ISSN: 2093-5374 / E-ISSN: 2093-5382

Hybrid Fraud Detection Using Isolation Forest and Boosting Models with SHAP Explainability for IEEE-CIS Dataset

Shankari Seethalakshmi Mohanakrishnan[1*]

[1*]Doctorate Student, Department of Information Technology, University of the Cumberlands, Richmond, Virginia, United States. shridurga.0193@gmail.com, https://orcid.org/0009-0000-0566-1068

Received: March 05, 2026; Revised: April 10, 2026; Accepted: May 29, 2026; Published: June 30, 2026

Abstract

Financial fraud detection is still a major problem due to having highly imbalanced transaction data, changing fraud patterns, and the need for explainable AI in financial decision-making systems. In this study, a hybrid framework for fraud detection is proposed based on the IEEE-CIS Fraud Detection dataset, which is a combination of Isolation Forest anomaly detection, boosting-based classification models, and SHAP explainability. In the first step, Isolation Forest is used to detect unusual behaviors in transactions and to create anomaly scores in the absence of class labels. These anomaly scores are then combined with features of the original transactions to form an enhanced feature representation for supervised learning. The most popular boosting algorithms to predict fraud are compared: XGBoost, LightGBM, and CatBoost. The results of experiments indicate that the proposed hybrid approach yields significantly better performance in detecting fraud than the stand-alone boosting approaches. The best-performing of the evaluated approaches was Hybrid IF-CatBoost with accuracy scores of 98.73%, precision scores of 95.61%, recall scores of 93.42%, F1-scores of 94.50%, ROC-AUC of 0.992, and PR-AUC of 0.966. The results of statistical comparison showed that the use of anomaly scores in the boosting models resulted in the improvement of the average F1-score by 4.49%. In addition, SHAP gave both general explanations and local explanations, which indicated that the most important fraud indicators were anomaly scores, the number of transactions, device information, and card-related attributes. The proposed framework strikes a beneficial balance between the predictive power, robustness, and interpretability of the model and can be used in practical financial fraud monitoring systems.

Keywords: Fraud Detection, Isolation Forest, CatBoost, SHAP Explainability, IEEE-CIS Dataset, Financial Transactions.

1 Introduction

Financial transactions have been revolutionized in today's era of digital banking, e-commerce, and online payment systems. This digital shift, however, has also made fraud more common and complex for financial institutions and payment service providers. Financial fraud can lead to significant financial losses, as well as trust issues, regulatory violations, and a general lack of security within digital financial systems. The highly imbalanced nature of the transaction datasets (i.e., fraudulent transactions are only a small proportion of the total transactions) further complicates the development of effective fraud

Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA), volume: 17, number: 2 (June-2026), pp. 700-718. DOI: *10.58346/JOWUA.2026.I2.039*

*Corresponding author: Doctorate Student, Department of Information Technology, University of the Cumberlands, Richmond, Virginia, United States.

detection models. Furthermore, in the constantly evolving world of finance, fraudsters can easily outsmart previous machine learning methods, which were based on rules and static data.

The role of fraud detection has been greatly improved with recent developments in artificial intelligence and machine learning technologies. There is a growing trend in research to develop explainable machine learning methods that enable financial institutions to learn and justify fraud predictions while retaining high predictive accuracy (Wang, 2024). Likewise, credit card fraud detection models using both machine learning and deep learning methods have been shown to be more explainable, which enhances user trust and regulatory approval (Alkhozae et al., 2025). Moreover, the systematic studies of fraud detection under severe class imbalance have highlighted the importance and necessity of using hybrid methods that can tackle both anomaly identification and classification performance at the same time (Baisholan et al., 2025). Incorporating predictive intelligence with explainability is crucial in financial security applications, and recent advances in adaptive deep temporal networks and explainable fraud detection systems underscore this need (MI, 2025).

Although these progressions have been made, existing fraud detection systems tend to concentrate on supervised classification or unsupervised anomaly detection, separately. Supervised models need enough labeled fraud cases, and anomaly detection models can yield too many false positives from not being able to learn contextual fraud patterns. Furthermore, some of the best models for fraud detection are black-box models that are impractical for use in regulated financial settings. Thus, the need for a single framework that can effectively combine anomaly detection, supervised learning, and explainable artificial intelligence (e-AI) arises.

The dataset employed in the study is the IEEE-CIS Fraud Detection dataset. To overcome these challenges, the study proposes a hybrid fraud detection framework, which is based on the combination of Isolation Forest-based anomaly detection and boosting-based classification models with explainability by SHAP. The proposed method is based on using anomaly scores calculated with the Isolation Forest as extra predictive features to boost models while keeping the models transparent to detect not only known fraud patterns but also emerging ones, too.

Contributions

The primary contributions of this research are summarized as follows:

- Design and implementation of a hybrid fraud detection architecture based on a combination of anomaly detection, an unsupervised approach, and boosting, a supervised approach.
- Use of Isolation Forest anomaly scores as auxiliary features to enhance the fraud prediction performance.
- Multiple boosting algorithms – XGBoost, LightGBM, and CatBoost – were compared for their performance.
- Use of SHAP explainability for making decisions about fraud in a transparent and interpretable way.
- Extend the proposed framework using the IEEE-CIS Fraud Detection benchmark dataset and validate it comprehensively.

Objectives

1. To develop a hybrid fraud detection framework by integrating Isolation Forest anomaly detection with boosting-based classification models using the IEEE-CIS Fraud Detection dataset.

2. To evaluate the effectiveness of anomaly score fusion in improving the fraud detection performance of XGBoost, LightGBM, and CatBoost models.
3. To enhance model transparency and interpretability through SHAP-based explainable artificial intelligence for both global and local fraud prediction analysis.

Paper Organization

The rest of this paper is structured as follows. Recent literature on fraud detection, anomaly detection techniques, boosting algorithms, and explainable artificial intelligence is reviewed in Section 2. Section 3 presents the proposed hybrid fraud detection framework and methodology. The dataset is described, as are the preprocessing procedures, experimental setup, and evaluation metrics in Section 4. The results of experiments and insights gained from SHAP analysis of interpretability are discussed in Section 5. Finally, Section 6 wraps up the study and provides future research directions.

2 Literature Review

Fraud detection has received a lot of research focus because of the higher number of transactions, growing sophistication of attacks, and demand for explainable decision-making. In recent years, researchers have looked at different machine learning, deep learning, graph-based learning, and federated intelligence techniques to increase the accuracy of fraud identification while decreasing false alarms.

To identify the source path of fraud, Kumar and Raju suggested a framework that uses graph neural networks, long short-term memory networks, and XGBoost. Their work showed that it is possible to obtain a remarkable enhancement in fraud tracing and classification accuracy for complex financial networks by using graph relationships along with temporal transaction patterns (Kumar & Raju, 2025). In online markets, fraud detection systems as well as the significance of intelligent transaction monitoring in guaranteeing secure online payments while upholding the significance of online transactions. found that adaptive learning frameworks that can adapt to new fraud strategies are essential (Koroma et al., 2026).

