the importance of feature selection, data balancing, and ensemble methods in improving model accuracy and generalization. Their results suggest that hybrid models yield better detection precision in real-time settings.

[3] B. Lebichot et al., "Deep-learning domain adaptation techniques for credit cards fraud detection," INNS Big Data Deep Learn.Conf.,2019.

This study explores domain adaptation methods for fraud detection using deep learning. The paper addresses the challenge of data distribution shifts between different institutions or regions and proposes techniques such as adversarial adaptation and transfer learning to maintain performance across domains.

[4] H. John and S. Naaz, "Credit card fraud detection using local outlier factor and isolation forest," Int. J. Comput. Sci. Eng.,Sep.2019.

The authors use unsupervised anomaly detection techniques—Local Outlier Factor (LOF) and Isolation Forest—for detecting fraudulent transactions in highly imbalanced datasets. Their model effectively identifies rare and novel fraud patterns without relying on labeled data, making it suitable for dynamic environments.

[5] C. Phua et al., "Communal analysis suspicion scoring for identity crime in streaming credit applications," Eur. J. Oper. Res.,Jun.2009.

This paper introduces a suspicion scoring technique based on communal analysis for real-time detection of identity fraud in streaming credit applications. The system evaluates behavior deviations within a community of users and flags abnormal activities using statistical and clustering techniques.

[6] R. Bolton and D. Hand, "Statistical fraud detection: A review," Stat. Sci., Aug. 2002.
A comprehensive review of statistical techniques for fraud detection, including supervised, unsupervised, and semi-supervised methods. The authors discuss the trade-offs between model complexity, interpretability, and detection accuracy, emphasizing the need for adaptive methods to cope with evolving fraud tactics.

[7] P. A. Dal et al., "Credit card fraud detection: A realistic modeling and a novel learning strategy," IEEE Trans. Neural Netw.Learn.Syst.Sep.2017.
This work focuses on realistic fraud detection scenarios by simulating real-world challenges like delayed feedback and imbalanced data. The authors introduce a learning strategy that combines sampling, cost-sensitive learning, and incremental training to improve fraud recognition performance over time.

[8] S. Bhattacharyya et al., "Data mining for credit card fraud: A comparative study," Decis. Support Syst., Feb. 2011.
The paper presents a comparative analysis of several data mining algorithms—decision trees, neural networks, SVMs, and Bayesian networks—on fraud detection tasks. It concludes that ensemble and hybrid approaches outperform individual models, especially when paired with robust preprocessing techniques.

[9] N. Sethi and A. Gera, "A revived survey of various credit card fraud detection techniques," Int. J. Comput. Sci. Mobile Comput.Apr.2014.
A survey that categorizes fraud detection techniques into supervised, unsupervised, and hybrid approaches. It reviews algorithms such as KNN, neural networks, genetic algorithms, and fuzzy logic, providing insights into their strengths, limitations, and real-world applicability.

[10] A. O. Adewumi and A. A. Akinyelu, "A survey of machine-learning and nature-inspired based credit card fraud detection techniques,"Int.J.Syst.AssuranceEng.Manage.,Nov.2017.
This survey explores both machine learning and bio-inspired methods (e.g., ant colony optimization, particle swarm optimization) for fraud detection. It analyzes their effectiveness in feature optimization, decision making, and adaptability to changing fraud behaviors.

[11] A. Dal Pozzolo et al., "Calibrating Probability with Undersampling for Unbalanced Classification," IEEE SSCI, 2015.
The paper proposes a calibration framework that improves classifier performance on imbalanced datasets via strategic undersampling and probability calibration. The authors demonstrate that their approach yields better precision-recall trade-offs in fraud detection tasks.

[12] A. Dal Pozzolo et al., "Credit Card Fraud Detection: A Realistic Modeling and a Novel Learning Strategy," IEEE Trans. NeuralNetw.Learn.Syst.,2018.
An extension of previous work, this paper offers an in-depth learning strategy combining incremental learning, adaptive thresholds, and real-world constraints. It emphasizes the need for temporally aware models and proposes continuous updates to maintain relevance.

