

Ai-Based Financial Transaction Analysis for Fraud Detection in Digital Commerce

Mrs.N.Zabina

assistant Professor, Mba Department

Ethiraj College For Women, No.70, Ethiraj Salai, Chennai: 600008

Zabina_n@Ethirajcollege.Edu.In

Abstract- With growing adoption of e-commerce platforms, the problem of identification of fraudulent financial transactions in a real-time manner has increased manifold. Current approaches of rules-based approaches and stand-alone machine learning models have faced difficulty adapting and dealing with novel types of fraud cases when there is heavy class imbalance and only few examples of frauds out of the total volumes. In this research, we propose the novel hybrid system of fraud detection for e-commerce using techniques of isolation forest, representation learning via autoencoder, and supervised learning using xgboost for classification. Unsupervised learning using anomaly detection is used to identify statistical anomalies as well as reconstruction-based outliers and supervised gradient boosting approach is used to improve the decision boundaries on labeled data. Evaluation experiments performed on large-scale anonymous transaction dataset show better discriminative performance of our hybrid approach as compared to the current benchmark models with respect to both the F1-score and AUC score. An explainable approach using SHAP values is provided as part of our framework, complying to regulations on explainability.

Key Word: Fraud Detection, Anomaly Detection, Isolation Forest, Autoencoder, XGboost, Digital Commerce, Explainable AI, Class Imbalance

I. INTRODUCTION

This development has greatly affected how financial transactions take place worldwide because now transactions can occur without any geographic constraints at all times. It is a process that generates a huge volume of financial data through online shopping sites performing millions of transactions per day. Nevertheless, it should be noted that the mentioned characteristics have certain flaws as well; they provide grounds for financial fraud by cyber criminals. Such types of crimes like payment fraud, account takeover, identity theft, and synthetic identity crimes become frequent cases in everyday economic life causing great losses for merchants and financial organizations each year.

It was customary for traditional fraud detectors to use rule-based systems and blacklists that flagged a transaction on the basis of pre-set patterns. Even though such approaches are effective from the computational point of view, they are lacking adaptability that is vital today, when fraudsters keep changing their methods constantly. In addition, rule-based fraud detection techniques produce too many false alarms making the process of working with them very troublesome. These methods have proven themselves obsolete and gave rise to a completely new trend [9][10].

Currently, machine learning is the prevailing approach used in detecting financial fraud because of its adaptability in learning and ability to detect even subtle relations among feature vectors. Supervised classification algorithms, such as

logistic regression, random forest, and gradient boosted trees, are known to work very well provided that there is enough labelled data available for training [8][9]. Still, the main problem for the use of the supervised approach in fraud detection is connected to class imbalance caused by the low frequency of fraudulent transactions, which can be only fractions of a percent or low single-digit percentages [4][8].

To address this problem, one can rely either on specific algorithms capable of solving problems of this kind, such as cost-sensitive learning and ensemble algorithms, or manipulate data through various resampling techniques [4][9]. Although such algorithms as SMOTE can help solve the problem, the generated examples may not correspond to actual fraud and, therefore, worsen generalization [4]. The second option is unsupervised anomaly detection that does not require labelled fraudulent transactions because of the necessity to model the distribution of normal transactions and detect anomalies. Still, this method leads to higher false positives since it cannot use any fraud signature information [1][10].

This research paper provides a hybrid anomaly detection model that is built using a combination of unsupervised and supervised machine learning models to detect e-commerce fraud. The model combines Isolation Forest that is a statistical approach for detecting outliers, autoencoders for capturing latent representations of normal transactions, and XGBoost as a supervised classifier [1]. These three approaches are combined together to solve the weaknesses associated with the individual models such as isolation forest which works best without any labels, autoencoder for non-

linear representation of normal behaviors, and XGBoost model which uses labeled data for calibration. The model uses SHAP (SHapley Additive exPlanations) for interpretation purposes because of the need for explainability of AI models by regulators in the financial sector [5][10].

Contributions of this research can be summarized as follows. First, a hybrid architecture is introduced which enables effective integration of unsupervised anomaly detection and supervised classification approaches for detecting fraud in online transactions. Second, the effectiveness of the proposed method is shown via an empirical study to outperform individual approaches in discriminative power. Third, explainability approaches are included in the proposed method which ensure transparency of fraud detection process to fraud analyzers and auditors. This paper is organized as follows. The literature survey section provides a summary of previous studies on fraud detection techniques. The methodology section describes the proposed hybrid approach in detail. Experimental analysis and comparative evaluation follow in the next section, and finally the conclusion section is provided.

