Journal of Information and Communications Technology: Algorithms, Systems and Applications

Research Article | *Open Access* |

Credit Card Fraud Detection Using Hybrid XGBoost and Autoencoder Models

Tejas V. Gandhi,[1] Pragati. P. Gupta,[1,*] Chirag. S. Gandhi,[1] Sarvesh. D. Gagare,[1] Vaishali Rajput[1] and Rohini Chavan[2]

[1] *Department of Artificial Intelligence and Data Science, Vishwakarma Institute of Technology, Pune, 411037, Maharashtra, India*
[2] *Department of Electronics and Telecommunication Engineering, Vishwakarma Institute of Technology, Pune, 411037, Maharashtra, India*
**Email:* pragati.gupta24@vit.edu (P. P. Gupta)

Abstract

The rapid growth of digital payment systems and e-commerce platforms has significantly increased the volume of credit card transactions worldwide, consequently raising the risk of fraudulent financial activities. Detecting fraudulent transactions remains a challenging problem due to the highly imbalanced nature of transaction datasets and the constantly evolving behavior of fraud patterns. Traditional rule-based detection systems and single machine learning models often struggle to identify both known and previously unseen fraud patterns effectively. To address these challenges, this study proposes a hybrid credit card fraud detection framework that integrates supervised and unsupervised machine learning approaches. In the proposed framework, XGBoost is employed as the primary supervised learning model to identify known fraud patterns from labeled transaction data due to its strong performance in handling nonlinear feature interactions and class imbalance. To complement this approach, an autoencoder-based anomaly detection model is used to identify unusual transaction behaviors by analyzing reconstruction errors from normal transaction patterns. The outputs of these two models are combined using a logistic regression–based meta-classifier, which learns an adaptive fusion strategy for integrating supervised fraud probabilities and unsupervised anomaly scores. The proposed hybrid system was evaluated using the BankSim transaction dataset, which simulates realistic financial transaction behavior. Experimental results demonstrate that the standalone XGBoost model achieved an Average Precision (AP) of 0.79, an F1-score of 0.719, precision of 0.743, and recall of 0.697. The final hybrid meta-classifier model achieved an AP of 0.724, F1-score of 0.703, precision of 0.689, and recall of 0.718, indicating improved recall stability and robustness in detecting fraudulent transactions. These results highlight the potential of hybrid machine learning architectures for enhancing fraud detection reliability in financial systems.

***Keywords*:** Credit card fraud detection; Hybrid machine learning; XGBoost; Autoencoder; Anomaly detection; Class imbalance.

Received: 15 January 2026; Revised: 12 February 2026; Accepted: 11 March 2026; Published Online: 16 March 2026.

1. Introduction

The rapid expansion of digital payment platforms and online financial services has significantly increased the number of credit card transactions worldwide. While this growth has improved convenience for consumers and businesses, it has also created new opportunities for financial fraud. Credit card fraud leads to substantial financial losses for banks, merchants, and customers each year, making it one of the most critical challenges in modern financial security systems. According to recent studies, fraudulent activities in digital payment systems continue to grow due to increasing transaction volumes and the widespread adoption of online

DOI: https://doi.org/10.64189/ict.25303

payment infrastructures.[1,2] Therefore, developing reliable and intelligent fraud detection systems has become an essential requirement for financial institutions.

Traditional fraud detection systems were primarily based on rule-based mechanisms and manual monitoring processes. These systems typically relied on predefined rules such as transaction thresholds, location mismatches, or unusual spending patterns. Although rule-based systems are simple to implement, they often struggle to detect complex or previously unseen fraud patterns. Fraudsters constantly adapt their strategies, making static rule-based approaches ineffective for modern financial environments.[3] Additionally, these systems generate a high number of false alarms and require continuous manual updates to maintain effectiveness.

To address these limitations, machine learning techniques have been widely adopted for credit card fraud detection. Machine learning models can learn patterns from historical transaction data and automatically identify suspicious activities. Various algorithms such as logistic regression, decision trees, random forests, and gradient boosting methods have been successfully applied to fraud detection problems.[4,5] Among these approaches, ensemble learning models such as XGBoost have shown strong performance due to their ability to capture complex nonlinear relationships within large-scale transaction datasets.

Despite the success of supervised learning techniques, these methods rely heavily on labeled data and may struggle to detect new or evolving fraud patterns that were not present during training. To overcome this limitation, unsupervised anomaly detection techniques have also been explored in fraud detection research. Methods such as autoencoders, isolation forests, and clustering-based approaches attempt to model normal transaction behavior and detect deviations from it as potential fraud.[6,7] Autoencoders, in particular, have been widely used for anomaly detection because they learn compact representations of normal data and identify anomalies through reconstruction error.

Several studies have attempted to combine supervised and unsupervised learning methods to improve fraud detection performance. Hybrid frameworks that integrate classification models with anomaly detection mechanisms have been proposed to leverage the strengths of both approaches.[8] Meta-learning techniques have also been explored, where multiple models are combined using a higher-level classifier that learns optimal decision boundaries from the outputs of individual models.[9,10] Such hybrid systems aim to improve robustness by detecting both known fraud patterns and previously unseen anomalies.