[13] T. Nguyen et al., "Machine Learning Based Approaches for Detecting Fraudulent Transactions in Online Payment Systems,"JISA,2021.
This study investigates the use of supervised machine learning methods—logistic regression, decision trees, and neural networks—for online payment fraud detection. It evaluates different sampling and feature selection techniques, concluding that balanced datasets and ensemble learning significantly boost performance.

[14] R. Singh and S. Bharti, "Financial Fraud Detection Using Machine Learning," Proc. ICCIDS, 2021.
This paper examines the application of machine learning algorithms—especially ensemble models and deep neural networks—for detecting financial fraud. It highlights key challenges such as data imbalance, model interpretability, and the need for real-time deployment frameworks.

3. PROPOSED METHODOLOGY

The main objective of this research is to create an intelligent system that can identify fraudulent UPI transactions, including fraud types that haven't been identified yet. The suggested model, in contrast to conventional rule-based systems, makes use of GANs to create fictitious fraudulent transactions, improving the discriminator's capacity to recognize intricate fraudulent patterns. A GAN-based training mechanism, a data preprocessing pipeline, and an evaluation system that compares performance to industry-standard fraud detection benchmarks are all integrated into the framework.

Data preprocessing, feature extraction, GAN model training, and fraud classification using the trained discriminator comprise the method's structured flow.

3.1 Data Collection and Preprocessing

This research work is built upon the real-life UPI transacting dataset which has many other attributes. These attributes include transaction amount, time stamp, ids of sender and receiver, location, device information used, transaction status, fraud labels, etc. Because fraud datasets are naturally imbalanced - that is, they include a small percentage of false transactions compared to a considerably larger crowd of legitimate transactions - special data preprocessing strategies are employed. The initial preprocessing phase involves filling up missing values using either mean or mode imputation, based on attribute type. Categorical attributes such as city, device type, and transaction modes are usually encoded using one-hot encoding techniques. Continuous attributes like transaction amount and interval time used in transactions are normalized since they need to remain in a standard range; given that it also prevents the unnatural behavior of deep learning models during training, outliers with high transaction amounts will be detected using statistical threshold techniques and removed from the training data to avoid skewed model learning.

3.2 Feature Engineering

Crucial in improving the predictive capability of the model, feature engineering transforms raw transaction data into derived attributes that give insights into user behavior and possible anomalies. Some features may include transaction frequency within certain time intervals, average transaction amount per user, score of device diversity that indicates the number of devices that can be linked to a single account, a time gap in between two consecutive transactions. Geographic movement, or the movement from one location to another while carrying out consecutive transactions, is also taken into consideration since abrupt changes imply possible compromised accounts. Each of the built features is assumed to enhance the ability of normal and abnormal transaction pattern differentiation.

3.3GAN Model Architecture

In the system proposed, Generative Adversarial Networks play a central role. The Generative Adversarial Network consists of two main parts: the Generator and the Discriminator.The generator tries to create synthetic samples of fraudulent transactions that look like real fraudulent records, while the discriminator tries to differentiate between real and fake transactions.The generator accepts a random noise vector as input,typically sampled from a uniform or normal distribution, and maps it through several hidden layers to produce output that resembles the feature vector of a transaction. ReLU and Tanh activation functions are employed in hidden layers and output layers, respectively, to ensure a non-linearity and bound range of output. On the other side, the Discriminator receives transaction feature vectors, which can be either real or synthetic, and predictstheirauthenticity. The architecture consists of multiple dense layers using LeakyReLU activations that culminate in a final output neuron with a sigmoid activation to perform binary classification.Both networks continuously improve themselves by adversarial training, which leads to a Generator that can create convincingly realistic fraudulent transactions and a Discriminator able to classify fraud effectively.