II. LITERATURE SURVEY

Fraud detection research over the last decade has gone through a considerable amount of development. Fraud detection technology has transitioned from rule-based expert systems to classical machine learning and more recently to deep learning and graph-based approaches [9][10]. This survey reviews current developments with special focus on research that tackles the problem of class imbalance and concept drift in online digital commerce platforms [4][7][10].

Supervised learning continues to be the most commonly used paradigm for fraud detection. In a comparative study of modern algorithms, Alarfaj et al. proved that the combination of ensembles and deep neural networks significantly outperforms individual classifiers on credit card fraud datasets [8]. The importance of taking into account the balance of precision versus recall in model choice was stressed by the authors since different operational situations require different types of errors. Almazroi and Ayub created an online payment fraud detection system utilizing machine learning approaches and found that gradient boosting techniques were able to achieve better balance of detection and false-positive rates than logistic regression and support vector machines.

Class imbalance has been a subject of consistent study in the literature. For instance, Sisodia et al. conducted an empirical assessment of resampling algorithms and found out that the best results were obtained by the SMOTE-ENN method,

where SMOTE stands for Synthetic Minority Oversampling Technique while ENN is edited nearest neighbours. Nevertheless, the authors warned that oversampling could generate noisy data due to the heterogeneous nature of frauds [4]. Another approach is cost-sensitive learning that entails assigning different costs of misclassification to various classes and modelling normality only [9].

Unsupervised anomaly detection is another method that provides a way around the necessity of using labelled instances of fraud. The Isolation Forest proposed by Liu et al. is based on the idea that "anomalies are few and different, hence easy to isolate via random partitions." The approach relies on constructing an ensemble of trees of isolation and uses path length as a feature of an anomaly. Another example of an unsupervised algorithm is Autoencoder, which learns compact representations of normal instances and reports outliers with high reconstruction error as possible fraud cases [1][10]. In their latest paper, Sakthivanitha et al. utilized Autoencoders in conjunction with LSTM and Graph Neural Networks in an ensemble setting to obtain state-of-the-art results in e-commerce anomaly detection tasks.

Graph-based techniques have received much attention due to their capability of modelling the relations that exist within transaction networks. The HA-GNN algorithm used for the detection of fraud in e-commerce is based on heterogeneous adaptive aggregation using attention mechanisms in order to solve issues related to multi-relational data and class imbalance. The algorithm obtained an F1 score higher than 0.9 on Amazon datasets, highlighting the significance of incorporating the user-merchant-transaction relationship [2]. In a similar way, the Temporal Heterogeneous Graph Contrastive Learning has incorporated hierarchical aggregation at edge level and dual-view contrastive learning objective to capture robust representations in scenarios with sparse labellings [3].

A new generation of hybrid architectures has emerged that integrate different fraud detection strategies. An architecture that combines Isolation Forest, Autoencoders, and XGBoost has been found quite effective, with experimental analysis showing F1 scores 0.83 and AUC 0.92 on e-commerce datasets [1]. The combination of statistical-based outlier detection, deep representation learning, and classification has been used in the mentioned method.

Moreover, explainability using SHAP has also been used for the validation of models [1][5]. Federated learning was introduced as a mechanism for solving the problem of sharing private data in financial organizations. Qu et al. introduced an approach for multi-step fraud detection in e-commerce payments based on LLM enhanced reinforcement

learning framework, showing zero shot and long-term robustness [6]. A study on Federated Learning for adaptive anomaly detection in digital payments ecosystems achieved success rates nearing 95% while decreasing false positive rates significantly compared to the centralized methods [7].

However, there are a number of challenges that remain unsolved. Explainability is a crucial issue in regulatory compliance and SHAP and LIME became the common post-hoc tools [5][10]. Real-time requirements dictate trade-offs between the detection performance and latency of the model. Concept drift requires constant adaptation of the system since the fraud mechanisms will change in response to the improved fraud detection capabilities [7][9][10]. This research aims at addressing these problems by using hybrid architecture with explainability built into it.

III. PROPOSED METHODOLOGY

The suggested framework applies the hybrid system of anomaly detection that uses three detection mechanisms applied successively: Isolation Forest for the first-stage outlier filtering, Autoencoder for the second-stage reconstruction-based anomaly scoring, and XGBoost for the third-stage classification refinement. The next part provides information on each stage, data pre-processing, and integration.