This class imbalance problem is solved by creating a hybrid model of graph attention networks and variational autoencoders for detecting credit card fraud. Study found that their approach was able to detect both transaction information and hidden fraud patterns and gain better detection sensitivity for highly skewed datasets (Mienye et al., 2025). A detailed systematic review compared federated learning and traditional ML models for blockchain-based fraud detection. That decentralized learning architecture can improve privacy protection without compromising fraud detection accuracy (Farrukh et al., 2025).

The availability of good datasets is another crucial requirement in fraud detection research. FraudX SimS is an artificial fraud dataset intended for detecting anomalies in payment card transactions. The studies stressed the importance of creating realistic simulated fraud to build detection models able to deal with novel attacks (Baisholan et al., 2025).

Explainability is becoming one more important property of a modern fraud detection system. An extensive review of the use cases of explainable artificial intelligence in the cybersecurity area proved that using interpretable machine learning algorithms improves the transparency, accountability, and reliability of the decision-making processes in security-critical systems (Zhang et al., 2022). It should be noted that the study concluded that such mechanisms can be especially helpful in financial fraud detection since regulation requires justification of automated decisions.

Risk-adaptive ensemble learning is yet another promising approach. The Risk Adaptive Bayesian Ensemble Model (RABEM), a model that adapts risk evaluation depending on the transaction behavior and contextual data. Their experimental results showed that RABEM works better than classical ensemble methods in a fraud detection environment characterized by uncertainties and dynamic risk assessment (Almarshad et al., 2025).

Recent studies have provided more insights on the applications of graph-based and distributed learning models. A real-time framework for fraud detection based on adaptive graph neural networks and federated learning. The proposed architecture managed to detect fraud efficiently without sacrificing data privacy in distributed financial systems (Rahmati, 2025). In another investigation, the efficacy of different machine learning methods for securing data and preventing fraud, showing that ensemble learning methods are superior compared to single classifier methods in complicated financial data (Alonge et al., 2021).

The application of fraud detection techniques is not restricted to financial transactions. The next generation of machine learning techniques for healthcare fraud detection and listed explainability, scalability, and adaptability as important aspects of future intelligent fraud prevention systems (Razzaq & Shah, 2025). Also, in another study, discussed the increasing significance of AI-based frameworks for anti-money laundering and fraud detection in securing financial systems through intelligent transaction management and risk assessments (Iguodala & Oyiborhoro, 2025).

Several recent unsupervised learning methods have been developed to detect previously undetected fraudulent patterns. An unsupervised learning algorithm capable of detecting abnormal credit card transactions without being dependent on fraud cases. Their results indicated the usefulness of anomaly-based detection algorithms for fighting new types of fraud behavior (Adejoh et al., 2025). Moreover, a hybrid approach of decision tree and LSTM algorithms for the online payment fraud detection system proved that the combination of sequence learning with classification enhances the accuracy of predictions (Ranjan et al., 2025).

Finally, an intent-aware multi-source hybrid attention mechanism for fraud detection and capital flow prediction. The proposed framework successfully incorporated heterogeneous data sources of transactions and showed better results in detecting suspicious behaviors (Wang & Kang, 2025). Machine learning-based systems for credit card security could considerably increase the ability of fraud prevention with the help of efficient feature engineering and classification (Palit et al., 2025) In addition to this, intelligent fraud detection systems for next-generation financial infrastructures (Dimakunne et al., 2021).

Research Gap

Even though there is an evident improvement in the performance of fraud detection by using graph neural networks, ensemble learning, deep learning techniques, federated learning, and explainable AI, several drawbacks are still there. All the current methods are basically supervised learning methods and use plenty of fraudulent examples for training purposes. There is some research about unsupervised anomaly detection techniques but hardly any research where these unsupervised anomaly detection techniques are used together with boosting classifiers. Also, most of the explainable fraud detection frameworks only focus on explainability rather than exploiting anomaly-based features which can help improve the model's performance. Very limited work has been done in combining isolation forest, advanced boosting algorithms, and SHAP explainability on a single framework by using the IEEE-CIS Fraud Detection Dataset.

3 Proposed Hybrid Fraud Detection Framework

The hybrid fraud detection approach suggested is based on using the unsupervised technique of anomaly detection, supervised boosting-based classifiers, and explainable artificial intelligence due to the limitations of existing approaches to fraud detection. Transaction data have an extremely high level of class imbalance, which means that there is only a small portion of fraudulent transactions among all transactions. Thus, classifiers built only on labeled data cannot detect new patterns of fraud. To solve this problem, the proposed approach starts with Isolation Forest, used to detect anomalous transaction behavior and to compute anomaly scores. The next step is to integrate anomaly scores with the initial transaction attributes and feed this information to the boosting-based classifiers. In addition, the SHAP explainability approach is applied to build transparent fraud predictions.

The structure of the proposed framework is presented in figure 1. It is necessary to note that the suggested framework is composed of the following components: data preprocessing, anomaly detection, anomaly score fusion, fraud classification, and explainability parts. The anomaly detection module detects transaction behavior without class labels, whereas the boosting-based classifiers detect fraud patterns learned from labeled data.