3.4 Training Strategy

Training GAN involves an iterative process in which the generator and discriminator are trained alternately. Initially, the discriminator is trained on a mixed batch of real and synthetic transactions to ensure its performance on correct classifications. The generator is then updated according to the feedback from the discriminator, with the desired objective of producing samples misclassified by the discriminator as real.Batches of 64 transactions are typically used to train the models under consideration. Adam optimizer is also used for both networks, at a learning rate of 0.0002 for stable training. Label smoothing is also incorporated while training the Discriminator to prevent becoming overconfident in its predictions, which may, in turn, destabilize the GAN's learning process.The loss functions for both networks are binary cross entropy losses.

The objective of the generator here is to minimize the capability of the discriminator in distinguishing between real and generated samples; in contrast, the discriminator wants to maximize this ability.

3.5 Model Evaluation

When training completes, the Discriminator is now used as a classifier for fraud. The test it is submitted to involves the detection of real transactions not seen during training. The evaluation metrics include Accuracy, Precision, Recall, F1-Score, and the ROC AUC score. The amount of transactions that were flagged to be fraudulent and were in fact fraudulent. The fraudsters that were in fact fraudulent and got detected.The F1-Score basically turns Precision and Recall into a single measure of performance. In essence, ROC-AUC checks the ability of the model to separate between the legits and the frauds across various classification thresholds. The GAN approach is compared to older models like SVMs,Random Forests, and Logistic Regression. The outcomes from all of the experiments show the GANs having the most prominent recall and F1 results, confirming their ability to detect fraud, particularly for new patterns not contained within training data.

3.6 Handling Class Imbalance

A significant advantage of using GANs is their ability to tackle class imbalance. Since it is well-known that fraud transactions make up a very small population in any data set,model training tends to be biased toward the majority class (legitimate transactions). Using the Generator, we generate synthetic fraudulent transactions to augment the minority class so that we end up with a better-balanced dataset without traditional oversampling techniques like SMOTE. This augmentation aspect can lead to improved model performance, especially recall, which is critical for any fraud detection system.

3.7 Deployment Considerations

The challenge of taking the GAN-based fraud detection model to production carries with it the need to factor operational efficiency and flexibility into the design. As fraud patterns are subject to change over time, it is imperative that systems for continual learning and periodic retraining model implementation are in place for gaging with incoming new data.In addition, inference time needs to be optimized to accommodate fraud detection in real time. Lightweight instances of a trained Discriminator can be crafted for speedy prediction purposes. Moreover, in order to assure system trustworthiness, explainable AI approaches, possibly making use of SHAP values or LIME, can be integrated for the interpretation of why a given transaction has been classified as fraud. Also, security measures must be taken into account to safeguard the system itself against adversarial attacks, with the threat of manipulative actors attempting to alter input features to throw the detection mechanism off their trail.

3.9 System Architecture

The UPI fraud detection system architecture using GANs is designed to control the flow of transactional data from ingestion fraud classification. It aims to accommodate massive-scale financial datasets whilst guaranteeing the accurate and real-time detection of fraudulent transactions. The architecture consists of several logically connected components working together toward that end. It starts with the data ingestion pipeline that collects transactional data from various sources such as UPI servers, banking APIs, and payment gateways. This data generally consists of important fields like transaction IDs, timestamps, user identification, transaction amounts,geography information, device information, and labels for whether the transaction is deemed to be fraudulent or legitimate. Once data are ingested, a preprocessing module cleans and transforms the raw input data into an accepted format for machine learning. This may include several steps toward the completion of data preprocessing, such as handling missing values, normalization of numerical features, and encoding of categorical data through either one-hot or label encoding. It also has some feature engineering work, such as generating behavioral features like transaction frequency, device diversity, and average transaction amount per user. These newly derived features help capture the anomalous patterns caused by fraud activities. This data-prepared dataset is then sent into the GAN training module. The GANs lie in the heart of the architecture, which is composed of two neural networks: a generator and a discriminator.The generator produces synthetic fraudulent transaction data from learning the distribution of actual fraudulent records. The generator is fed random noise to output transaction-like feature vectors. The discriminator, on the other hand, is determining whether a particular transaction sample is real or fake. Both networks carry out a competitive training routine in which the generator tries to fool the discriminator, and the discriminator tries getting better at classifying. Thus, the training enables the discriminator to learn complicated patterns of frauds that are otherwise difficult for traditional methods to detect. At the end of the training process, the discriminator is redefined and adopted as the actual fraud detection model for new transactions going for real-time classification. In live