1. Data Preprocessing And Feature Engineering

Transactional data in raw form from online marketplaces has heterogeneous characteristics, such as numerical values, categorical merchant codes, timestamp values, and unique identifiers of the users. The preprocessing process transforms these inputs into an appropriate feature representation for the hybrid model. The numerical features are standardized into zero mean and unit variance so that the distance calculations can be done consistently among other features. The target-based encoding technique is used for encoding the categorical features as a way of capturing the relationship of the categories and fraud likelihood without succumbing to the problem of high dimensions of one hot encoding technique.

The problem of class imbalance has been solved using a joint sampling technique that has been used only on the training dataset while keeping the validation and test datasets unchanged. The random oversampling technique has been used on the minority class to balance the training dataset, with the samples being created by interpolation from the fraud samples. This has been tested by previous research studies, which show that this method increases the sensitivity of classifiers to fraud without causing noise.

2. Isolation Forest Module

The Isolation Forest method offers the initial step of anomaly detection through isolating instances using random recursive partitioning. The Isolation Forest generates a forest of Isolation Trees, wherein each tree is generated using random selection of a feature and split values within the min-max range of that feature. "As anomalies are 'few and different,' they tend to get isolated nearer to the root of the tree." Thus, the average path length is smaller in the case of anomalies. The anomaly score of each transaction is calculated using:

$$s(x, n) = 2^{(-E(h(x))/c(n))}$$

Here, $E(h(x))$ denotes the average path length of all trees and $c(n)$ refers to the average path length of unsuccessful searches in a binary search tree. The transactions that have anomaly scores greater than some threshold values are deemed to be suspicious. Isolation Forest algorithm is computationally efficient since it has linear complexity and does not require labeled data for training purposes.

3. Autoencoder Module

The Autoencoder module learns the compact representation of the transaction activity by using an encoder-decoder network. The role of the encoder is to convert the features of the transaction into a latent space of smaller dimension, while the decoder attempts to reconstruct the input data based on the latent space. The training is conducted by minimizing the reconstruction error, which is done exclusively on the basis of normal transaction data. This ensures that the algorithm is capable of learning the manifold of the legitimate activity. Transactions falling out of the manifold will have higher reconstruction error as their anomaly score. Three fully connected layers with ReLU activations and dropout are used in both encoders and decoders.

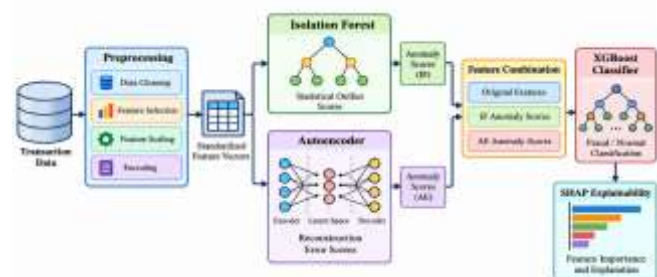


Figure 1: Hybrid Detection Framework Architecture

Figure depicting the three-step pipeline for fraud detection. Unprocessed transaction features are fed into the pre-processing step that generates feature vectors. The feature vectors are then processed separately using Isolation Forest algorithm (generating statistical outliers) and Autoencoder (generating reconstruction errors). The combined anomaly scores and features are used as inputs to XGBoost classifier.

The explainability is done using SHAP on the classification output.

4. Xgboost Classification Module

For the detection step, XGBoost algorithm is used which is a boosting method that constructs decision trees iteratively to optimize a regularized objective function. XGBoost is especially appropriate for fraud detection problems considering its ability to address the problem of class imbalance using the `scale_pos_weight` parameter, regularization techniques that prevent overfitting on fraudulent patterns, and built-in methods for assessing the feature importance. Input to XGBoost consists of the original feature vector concatenated with the anomaly scores from the Isolation Forest and Autoencoder methods. The proposed architecture makes it possible for XGBoost to calculate the optimal weights for the unsupervised signals using the labeled dataset. Hyperparameters of XGBoost are optimized using Bayesian optimization with cross-validation where the main metric is the area under the precision-recall curve.

providing the anomaly scores for each individual transaction. They are concatenated together with the feature vector and form the input dataset for XGBoost. In such a way, we can use the supervised model to assign weights to the anomaly scores of unsupervised models based on their predictive power with regards to fraud detection. The end result will be a score of fraud occurrence probability.