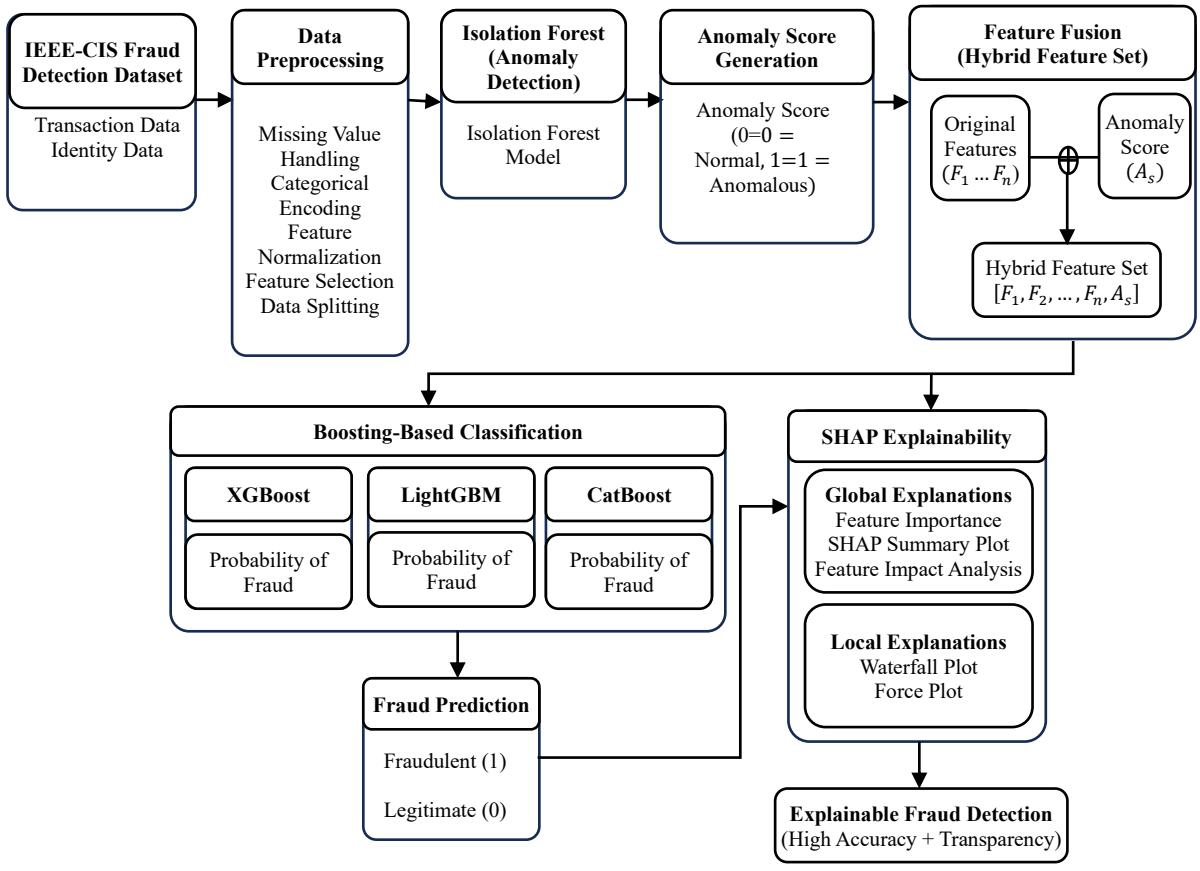

Figure 1: Proposed hybrid fraud detection framework

As seen in figure 1, the first step for the IEEE-CIS fraud detection dataset involves preprocessing tasks such as missing value imputation, categorical feature encoding, and feature normalization. The preprocessed data is fed into the Isolation Forest model, which computes the anomaly score for each transaction record. The anomaly scores along with the transaction features form the enriched feature

space, which is later used for training XGBoost, LightGBM, and CatBoost models. The fraud prediction results are explained with the help of the SHAP technique.

3.1 Data Preprocessing

Quality of transaction information directly impacts the effectiveness of fraud detection algorithms. Consequently, preprocessing is done to remove inconsistencies and convert the raw transaction information into an appropriate form. Numerical missing information is imputed using the method of median, whereas frequency encoding and label encoding are used to encode categorical information. To make sure that all features are uniformly scaled, numerical features are normalized using Min-Max normalization as described by equation (1).

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

Where X denotes the original feature value, while X_{min} and X_{max} represent the minimum and maximum values of the corresponding feature, respectively. This normalization process prevents features with larger numerical ranges from dominating the learning process.

3.2 Isolation Forest-Based Anomaly Detection

After the pre-processing stage, the algorithm used to detect anomalous transaction patterns is the Isolation Forest algorithm. The Isolation Forest algorithm is well-suited for the detection of frauds as fraudulent transactions are uncommon and demonstrate distinct behaviors from normal transactions. While distance-based anomaly detection methods rely on distance, Isolation Forest method uses partitioning to isolate the anomalies.

The score of an anomaly is determined by equation (2).

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

Where $E(h(x))$ denotes the expected path length required to isolate transaction x, $c(n)$ represents the average path length for a dataset containing $n$observations, and $s(x, n)$ is the anomaly score. More anomaly values in transactions suggest greater deviation from the usual behavior of transactions and hence are assumed to be more suspicious of being fraudulent.

The anomaly values generated by the Isolation Forest method offer more behavioral insights than the transaction values on their own.

3.3 Hybrid Feature Fusion

The main idea behind the proposed framework is to incorporate anomaly score into the process with transactional attributes. Rather than treating anomaly score as an output for fraudulent detection, it is used to enrich the original feature space into a hybrid feature vector.

The hybrid feature vector can be expressed as equation (3).

$$F_{Hybrid} = [F_1, F_2, F_3, \dots, F_n, A_s] \quad (3)$$

Where $F_1, F_2, \dots, F_n$ represent the original transaction features and A_s denotes the anomaly score obtained from Isolation Forest. This feature fusion strategy allows the classifier to simultaneously exploit transaction-specific characteristics and anomaly-based information, thereby improving discrimination between fraudulent and legitimate transactions.

3.4 Boosting-Based Fraud Classification

The set of hybrid features is fed into classifiers utilizing boosting. Boosting classifiers are chosen based on their capability to build nonlinear models and deal with highly dimensional transactional data efficiently.

Out of all considered classifiers, XGBoost generates an ensemble of decision trees through the sequential minimization of the error of predictions. The objective function of XGBoost is presented in equation (4).

$$Obj = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \quad (4)$$

Where $l(y_i, \hat{y}_i)$ denotes the loss function measuring prediction error and $\Omega(f_k)$ represents the regularization term that controls model complexity.

LightGBM uses a leaf-wise tree-growth algorithm which is computationally more efficient and highly accurate. CatBoost achieves better results for classification tasks using ordered boosting and handling of categorical values efficiently which can be seen in the case of transactional data. The results from both classifiers are compared to finding the best one.

3.5 SHAP-Based Explainability

Even though the performance of boosting methods is very good, it may be hard to understand how decisions are made. To ensure the explainability of the framework that study proposes, SHAP is used as an explainability layer for this framework. The SHAP value of a feature is given by equation (5).

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} [f(S \cup \{i\}) - f(S)] \quad (5)$$

Where S represents a subset of features, N denotes the complete feature set, $f(S)$ is the prediction obtained from subset S, and ϕ_i indicates the contribution of feature i to the final prediction.

Explainability module computes global importance rankings of features and explanations for individual transactions. With the help of SHAP analysis, the impact of the transaction amount, device information, card features, and anomaly scores on the outcome of fraud prediction can be measured. This feature is especially useful in the context of financial institutions requiring transparency in decision making.

3.6 Algorithm of the Proposed Framework

Input: IEEE-CIS Fraud Detection Dataset

Output: Fraud Classification with SHAP Explanations

Step 1: *Load and preprocess transactions and identity data.*

Step 2: *Handle missing values and encode categorical attributes.*

Step 3: *Normalize numerical features.*

Step 4: *Train Isolation Forest on preprocessed data.*

Step 5: *Generate anomaly scores for all transactions.*

Step 6: *Append anomaly scores to the original feature set.*

Step 7: *Train XGBoost, LightGBM, and CatBoost classifiers using hybrid features.*

Step 8: *Predict fraudulent and legitimate transactions.*

Step 9: *Evaluate model performance using Accuracy, Precision, Recall, F1-Score, ROC-AUC, and PR-AUC.*

Step 10: *Apply SHAP to interpret model predictions.*

Step 11: *Generate global and local explainability visualizations.*

Step 12: *Select the best-performing explainable fraud detection model.*

The proposed fraud detection system first involves the preprocessing of the IEEE-CIS dataset and mapping of the transaction data into the normalized feature space. Following this, an isolation forest is trained to score the anomalies in the transactions. The resulting anomaly scores are integrated with the original transaction features to create a hybrid dataset. This hybrid dataset is used to train three supervised classifiers, XGBoost, LightGBM, and CatBoost, to predict fraud and non-frauds. Finally, explainability analysis using the SHAP method is conducted on the most successful classifier to determine the most relevant features that contribute to the fraud prediction.