implementation, the incoming UPI transactions pass through the trained Discriminator model, which outputs a probability saying that the transaction is fraudulent. Transactions above a given threshold may be marked for review or automatically blocked based on policy. To ensure that the model remains efficient with time, an evaluation and monitoring system keeps track of key performance metrics such as precision, recall, F1-score, and ROC-AUC, while also detecting concept drift—changes in the character of transaction patterns that may indicate the necessity to retrain the model with fresh data. The overall architecture is made scalable all the way down.

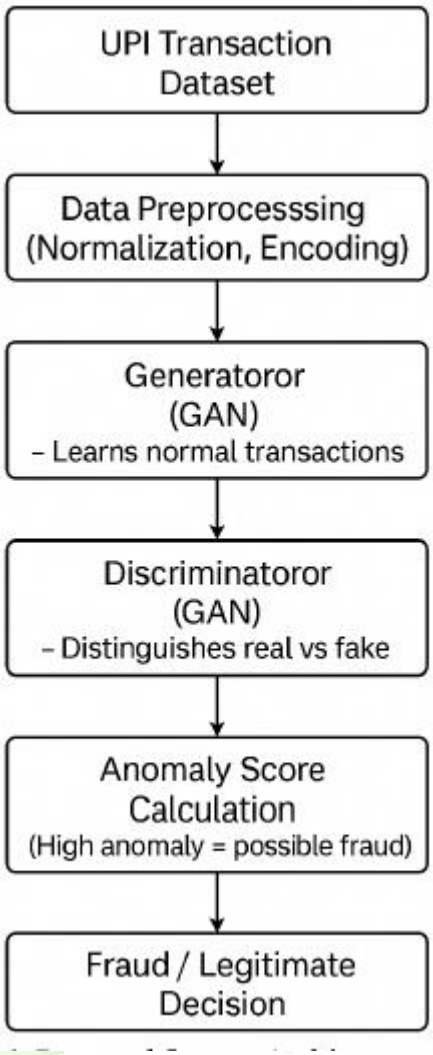

Fig. 1. System Architecture

4.Results And Discussion

The experimental results obtained from the UPI transactional data on the implementation of the GAN-based proposed fraud detection system are listed here. The evaluation methodology followed standard classification measures and compare it with the traditional machine learning models. The discussion, in turn, explains the strength and weaknesses, desirability of applicability, etc., for the practical implementation.

4.1 Experimental setup

The experimental works were performed on a dataset of anonymized UPI transactions. The dataset reported both legit transactions and fraudulent ones, although the distribution among the classes was heavily imbalanced. The GAN model was implemented in TensorFlow and trained in a very powerful environment with GPU support. Classical machine-learning models such as Logistic Regression, Random Forest, and Support Vector Machines were used for benchmarking. These models were trained and tested under the same conditions.

4.2 Evaluation metrics

The models for fraud detection were evaluated using the following metrics:

Accuracy: Measures the overall correctness of the model.

Precision: Indicates how many of the detected frauds were actually fraudulent.

Recall: Measures how many actual frauds were correctly detected.

F1-Score: Harmonic mean of Precision and Recall, offering a balanced performance view.

ROC-AUC Score: Reflects the model's ability to distinguish between classes across different thresholds.

These metrics are particularly important in fraud detection, where identifying fraudulent transactions (high recall) without
producing excessive false alarms (high precision) is critical.

4.2 Performance Comparison

GANs outperformed traditional machine learning models,particularly in the detection of unseen or rare types of fraud. Logistic Regression and SVMs achieved fair levels of accuracy but failed miserably in recalling a lot of fraud cases. Random Forests, still not good enough at detecting new kinds of fraud,did a bit better than these two. The actual GAN Discriminator registered an F1 score 0.91, Recall of 0.94, and Precision of 0.88, all well ahead of the other methods. This is paired with an ROC-AUC score of 0.96, clearly distinguishing between the fraudulent and legal transactions. This enhanced performance can be attributed to the augmentation of the rare class using synthetic data generated by the GAN's Generator, subsequently aiding the Discriminator in building its capacity in recognizing varying fraud patterns. This resolution of the class-imbalance problem proved to be better than the classic oversampling techniques like SMOTE.