Although previous studies have demonstrated promising results using ensemble and hybrid learning strategies, many existing approaches either focus on a single learning paradigm or rely on fixed-weight combinations of model outputs. These strategies may not effectively adapt to different transaction behaviors or varying fraud patterns across datasets. Furthermore, the integration of supervised and unsupervised models in a unified decision framework remains an area that requires further exploration.

To address these challenges, this study proposes a hybrid credit card fraud detection framework that integrates supervised classification and unsupervised anomaly detection using a meta-classifier fusion strategy. In the proposed system, XGBoost is used as the primary supervised learning model to detect known fraud patterns from labeled transaction data, while an autoencoder-based anomaly detection model is employed to identify unusual transaction behavior. The outputs of these models are combined using a logistic regression–based meta-classifier, which learns how to optimally fuse supervised fraud probabilities and unsupervised anomaly scores.

The main objective of this work is to design a practical and scalable fraud detection system that improves detection robustness while maintaining interpretability and computational efficiency. The proposed approach is evaluated using the BankSim transaction dataset, which simulates realistic financial transaction behavior. By integrating complementary machine learning paradigms within a unified framework, the study aims to provide an effective fraud detection solution capable of supporting modern financial monitoring systems.

2. Methodology

This section provides a detailed explanation of the dataset used, the methodology for preprocessing, the supervised and unsupervised learning models, the hybrid fusion strategy, the meta-classifier design, the evaluation metrics, and the implementation of the system. The proposed methodology is designed to address class imbalance and evolving fraud patterns, following the challenges identified in prior studies.

2.1 Dataset description

For the experimental evaluation of the proposed fraud detection model, the chosen simulation dataset is the BankSim dataset. BankSim is a synthetic dataset, and synthetic datasets such as this one have been widely used in the field of fraud detection, since real banking data sources are inaccessible in some regions because of privacy and legal issues related to the confidentiality of customer information, as stated in [11].

The data are in the form of many transaction records made during consecutive time steps. In each transaction, the parameters include the transaction time or step, transaction amount, customer age group, gender, transaction type, and transaction class or label indicating fraud or a genuine transaction. The transaction types cover several areas such as food, transportation services, travel services, technology transactions, and lifestyle transactions. This allows research to conduct fraud analysis that depends on transaction type, as in previous works on anomaly detection.

Although BankSim is a synthetic dataset, it has been

widely adopted in fraud detection research because of its ability to realistically simulate customer behavior, transaction dynamics, and fraud patterns under controlled conditions. The use of BankSim enables reproducible experimentation while avoiding privacy and legal constraints associated with real financial data.

2.2 Data preprocessing

Data preprocessing is considered one of the most critical phases in fraud detection systems; the raw data of a transaction often contain irrelevant attributes, categorical variables, and features that are heterogeneously scaled. Poor preprocessing can severely degrade model performance and generalizability, as has been reported in several studies.[12]

2.2.1 Feature selection and cleaning

Initially, identifier-based features such as customer ID, merchant ID, and geographic code were filtered out. These features act more like unique identifiers rather than behavioral indicators and tend to be risky for model overfitting if used in training. A similar strategy for feature filtering has been followed in previous fraud detection frameworks to eliminate noise and help in generalizing better performance for unseen data.[13]

2.2.2 Categorical feature encoding

One-hot encoding was applied to transform categorical attributes, including customer age group, gender, and transaction category. In this encoding, each categorical value is composed of a binary feature so that no ordinal relationship unintended between categories is imposed by the machine learning models. One-hot encoding is widely applied in fraud detection pipelines because it is simple yet effective in preserving the categorical information of data examples.[14]

2.2.3 Feature scaling

The number of transactions has a wide numerical range and was therefore normalized to have zero mean and unit variance using standard scaling. Feature scaling is essential in gradient-based models and neural networks because unscaled features may dominate the learning process and negatively affect model convergence and learning stability. Normalization techniques are commonly used in machine learning–based anomaly and intrusion detection systems to improve classification effectiveness and ensure balanced feature contribution during model training.[15]

2.2.4 Train–validation–test split

The dataset is split into training, validation, and test datasets with stratified sampling to maintain the fraud-to-nonfraud ratio of the original dataset across all the splits. According to [16] and [17], stratification is important for imbalanced datasets to avoid biased validation results.

2.2.5 Handling class imbalance

To counter the class imbalance in supervised learning the synthetic minority over-sampling technique (SMOTE) was used solely for the training data. SMOTE helps in creating artificial samples for the minority class and thus enables the classifier to better detect the fraud class. Note, however, that oversampling is not performed for validation and test data to avoid inflating performance measures, and techniques to be used are as stated in [6,8].