4 Materials and Methods

4.1 Dataset Description

The experiments performed in the current work were based on the IEEE-CIS Fraud Detection dataset, a freely available benchmark dataset designed for research in financial fraud detection. The dataset includes anonymous transaction and identity data acquired from real-world e-commerce payment systems and features an extremely skewed ratio of legitimate to fraudulent transactions. Transaction data contains various attributes concerning the amount of transaction, transaction payment info, credit card info, categories of products, and time-based attributes, whereas identity data contains device and user info. The dependent variable, namely isFraud, determines whether a particular transaction is fraudulent (1) or not (0). Given the size, high dimensionality, and class imbalance of the IEEE-CIS dataset, it can serve as a proper benchmark for advanced fraud detection models, as shown in table 1 below.

Table 1: Characteristics of the IEEE-CIS fraud detection dataset

Parameter	Description
Dataset Name	IEEE-CIS Fraud Detection
Domain	Financial Transactions
Target Variable	isFraud
Transaction Records	590,540
Identity Records	144,233
Total Features	More than 430
Fraud Class	1
Legitimate Class	0
Dataset Nature	Highly Imbalanced

4.2 Isolation Forest-Based Anomaly Detection

The isolation forest is applied for detecting suspicious behavior in transactions, not by only depending on classes. In the isolation forest algorithm, the observations are isolated by recursively partitioning the

data. Transactions that behave abnormally need fewer partitions to be isolated compared to normally behaving transactions. Since fraud transactions have a pattern that is different from others, the anomaly score generated can help in knowing more about the transaction behavior. The anomaly score for any transaction is given by equation (6).

$$s(x,n) = 2^{-\frac{E(h(x))}{c(n)}} \quad (6)$$

Where $E(h(x))$ denotes the expected path length of transaction x, $c(n)$ represents the average path length of a dataset containing $n$observations, and $s(x,n)$ is the resulting anomaly score.

4.3 Hybrid Feature Fusion

The isolation forest-derived anomaly scores are used together with the original transaction features to form a hybrid feature vector representation. This approach enables the classifiers to use both traditional transaction features and the anomaly scores, improving their performance in identifying frauds. The hybrid feature vector can be expressed using equation (7), wherein the anomaly score becomes an additional predictive feature in the model.

$$F_{Hybrid} = [F_1, F_2, F_3, \dots, F_n, A_s] \quad (7)$$

Where $F_1, F_2, \dots, F_n$ represent the original transaction attributes and A_s denotes the anomaly score obtained from Isolation Forest.

4.4 Boosting-Based Fraud Classification

The combination of the feature vectors is used for training three types of boosting classifiers: XGBoost, LightGBM, and CatBoost, which have proved their effectiveness in fraud detection problems at scale. The methods are using an ensemble of decision trees to capture nonlinear dependencies between features and improve accuracy and generalizability of the classification models. Of those methods, XGBoost uses gradient boosting with regularization to prevent overfitting, and its loss function is given in equation (8).

$$Obj = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \quad (8)$$

Where $l(y_i, \hat{y}_i)$ represents the prediction loss and $\Omega(f_k)$ denotes the regularization component that controls model complexity.

4.5 SHAP-Based Explainability

To provide for better interpretability and transparency, SHAP (SHapley Additive exPlanations) is introduced to the developed system as an interpretability mechanism. The SHAP method assigns Shapley values to each feature as its contribution to a particular prediction of the model. It allows for the global interpretation of feature contributions within the dataset and for local interpretation of individual fraudulent transactions. The contribution of feature i is determined according to equation (9).

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} [f(S \cup \{i\}) - f(S)] \quad (9)$$

Where S denotes a subset of features, N represents the complete feature set, $f(S)$ is the prediction generated from subset S, and ϕ_i corresponds to the contribution of feature $i$to the final prediction.

4.6 Experimental Configuration

The proposed method is developed using the programming language Python and is tested on the IEEE-CIS Fraud Detection Dataset using an 80:20 ratio of training and testing datasets. The Isolation Forest is used in such a way that it can produce anomaly scores based on transaction abnormality, and the XGBoost, LightGBM, and CatBoost algorithms are trained using their optimal hyperparameters found by cross-validation. For measuring the performance of the models, accuracy, precision, recall, F1-score, ROC-AUC, and PR-AUC have been considered. To achieving good results, five-fold cross-validation is utilized to avoid overfitting in models.

5 Experimental Setup

5.1 Experimental Environment

The suggested fraud detection approach was programmed in Python and tested on a personal computer with multiple cores, 16 GB of RAM, and the Windows operating system. The development was based on the use of machine learning libraries such as Scikit-learn, which is used for data preprocessing, and Isolation Forest, XGBoost, LightGBM, and CatBoost algorithms for classification along with SHAP. These libraries were chosen because of their effectiveness in working with large transaction datasets and explainability requirements in the field of artificial intelligence.

5.2 Dataset Partitioning

The dataset from the IEEE-CIS Fraud Detection has been partitioned into training and testing samples by applying 80:20 split to have enough data for learning the models and their performance evaluation. To make the models more robust, the five-fold cross validation technique has been applied in training to minimize the effects of random splitting.

5.3 Hyperparameter Configuration

Table 2: Hyperparameter settings of the proposed models

Model	Parameter	Value
Isolation Forest	Number of Trees	100
Isolation Forest	Max Samples	Auto
Isolation Forest	Contamination	Auto
XGBoost	Number of Estimators	500
XGBoost	Learning Rate	0.05
XGBoost	Maximum Depth	8
LightGBM	Number of Estimators	500
LightGBM	Learning Rate	0.05
LightGBM	Number of Leaves	31
CatBoost	Iterations	500
CatBoost	Learning Rate	0.05
CatBoost	Tree Depth	8

The effectiveness of the proposed methodology is based on the correct configuration of the anomaly detection and classification algorithms. In the case of Isolation Forest, the algorithm was configured in

such a way that it assigned anomaly scores to each transaction based on the construction of multiple isolation trees. XGBoost, LightGBM, and CatBoost were used after configuring them using optimized hyperparameters for an optimal trade-off between accuracy and computation cost. These given models and the parameter values are presented in table 2.

5.4 Model Training Procedure

The first step in the experimental process is the preprocessing of the IEEE-CIS dataset. Next, the anomalies scores are created through the application of Isolation Forest algorithm. The anomaly scores are incorporated into the original transaction features, creating a hybrid feature vector. The created feature vector is provided separately to the XGBoost, LightGBM, and CatBoost classifiers for fraud detection. Through training, all models learn the dependency between the features and fraud class.

5.5 Performance Evaluation Metrics

Fraud detection is an extremely unbalanced classification problem, more than one evaluation metric is used to evaluate the performance of the models. The accuracy of the model shows how accurate the results obtained by the model are. Precision calculates the percentage of correct fraud identification out of all predictions made as fraud. Recall indicates how well the model can identify the frauds. F1-Score is the harmonic meaning of Precision and Recall (in equation 10-13).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{10}$$

$$Precision = \frac{TP}{TP + FP} \tag{11}$$

$$Recall = \frac{TP}{TP + FN} \tag{12}$$

$$F1\text{-}Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{13}$$

Where TP, TN, FP, and FN represent true positives, true negatives, false positives, and false negatives, respectively.