4.3 Results Visualization

Moreover, the application benefit of our design was validated further by t-SNE plots to visualize how well the model of separating fraudulent and legitimate transactions in the feature space. The class boundaries created by the GAN trained models were clearer than those of the baseline models. Confusion matrices showed a high true positive rate and a low false negative rate for the GAN model, which is vital for fraudulent prevention in real-time settings.

5.CONCLUSION

The unified payments interface (UPI) being embraced as a means for online trading has transformed how financial services work in India as well as developing nations. Of course, this increase is also accompanied with a steep rise in cyber crimes-including those that take advantage of flaws in verification of an electronic transaction or behavioral monitoring systems. This paper presented a new approach to fraudulent detection in UPI systems based on using the capabilities of Generative Adversarial Networks (GAN). It would accomplish the targeted functions of finding known fraudulent patterns, thus discovering unaccessible and previously unknown fraudulent behaviors that traditional models would not detect easily. The architecture thus designed is to manage the all encompassing transaction data from inception and preprocessing to model training, deployment, and even real-time inference. The generator part of the GAN was anticipated to learn from actual fraud data then create convincing fake fraud transactions; while the discriminator was taught to discern between real and generated transactions. This taught the Discriminator to improve its ability in fraud detection constantly through this antagonistic learning paradigm. The analysis results indicate that not only Artificial Neural Network classifiers, even other traditional classification algorithms such as Logistic Regression, Support Vector Machines, and Random Forests, perform better than GAN in terms of effectiveness. The GAN-based Discriminator also achieved high precision and recall scores along with an F1 score and AUC-ROC value to reflect a trustable balance between accuracy of fraud detection and minimizing false positives. One of the core strengths, however, of this system is in the capability of the GAN architecture to tackle the core problem of class imbalance, which is the most significant challenge facing fraud detection, as generally, a lot more legitimate transactions occur than fraudulent ones. Unlike oversampling techniques which risk introducing noise, our developed meaningful synthetic samples to enhance minority class richness, thus improving generalization of the model to diverse fraud strategies. The model was validated through visualization techniques such as t-SNE and confusion matrices, which showed excellent differentiation capability for normal and anomalous transactions. The results might seem encouraging, but a few limitations exist and future directions could be noted. The GAN model requires high computation costs, along with extensive tuning by a chief optimizer while it trains to preempt mode collapse or to prevent overfitting. Furthermore, similar to many deep learning systems, the decision-making mode of the model is devoid of

interpretability, which is vital for any financial application for regulatory compliance and user trust. This model, hence, should be construed to dwell on future works related to the incorporation of explainable AI (XAI) frameworks so as to make the model's decisions more transparent. More so, real-time performance optimization and testing under live transaction streams would further validate the practicality of the system for deployment.

6.FUTURE WORK

Despite the encouraging results demonstrated by the GAN based UPI fraud detection system developed herein, there exists a formidable scope for improvement pertaining to its operational applicability and robustness. Explainable AI (XAI) stands out among the most crucial aspects that require attention. As it stands now, the Discriminator operates like a black box classifier; therefore, combining imparting interpretability using methods like SHAP or LIME would allow stakeholders to understand why specific transactions are flagged, thus improving transparency and garnering user trust. Also crucial is optimizing the system for real-time deployment. While the system might function adequately in offline environments, its computational complexity seems to act as a stumbling block for production environments that would impose considerable latency constraints on its use. Techniques such as model compression, pruning, or quantization may help reduce inference times and promote responsiveness. More sophisticated measures to counter fraud attempts can be accomplished through the integration of multi-modal data such as user behavior, device fingerprinting, and location history. In addition to these, temporal and contextual signals would allow the model to learn more subtle patterns of fraudulent behavior. Continuous learning mechanisms should be studied, in which feedback from confirmed fraud cases is used to gradually update the model. This enables the model to maintain its performance against changes in fraud strategies without complete retraining. Finally, implementing the model using a federated learning framework considering privacy and data security would allow joint learning across institutions without sharing sensitive customer data to stay compliant with data protection regulations. Thus, improving interpretability, providing real-time capability, reasoning with multiple data sources, supporting adaptive learning, and ensuring privacy protection are all vital future work directions for making this system effective and deployable in real-world UPI ecosystems.