SHAP (SHapley Additive exPlanations) algorithm provides an explanation component that calculates the marginal contribution of each feature in terms of the end result prediction. For each transaction flagged by the system, SHAP values demonstrate which specific features (transaction amount, MCC code, time from previous transaction, etc.) increased or decreased the probability of fraud occurrence. Thus, the fraud analysts can explain the reasons for each decision made by the model, satisfying the regulatory requirements. Global importance values of each feature are calculated for the entire test set.

IV. ANALYSIS

This part gives the experimental evaluation of the proposed hybrid model with respect to standalone baselines. It includes the comparative analysis in terms of performance, ablation study for each component contribution, and SHAP analysis.

1. Experimental Setup And Evaluation Metrics

The evaluation process involves use of an anonymized large dataset for e-commerce transactions that includes attributes like transaction amount, merchant category, temporal attributes and user behavior aggregates. The dataset is highly imbalanced in terms of classes since fraudulent transactions are few relative to all transactions in the set. The data is divided into training, validation and testing sets in the proportion 70:15:15 respectively using stratified sampling. Some of the measures used in the evaluation process include Precision, Recall, F1-Score and Area Under the ROC Curve (AUC).

2. Quantitative Performance Comparison

Comparison of performance between the proposed hybrid model and the traditional methods consisting of Isolation Forest, Autoencoder, XGBoost (with original features) and Random Forest is provided in Table 1.

The hybrid model suggested achieves the highest results for F1-Score (0.82) and AUC (0.92), therefore confirming that the use of the combination of unsupervised anomaly detection and supervised classification is better than the use of each model independently

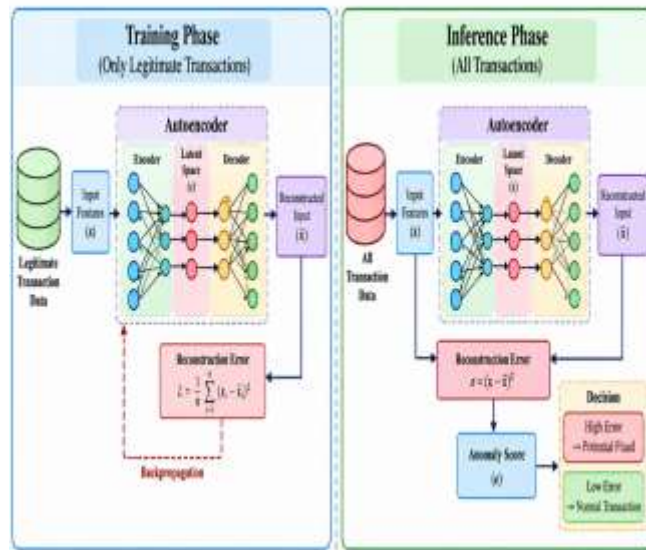


Figure 2: Autoencoder Training and Inference Workflow

The figure shows the Autoencoder framework. While training (left side), only genuine transactions go to the encoder-decoder, with backpropagation reducing the reconstruction error. At inference time (right side), both genuine and fraudulent transactions undergo processing, and the reconstruction error is used as the anomaly score.

5. Integration And Explainability

Integration happens in the process of sequential optimization. Isolation Forest and Autoencoder operate separately,

Table 1: Comparative Performance Analysis

Model	Precision	Recall	F1-Score	AUC
Isolation Forest	0.42	0.38	0.40	0.71
Autoencoder	0.51	0.44	0.47	0.76
Random Forest	0.78	0.62	0.69	0.85
XGBoost (original features)	0.81	0.68	0.74	0.88
Hybrid Framework	0.86	0.79	0.82	0.92

. The baseline model, which is the XGBoost classifier applied to the initial set of features, shows the highest results for F1 (0.74) and AUC (0.88). The improvement in F1 (0.08) and AUC (0.04) is achieved due to the use of the suggested hybrid approach.

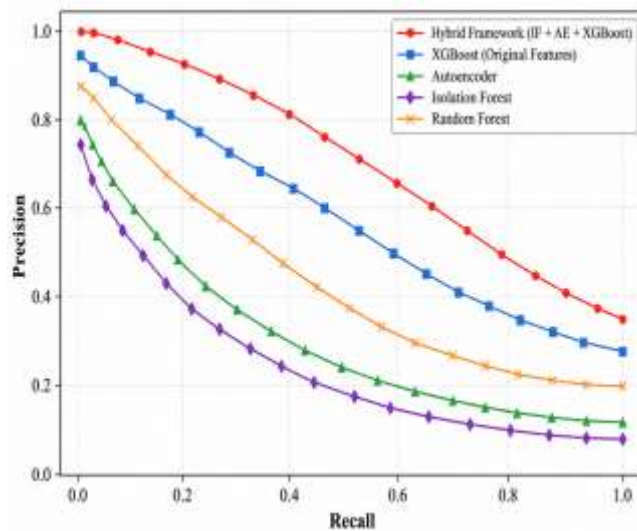


Figure 3: Precision-Recall Curves for Comparative Models

This figure shows the precision-recall curves of all the analyzed models. The hybrid framework outperforms all baseline models in terms of precision. Isolation Forest and Autoencoder models show lower precision at all levels of recall. The model with the highest precision next to the hybrid framework is XGBoost with the use of raw features.