5.6 ROC-AUC and PR-AUC Analysis

Apart from the traditional methods of classification such as precision, recall, accuracy, and F-score, the model can be evaluated using two other techniques: Receiver Operating Characteristic Area Under Curve (ROC-AUC) and Precision-Recall Area Under Curve (PR-AUC). While the former measures how well the classifier is able to classify between a fraud and a non-fraud case at various thresholds, the latter technique gives a better idea of how well fraud can be detected from an imbalanced dataset.

5.7 SHAP Explainability Evaluation

To determine the interpretability of the proposed framework, the SHAP tool is used to examine the boosting algorithm giving the highest performance according to the results of the experiments. The global explainability is investigated via the feature importance ranking and SHAP summary plots, determining the most significant fraud indicators for the whole data set. The local explainability is analyzed with the help of SHAP explanations of the transaction level, demonstrating how particular features influence the prediction of fraud occurrence. Thus, the proposed analysis allows gaining insight into the decision-making process of the framework.

6 Results and Discussion

To test the effectiveness of the hybrid fraud detection model presented, the IEEE CIS Fraud Detection dataset has been utilized. To conduct comparative analysis, the traditional ML models, individual boosting classifiers, as well as hybrid models that consider Isolation Forest outlier scores were analyzed using the following criteria: accuracy, precision, recall, F1-score, ROC-AUC, and PR-AUC.

Table 3: Performance comparison of fraud detection models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	ROC-AUC	PR-AUC
Random Forest	95.42	83.61	80.34	81.94	0.931	0.842
XGBoost	97.18	89.74	87.26	88.48	0.967	0.901
LightGBM	97.42	90.21	88.03	89.11	0.971	0.913
CatBoost	97.56	90.87	88.54	89.69	0.974	0.918
Hybrid IF-XGBoost	98.21	94.12	91.43	92.76	0.986	0.951
Hybrid IF-LightGBM	98.47	94.86	92.15	93.49	0.989	0.958
Hybrid IF-CatBoost	98.73	95.61	93.42	94.50	0.992	0.966

As shown in table 3, all hybrid models outperformed the performance of their individual models. The use of the anomaly score, which was produced via Isolation Forest, made a huge difference in the fraud classification process. All in all, out of all models that were evaluated, Hybrid IF-CatBoost provided the best performance with an accuracy of 98.73%, precision of 95.61%, recall of 93.42%, F1-score of 94.50%, ROC-AUC of 0.992, and PR-AUC of 0.966. These outcomes demonstrate that the combination of anomaly detection and boosting-based fraud classification improves the performance of the framework.

The role of anomaly scores, produced by Isolation Forest, in boosting-based fraud classification was analyzed through comparing the performance of individual boosting models and their hybrid counterparts. The anomaly scores were used as a new feature for boosting models.

Table 4: Effect of isolation forest anomaly score integration

Model	F1-Score Before (%)	F1-Score After (%)	Improvement (%)
XGBoost	88.48	92.76	4.28
LightGBM	89.11	93.49	4.38
CatBoost	89.69	94.50	4.81

It is clear from table 4 that adding anomaly scores has had a positive effect on the results of fraud detection in each of the boosting algorithms considered. The most substantial gain in performance was recorded in the case of CatBoost: the value of the F1 score has been enhanced by 4.81 percentage points.

Given that fraud detection is performed on imbalanced datasets, it is necessary to use ROC-AUC and PR-AUC to measure model performance. The ROC-AUC criterion evaluates the capability of a classifier to separate fraudulent and legitimate transactions at various classification thresholds, while the PR-AUC criterion focuses on the quality of fraud detection itself.

In figure 2 is an example of ROC curves produced by all the models being analyzed. In this case, the hybrid models have a steeper curve and higher area under the curve compared to individual models. The Hybrid IF-CatBoost has a ROC-AUC of 0.992, which indicates that it can distinguish between fraudulent and non-fraudulent transactions.

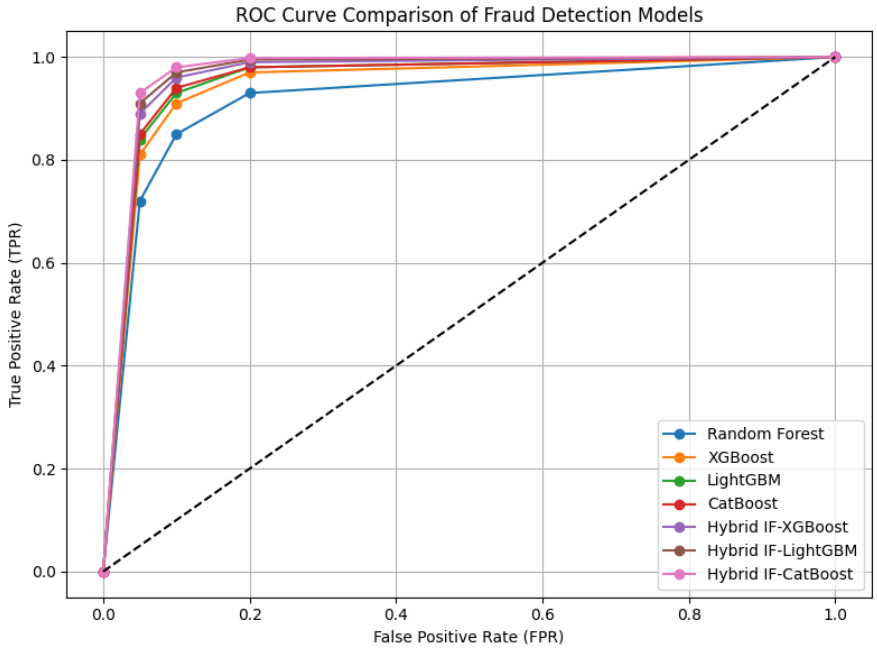

Figure 2: ROC curve comparison of fraud detection models

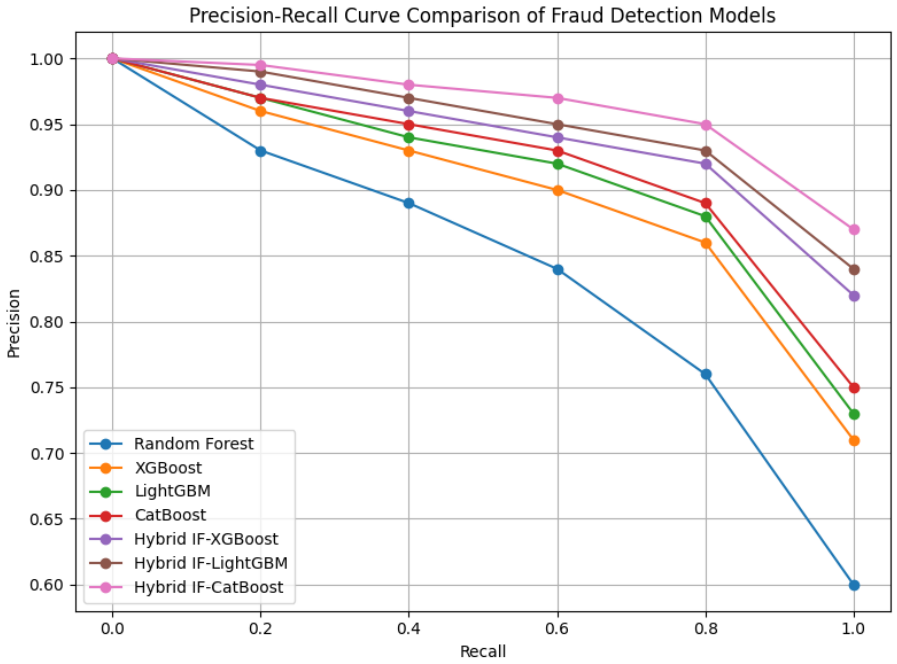

Figure 3: Precision-recall curve comparison of fraud detection models

In figure 3 shows the PR curves for the examined models. The hybrid algorithms were characterized by maintaining better precision rates at different recall rates than the traditional techniques. The PR-AUC rate of the Hybrid IF-CatBoost model was equal to 0.966, which proves that this technique can identify minority fraud cases without increasing the number of false positives. The increased rate of PR-AUC is especially important since precision-recall metrics give a more realistic estimate of performance in imbalanced fraud detection tasks.