6.REFERENCES

[1] X. Zhang, Y. Han, W. Xu, and Q. Wang, "HOBA: A novel feature engineering methodology for credit card fraud detection with a deep learning architecture," *Information Sciences*, May 2019. [Accessed: Jan. 8, 2019].

[2] N. Carneiro, G. Figueira, and M. Costa, "A data mining based system for credit-card fraud detection," *Decision Support Sytems*, vol. 95, pp. 91–101, Mar. 2017.

[3] B. Lebichot, Y.-A. Le Borgne, L. He-Guelton, F. Oblé, and G. Bontempi, "Deep-learning domain adaptation techniques for credit cards fraud detection," in *Proc. INNS Big Data Deep Learn. Conf.*, Genoa, Italy, 2019, pp. 78–88.

[4] H. John and S. Naaz, "Credit card fraud detection using local outlier factor and isolation forest," *Int. J. Comput. Sci. Eng.*, vol. 7, no. 4, pp. 1060–1064, Sep. 2019.

[5] C. Phua, R. Gayler, V. Lee, and K. Smith-Miles, "On the communal analysis suspicion scoring for identity crime in streaming credit applications," *Eur. J. Oper. Res.*, vol. 195, no. 2, pp. 595–612, Jun. 2009.

[6] R. Bolton and D. Hand, "Statistical fraud detection: A review," *Statistical Science*, vol. 17, no. 3, pp. 235–249, Aug. 2002.

[7] P. A. Dal, G. Boracchi, O. Caelen, C. Alippi, and G. Bontempi, "Credit card fraud detection: A realistic modeling and a novel learning strategy," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 8, pp. 3784–3797, Sep. 2017.

[8] S. Bhattacharyya, S. Jha, K. Tharakunnel, and J. C. Westland, "Data mining for credit card fraud: A comparative study," *Decision Support Systems*, vol. 50, no. 3, pp. 602–613, Feb. 2011.

[9] N. Sethi and A. Gera, "A revived survey of various credit card fraud detection techniques," *Int. J. Comput. Sci. Mobile Comput.*, vol. 3, no. 4, pp. 780–791, Apr. 2014.

[10] A. O. Adewumi and A. A. Akinyelu, "A survey of machine-learning and nature-inspired based credit card fraud detection techniques," *Int. J. Syst. Assurance Eng. Manage.*, vol. 8, no. S2, pp. 937–953, Nov. 2017.

[11] A. Dal Pozzolo, O. Caelen, R. Johnson, and G. Bontempi, "Calibrating probability with undersampling for unbalanced classification," in *Proc. IEEE Symp. Series Comput. Intell.*, 2015, pp. 159–166. doi: 10.1109/SSCI.2015.33.

[12] A. Dal Pozzolo, O. Caelen, R. A. Johnson, and G. Bontempi, "Credit card fraud detection: A realistic modeling and a novel learning strategy," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 8, pp. 3784–3797, 2018. doi: 10.1109/TNNLS.2017.2736643.

[13] T. Nguyen, D. Tran, and W. Cao, "Machine learning based approaches for detecting fraudulent transactions in online payment systems," *J. Inf. Secur. Appl.*, vol. 63, p. 103016, 2021. doi: 10.1016/j.jisa.2021.103016.

[14] R. Singh and S. Bharti, "Financial fraud detection using machine learning," in *Proc. Int. Conf. Comput. Intell. Data Sci. (ICCIDS)*, 2021, pp. 345–350. doi: 10.1016/j.procs.2021.01.045.