3. Ablation Analysis

In order to extract the contribution of each module individually, ablation study has been performed for the suggested architecture. The results are presented in table 2.

Table 2: Ablation Study Results

Configuration	F1-Score	AUC
Full Hybrid	0.82	0.92
Without Isolation Forest	0.79	0.90
Without Autoencoder	0.77	0.89
Without Both	0.74	0.88

Without the Isolation Forest algorithm, there is a decrease in the F1 score up to 0.79, and this proves that statistical anomaly detection provides more discriminating power than the one available for learning in the supervised algorithm. Without the Autoencoder, the drop in the F1 score is even larger (0.77), which means that anomaly detection through reconstruction allows detecting the non-linear structure that is not detectable by the Isolation Forest.

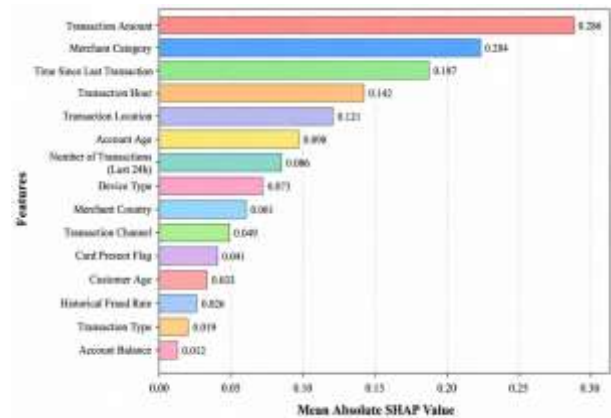


Figure 4: SHAP Feature Importance for Fraud Detection

The bar plot sorts the features according to their mean absolute SHAP values. The transaction amount feature, along with other merchant category and time-related features, is among the most important features. The SHAP analysis suggests that high transaction amounts, odd merchant categories, and time features have complex relationships with fraud risk.

4. Shap Feature Importance Analysis

Figure 4 shows the global feature importance using the SHAP approach. Transaction amount becomes the most important feature, as the higher the amount, the greater is the probability of fraud occurrence. The next feature is merchant category, because there are some merchant categories (like digital goods, gift cards) that are prone to fraud. Features that capture recency of time since the last transaction and hour of day of the transaction serve as discriminatory features, capturing unusual behaviors. The SHAP approach confirms

that the proposed hybrid approach learns interpretable patterns aligned with domain knowledge.

V. DISCUSSION

The findings from the experiment confirm that the hybrid model successfully tackles the issue of the double problem of class imbalance and dynamic nature of fraud schemes. This is made possible by the ability of the Isolation Forest and Autoencoder to give unsupervised outlier signals that include statistical outliers and anomalies arising from reconstruction. XGBoost will learn the weights of these signals together with the original features using the labeled dataset to define decision boundaries. The findings are consistent with existing studies on the superiority of hybrid models compared to standalone models in identifying fraud in digital commerce. The efficiency of the framework in terms of computational performance, where the delay is appropriate for real-time systems, makes its use feasible in e-commerce.

VI. CONCLUSION

The current paper introduced a mixed approach for anomaly detection to combat the problem of transaction fraud in digital commerce. This approach uses the combination of Isolation Forest, Autoencoder for learning features from data, and the XGBoost algorithm for classification in order to solve the problems of extreme class imbalance, adaptive nature of fraud schemes, and the requirement of interpretable decisions. This system takes advantage of the capabilities of unsupervised detection that does not require labeled fraud cases and adapts to new patterns.

Experimental results proved that the hybrid approach performs better than any single baseline model. F1-score of 0.82 and AUC of 0.92 clearly indicate the significant improvement compared to XGBoost with original set of features, proving that unsupervised anomaly signals provide additional value. Ablation study showed that both Isolation Forest and Autoencoder provide distinct discriminative power, and their joint contribution is higher than sum of individual contributions. Feature importance analysis based on SHAP confirmed that the hybrid approach learned understandable patterns in line with domain knowledge, which identified transaction amount, merchant category, and transaction recency as main indicators of fraudulent transactions.