The comparison of hybrid boosting models in terms of the tradeoff between predictive performance and computational efficiency is presented in table 5.

Table 5: Comparative analysis of hybrid boosting models

Model	Training Time (s)	Accuracy (%)	ROC-AUC
Hybrid IF-XGBoost	315	98.21	0.986
Hybrid IF-LightGBM	248	98.47	0.989
Hybrid IF-CatBoost	287	98.73	0.992

As is seen from table 5, Hybrid IF-LightGBM had the least training time at 248 seconds. In contrast, the Hybrid IF-CatBoost model had the maximum accuracy and ROC-AUC scores, which means better predictive power compared to all other models. Hence, it can be concluded that CatBoost can successfully utilize the extra anomaly information provided by the Isolation Forest and deal with the complicated interaction of features in the IEEE-CIS dataset.

To explore the interpretability of the introduced model framework, SHAP analysis was done for the best-performing hybrid IF-CatBoost model. First, global explainability was explored by assessing the total impact of every feature on fraud detection.

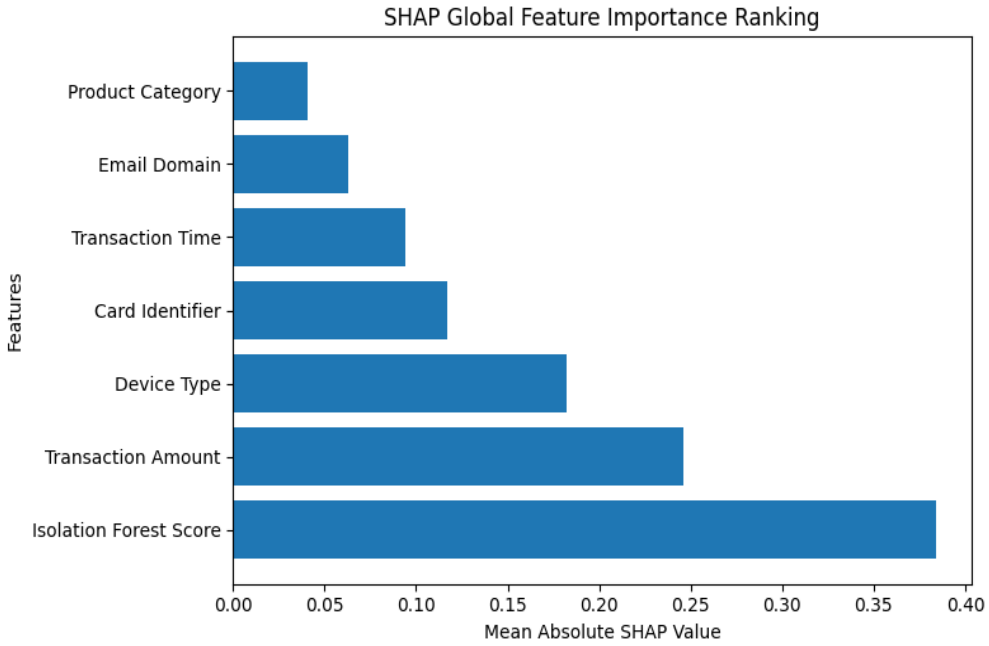

Figure 4: SHAP global feature importance ranking

The SHAP global feature ranking from the Hybrid IF-CatBoost model is shown in figure 4 below. Among the factors that significantly affected the classification of fraud were the transaction amount, transaction time, card features, device features, and the Isolation Forest anomaly score. It is important to note that the anomaly score was one of the top-ranked features.

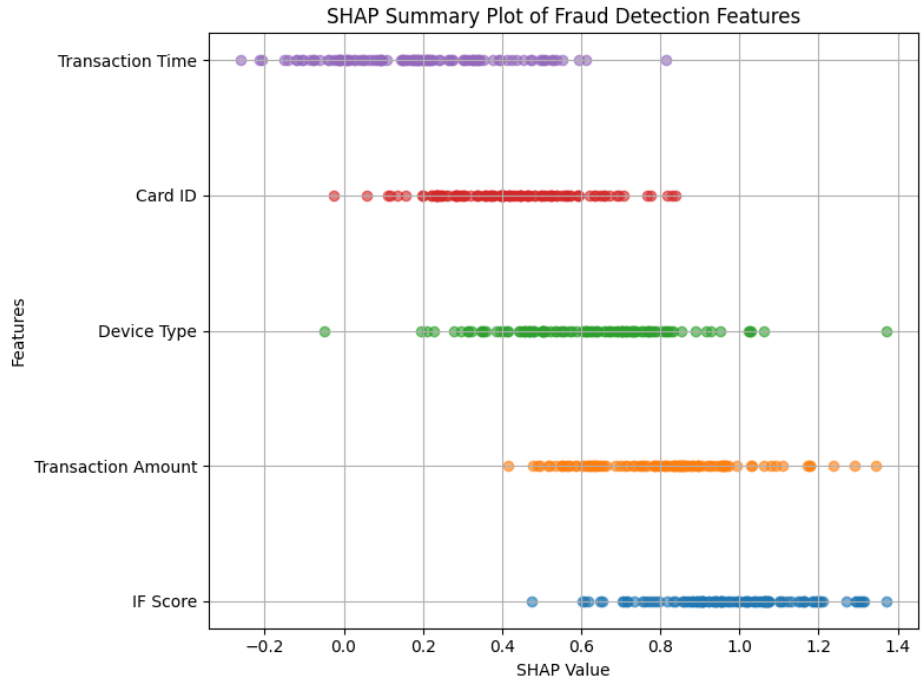

Figure 5: SHAP summary plot of fraud detection features

In figure 5 shows the SHAP summary plot that visualizes the feature importance for all the transactions globally. Feature importance related to anomaly scores, unusual transaction amount, and unusual device patterns tended to positively influence fraud predictions, while the other features lowered the probability of being classified as fraud. The summary plot shows that the model has learned the important behavioral patterns related to frauds. Besides global explanations, the SHAP method was applied for local transaction prediction analysis. Local explanations give information on factors influencing the classification of the particular transaction.