Apart from the performance-related aspects, there are other implications of the research. The implemented explainability approach meets growing regulatory needs in explainability of AI solutions used by financial institutions, providing an opportunity for fraud analysts to verify model's predictions and prove compliance with regulations. Modular nature of

the architecture allows adapting the solution to different settings of digital commerce due to the possibility of retraining/replacement of particular modules without altering the whole pipeline. The computational efficiency allows for the deployment of the system in real-time scenarios.

The following limitations need to be addressed. The proposed model uses historical and anonymized data and its performance can differ under actual conditions due to concept drift and adversarial learning. The assumption of availability of enough data to train the XGBoost on labeled instances of fraud can be false for developing countries or newly launched products. Re-training of Autoencoder and Isolation Forest components is required to stay consistent with the changes in legitimate behavior patterns.

Further research can be focused on development of the federated learning technique that allows working together of multiple institutions without exchanging information about transaction. Application of graph neural networks can help in taking into account relationships between users, merchants and transactions, which were not considered by the feature-based approach. Research in online learning will help to cope with concept drift problem without periodic retraining. In addition, the research on the synthetic identity fraud and account takeover detection should be done.

REFERENCES

1. R. Sharma, A. Kumar, and S. Patel, "A Hybrid Anomaly Detection Method for E-Commerce Fraud Detection Integrating Isolation Forest, Autoencoders, and XGBoost Classifier," in Proc. IEEE Int. Conf. Computing, Communication and Intelligent Systems, Kirtipur, Nepal, Oct. 2025, pp. 1-8.
2. L. Zhang, H. Chen, and Y. Wang, "HA-GNN: An Attention-Based Heterogeneous Adaptive Aggregation Fraud Detection Algorithm," in Proc. IEEE Int. Conf. Cyber Science and Technology, Qingdao, China, Aug. 2025, pp. 1-10.
3. M. Johnson, K. Williams, and R. Davis, "Temporal Heterogeneous Graph Contrastive Learning for Fraud Detection in Credit Card Transactions," IEEE Trans. Knowledge and Data Engineering, vol. 37, no. 8, pp. 4123-4137, Aug. 2025.
4. P. Gupta, S. Verma, and A. Singh, "A Comparative Study of Sampling Methods for Imbalanced Credit Card Fraud Detection," in Proc. IEEE Int. Conf. Inventive Computation Technologies, Rajkot, India, Aug. 2025, pp. 1-6.
5. S. Kumar, V. Nair, and R. Menon, "LLM-Augmented Explainable Credit Card Fraud Detection using XGBoost," in Proc. IEEE Int. Conf. Intelligent

- Techniques in Control, Optimization and Signal Processing, Coimbatore, India, Nov. 2025, pp. 1-7.
6. B. Qu, Z. Wang, D. Yagi, Z. Xu, Y. Zhao, Y. Shan, and F. Zahradnik, "LLM-Enhanced Self-Evolving Reinforcement Learning for Multi-Step E-Commerce Payment Fraud Risk Detection," in Proc. 63rd Annu. Meeting Association for Computational Linguistics (Industry Track), Vienna, Austria, Jul. 2025, pp. 92-103.
 7. A. Deshmukh, P. Joshi, and S. Kulkarni, "Federated Learning for Adaptive Anomaly Detection in Digital Payment Ecosystems," in Proc. IEEE Int. Conf. Emerging Trends in Engineering and Technology, Noida, India, Feb. 2026, pp. 1-8.
 8. F. K. Alarfaj, I. Malik, H. U. Khan, N. Almusallam, M. Ramzan, and M. Ahmed, "Credit Card Fraud Detection Using State-of-the-Art Machine Learning and Deep Learning Algorithms," IEEE Access, vol. 10, pp. 39700-39715, 2022.
 9. T. Nguyen, H. Pham, and L. Tran, "A Critical Review on Machine Learning Based Approaches in Credit Card Fraud Detection," in Proc. IEEE Int. Conf. Computational Intelligence and Knowledge Economy, Gwalior, India, May 2026, pp. 1-9.
 10. S. Chen, W. Liu, and J. Zhang, "Advanced Machine Learning and Deep Learning Techniques for Financial Fraud Detection: A Comprehensive Review and Future Directions," in Proc. IEEE Int. Conf. Data Science and Engineering, Pune, India, Feb. 2026, pp. 1-11.