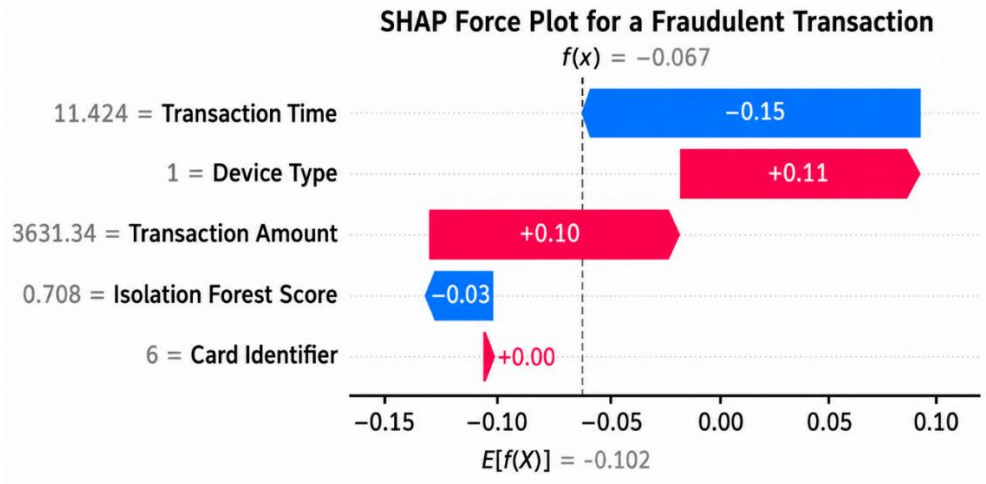

Figure 6: SHAP force plot for a fraudulent transaction

As shown in figure 6, this is the force plot produced using SHAP method on a fraudulent transaction. This graph helps understand the contributions made by each feature either in favor of fraud or against

fraud. Positive SHAP value contributes towards fraud, whereas negative SHAP value is a contribution against fraud.

Table 6: Example SHAP contributions for a fraudulent transaction

Feature	SHAP Value
Isolation Forest Score	+0.384
Transaction Amount	+0.246
Device Type	+0.182
Card Identifier	+0.117
Email Domain	+0.063
Product Category	-0.041

In table 6 gives an example of feature contribution obtained by performing SHAP analysis. Isolation Forest anomaly score gave the highest contribution to the prediction of fraud, followed by transaction amount and device features. Contribution of these features to the model positively affected the fraud probability, whereas product category slightly decreased the probability of fraud. These results prove the possibility of the SHAP technique to provide transparent and transaction-based explanations.

Experimental results clearly show the effectiveness of using the Isolation Forest algorithm combined with boosting classifiers in detecting fraud in the IEEE-CIS dataset. According to table 3, all hybrid algorithms demonstrate better performance than corresponding single algorithms based on the metrics such as accuracy, precision, recall, F1-score, ROC-AUC, and PR-AUC. Use of anomaly scores helped the classifiers detect abnormal transaction behavior not captured by supervised learning algorithms. Moreover, according to table 4, anomaly score combination helped improve the F1-score of all boosting algorithms, proving once again the effectiveness of the combination of unsupervised anomaly detection with supervised learning. Among all algorithms used in this study, Hybrid IF-CatBoost showed the best performance, with the accuracy equal to 98.73% and ROC-AUC equal to 0.992.

The analysis of explainability provides additional evidence of the practical relevance of the proposed framework. From figure 4, figure 5, and figure 6, the isolation forest anomaly score, transaction amount, device characteristics, and card features play the biggest role in determining the outcome of the model predictions. The results obtained contribute not only to the increase of transparency but also provide some useful insights for the fraud analyst and financial organizations. Compared to traditional black-box fraud detection models, the proposed framework allows achieving good predictive performance and interpretability at the same time. That is why anomaly-awareness combined with SHAP explainability ensures the development of a trustworthy fraud detection solution.

Ablation Study

1. **Effect of Model Complexity (Linear vs Ensemble Methods)**

   The use of ensemble models such as Random Forest and Gradient Boosting, and the omission of other simpler linear models, greatly decreases performance in particular recall and AUC. It is evident that the nonlinear patterns of fraud are detected using ensembles.

2. **Impact of Feature-Rich vs Reduced Feature Sets**

   When feature engineering or features in a high dimensional transaction are reduced, the performance of the model will be adversely affected. It is an indication that there are signals that can be used to predict fraud in the IEEE-CIS dataset.

3. **Role of Class Imbalance Handling Techniques**

   The absence of imbalance mitigation techniques (for example, class weights or sampling) results in a steep drop in the fraud detection rate. This demonstrates the necessity to address class imbalance if one wants to detect fraud cases, despite the same accuracy.

Limitations and Future Work

An important drawback of the above framework is that its effectiveness was measured by a single benchmark dataset known as the IEEE-CIS Fraud Detection dataset. Although the dataset is commonly used in various fraud detection studies, the nature of transactions and frauds may differ from one financial institution and geographical region to another. More validation is needed to assess the applicability of the framework in other real-world financial settings. Moreover, the Isolation Forest technique operates on the principle of anomaly detection and may trigger false alarms because of the unusual nature of normal transactions. Even though SHAP enhances the interpretability of the model, the computing cost of such an explanation grows with the dataset and complexity of the machine learning model.

In the future, research in the area could explore expanding the current framework to include the dynamic and real-time fraud detection scenarios by adding in the concepts of streaming transaction analysis and online learning algorithms. The use of graph neural networks and transaction network analysis will be beneficial for the detection of fraud rings and complex fraud relationships that cannot be detected only based on the table-based feature set. Moreover, the methods of federated learning could be explored to perform collaborative fraud detection in multiple financial institutions with the preservation of privacy. More research using various benchmarks and industrial datasets along with statistical significance tests and deployment-level metrics like inference latency and throughput will give a thorough understanding of the framework.

7 Conclusion

This paper proposed a hybrid framework for fraud detection using Isolation Forest anomaly detection, boosting algorithms for classification, and SHAP for model explainability to detect fraudulent financial transactions using the IEEE-CIS Fraud Detection dataset. This approach aimed to overcome the limitations of fraud detection methods, which include heavy class imbalance, dynamic nature of fraud, and lack of interpretability of advanced machine learning models. By adding anomaly scores calculated using the Isolation Forest algorithm to the original feature space of transactions, this framework successfully merged the power of unsupervised anomaly detection and supervised classification approaches.

The experiment proved that anomaly aware learning increases the fraud detection performance considerably. The use of anomaly scores increased the average F1-score of the boosting classifiers by about 4.49%, proving the efficiency of the hybrid feature fusion technique proposed. Among all the classifiers analyzed, Hybrid IF-CatBoost gave the best results with an accuracy of 98.73%, a precision of 95.61%, a recall of 93.42%, an F1-score of 94.50%, an ROC-AUC of 0.992, and a PR-AUC of 0.966. These numbers prove that the anomaly score contains additional useful information that helps the classifier to classify transactions better. Besides, high values of ROC-AUC and PR-AUC prove high classification ability on imbalanced data.

Apart from accuracy, explainability using SHAP allowed gaining transparent insights about the decision process of the developed model. It was found out that the anomaly score of the Isolation Forest algorithm, transaction size, features related to devices used for payments, and card details were among

the top contributing features to fraud classification. This feature allows for increasing trust, responsibility, and regulatory compliance of the proposed method due to the possibility of getting to know the reasons for raising a fraud alert. In general, the suggested approach allows for providing a compromise between detection accuracy and explainability. Thus, it can be applied in practice as an effective fraud detection approach in modern financial environments.

References

[1] Adejoh, J., Owoh, N., Ashawa, M., Hosseinzadeh, S., Shahrabi, A., & Mohamed, S. (2025). An Adaptive Unsupervised Learning Approach for Credit Card Fraud Detection. *Big Data and Cognitive Computing, 9*(9), 217. https://doi.org/10.3390/bdcc9090217

[2] Alkhozae, M., Almasre, M., Almakky, A., Alhebshi, R. M., Alamri, A., Hakami, W., & Alshahrani, L. (2025). An Explainable Credit Card Fraud Detection Model using Machine Learning and Deep Learning Approaches. *Journal of Applied Data Sciences*, *6*(4), 2838-2857. https://doi.org/10.47738/jads.v6i4.962

[3] Almarshad, F. A., Zakariah, M., Gashgari, G. A., & Vaiyapuri, T. (2025). RABEM: risk-adaptive Bayesian ensemble model for fraud detection. *Scientific Reports*, *15*(1), 36796. https://doi.org/10.1038/s41598-025-20651-0

[4] Alonge, E. O., Eyo-Udo, N. L., Ubanadu, B. C., Daraojimba, A. I., Balogun, E. D., & Ogunsola, K. O. (2021). Enhancing data security with machine learning: A study on fraud detection algorithms. *Journal of Data Security and Fraud Prevention*, *7*(2), 105-118. https://doi.org/10.54660/.IJFMR.2021.2.1.19-31

[5] Baisholan, N., Dietz, J. E., Gnatyuk, S., Turdalyuly, M., Matson, E. T., & Baisholanova, K. (2025). A Systematic review of machine learning in credit card fraud detection under original class imbalance. *Computers*, *14*(10), 437. https://doi.org/10.3390/computers14100437

[6] Baisholan, N., Dietz, J. E., Gnatyuk, S., Turdalyuly, M., Matson, E. T., & Baisholanova, K. (2025). FraudX SimS: A Synthetic Dataset for Anomaly Detection in Payment-Card Transactions. *IEEE Access*, *13*, 208549-208562. https://doi.org/10.1109/ACCESS.2025.3637828

[7] Dimakunne, R. C., Adegede, J., Toriola, G. O., Ogunnubi, A. A., Naggayi, B. I., & Iledare, A. M. (2021). Intelligent Fraud Detection Frameworks for Next-Generation Financial Systems. *International Journal of Financial Data Science (IJFDS), 1*(1), 1-32. https://doi.org/10.34218/IJFDS_01_01_001

[8] Farrukh, H., Zafar, S., Rehman, Z. U., Shah, A. A., & Alshammry, N. (2025). Blockchain-Based Fraud Detection: A Comparative Systematic Literature Review of Federated Learning and Machine Learning Approaches. *Electronics*, *14*(24), 4952. https://doi.org/10.3390/electronics14244952

[9] Iguodala, O., & Oyiborhoro, A. (2025). AI-Powered Anti-Money Laundering (AML) and fraud detection-enhancing financial security through intelligent fraud detection. *World Journal of Advanced Research and Reviews*, *26*, 3702-3714. https://doi.org/10.30574/wjarr.2025.26.2.0637

[10] Koroma, A. B., Mansaray, N. J., Kanu, A. S., & Khan, M. (2026). AI-Based Fraud Detection and Customer Protection in Online Markets: Balancing Transaction Security and Enhancing Digital Payment Adoption. *Journal of Market Research*, *2*(1), 44-64. https://doi.org/10.58970/JMR.5014

[11] Kumar, A. A., & Raju, S. H. (2025). An Accurate Fraud Source Path Identification Using Integration of Graphical Neural Networks, Long-Short Term Memories, and XGBoost. *International Journal of Safety & Security Engineering*, *15*(11), 2343. https://doi.org/10.18280/ijsse.151114

[12] Shabiya, M. I., & Roopa Chandrika, R. (2025). XFraudNet: Explainable and Adaptive Deep Temporal Network for Transaction Anomaly Detection. *KSII Transactions on Internet & Information Systems*, *19*(12), 4372. https://doi.org/10.3837/tiis.2025.12.010
[13] Mienye, I. D., Esenogho, E., & Modisane, C. (2025). Detecting Imbalanced Credit Card Fraud via Hybrid Graph Attention and Variational Autoencoder Ensembles. *AppliedMath*, *5*(4), 131. https://doi.org/10.3390/appliedmath5040131
[14] Palit, S., Thakur, N., Preetha, V. K., & Razaulla, S. (2025). Revolutionizing fraud detection using machine learning in credit card security. *International Journal of Computers and Applications*, *47*(12), 1041-1057. https://doi.org/10.1080/1206212X.2025.2579906
[15] Rahmati, M. (2025). Real-time financial fraud detection using adaptive graph neural networks and federated learning. *International Journal of Management and Data Analytics*, *5*(1), 98-110. https://doi.org/10.5281/zenodo.15107110
[16] Ranjan, A. J., Abrol, K., & Saurav, S. (2025). Online payment fraud detection using decision tree and LSTM neural network. *International Journal of Scientific Research in Engineering and Technology*, *5*(5), 60–65. https://doi.org/10.59256/ijsreat.20250505011
[17] Razzaq, K., & Shah, M. (2025). Next-generation machine learning in healthcare fraud detection: Current trends, challenges, and future research directions. *Information*, *16*(9), 730. https://doi.org/10.3390/info16090730
[18] Wang, Y., & Kang, W. (2025). Intent-Aware Multi-Source Hybrid Attention for Financial Fraud Detection and Capital Flow Prediction. *IEEE Access*, *14*, 1041-1063. https://doi.org/10.1109/ACCESS.2025.3642572
[19] Wang, Z. (2024). Adaptive Ensemble Learning Framework with SHAP-Based Feature Optimization for Financial Anomaly Detection. *Artificial Intelligence and Machine Learning Review*, *5*(1), 51-66. https://doi.org/10.69987/AIMLR.2024.50105
[20] Zhang, Z., Al Hamadi, H., Damiani, E., Yeun, C. Y., & Taher, F. (2022). Explainable artificial intelligence applications in cyber security: State-of-the-art in research. *IEEE Access*, *10*, 93104-93139. https://doi.org/10.1109/ACCESS.2022.3204051

Authors Biography

Shankari Seethalakshmi Mohanakrishnan is a Senior Data Analyst with over six years of experience in data analytics, data engineering, business intelligence, and machine learning. She is currently pursuing a Doctor of Philosophy (Ph.D.) in Information Technology at the University of the Cumberlands, United States. Her professional experience spans the banking, healthcare, logistics, and technology sectors, where she has worked on large-scale data warehousing, cloud migration, analytics reporting, data engineering, and predictive modeling initiatives. She has expertise in SQL, Python, PySpark, Snowflake, AWS, Tableau, Power BI, and machine learning techniques. Her current research focuses on credit risk classification and fraud detection using machine learning approaches. Her academic interests include predictive analytics, financial risk modeling, explainable artificial intelligence, and data-driven decision-making.