The full preprocessing pipeline was serialized and stored so that it would work identically in both the training and deployment environments to avoid the problems with data leakage and preprocessing differences mentioned in other studies on model deployment.[6]

2.3 Supervised learning models

The major mechanism for known fraud pattern detection is supervised learning models, which are based on labeled transaction data. Several supervised classifiers were assessed to select the best model that can be integrated into the hybrid framework.

2.3.1 Logistic regression

Owing to its interpretability and computational efficiency, logistic regression was used as a baseline among the tested supervised classifiers. It models fraud probability by adopting a linear decision boundary with a sigmoid activation. Logistic regression has been widely used in early fraud detection systems, but its application essentially faces serious limitations because it can hardly capture any complex nonlinear relationships that might appear in transaction data. The probability of fraud can be expressed as

$$\mathrm{P(y = 1|x)} = \frac{1}{1+\mathrm{e}^{-(\mathrm{w^T\,x+b})}} \tag{1}$$

where;

x represents the input feature vector consisting of the supervised fraud probability and the anomaly score, w represents the learned weight vector, and b is the bias term. The output of the logistic regression model corresponds to the final fraud probability produced by the hybrid system.

During training, the parameters are optimized by minimizing the logistic loss function given by

$$L = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(p_i) + (1 - y_i)\log(1 - p_i)] \tag{2}$$

where;

y_i represents the true class label and p_i represents the predicted probability of fraud.

2.3.2 Random forest classifier

Random forest is an ensemble learning algorithm that makes use of the combined outputs of a series of decision trees that have been trained on random subsets of features. This helps overcome the weaknesses of linear models. Various previous studies have proven the efficacy of the random forest model in the area of fraud detection even in comparison to the linear

model.[7,13] The disadvantage of the random forest model is that it becomes complex in larger datasets.

2.3.3 XGBoost classifier

Among the supervised models that were considered for evaluation, XGBoost (Extreme Gradient Boosting), which demonstrated excellent performance, was chosen as the chief supervised model. XGBoost is an algorithm that builds an ensemble of weak models sequentially by maximizing and regularizing an objective function. Its efficacy in addressing missing values, handling complex interactions in the feature space, and controlling overfitting via regularization has been recognized the literature on fraud detection tasks.[2,4,12]

Experimental results have proven that XGBoost outperforms logistic regression and random forest in terms of precision, recall, F1-measure, and average precision. The general objective function used in XGBoost can be expressed as

$$\mathrm{Obj} = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \quad (3)$$

where,

$l(y_i, \hat{y}_i)$ represents the loss function measuring the difference between the predicted value $\hat{y}_i$ and the true label y_i and $\Omega(f_k)$ represents the regularization term that penalizes model complexity. The regularization term is defined as

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \quad (4)$$

where,

T is the number of leaves in the decision tree, wj represents the weight of leaf j, and γ and λ are regularization parameters used to prevent overfitting.

2.4 Unsupervised learning model

Although it has proven to be quite successful in finding known fraud patterns, supervised models are limited in that they rely on labeled data. To improve the detection capabilities of new and developing fraud patterns, an unsupervised anomaly detection model was added.

2.4.1 Autoencoder-based anomaly detection

For the unsupervised part of the learning process, a deep autoencoder was used. An autoencoder is a neural network that is used to learn a representation of the input data that has to be reconstructed through self-learning by reducing the error in reconstruction. When trained for normal transactions alone, autoencoders can effectively learn normal behavior in the pattern of legitimate transactions. Those that are not anomalies are deemed to have larger errors in reconstruction.[14,17]

For this purpose, the AE architecture had more than one hidden layer for learning nonlinear relationships among the features. The features were scaled before training the AE. The reconstruction error which was calculated as the mean squared error between the original and reconstructed outputs was used as the AE score to measure the anomaly of each transaction.

It has been observed in the literature that the use of autoencoders is effective at uncovering novel patterns of fraudulent activity, but in a standalone manner, they may produce high false positive rates.[1,7] The drawback of the method sparks interest in using them in a hybrid model.

The autoencoder is not expected to be a substitute classifier by itself for fraud detection, but rather a supporting source of anomalies to be used jointly with the results of supervised classification. The reconstruction process of the autoencoder can be expressed as

$$\hat{x} = g(f(x)) \quad (5)$$

where,

(x) represents the encoder function that maps the input data to a lower-dimensional latent representation and $(\cdot)$ represents the decoder function that reconstructs the input from the latent space.

The reconstruction error used for anomaly detection is computed using the mean squared error (MSE) loss function

$$L_{AE} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \hat{x}_i)^2 \quad (6)$$

where,

xi represents the original transaction features and $\hat{x}_i$ represents the reconstructed output generated by the autoencoder. Transactions with higher reconstruction errors are considered anomalous and may indicate potential fraud.

2.5 Hybrid model architecture

The proposed hybrid fraud detection system combines the outputs of the supervised XGBoost model and the unsupervised autoencoder anomaly score using a meta-classifier. Hybrid learning strategies that integrate supervised and unsupervised models have been shown to improve fraud detection performance by leveraging complementary strengths of different algorithms.[8,10] In such systems, supervised models capture known fraud patterns from labeled data, while unsupervised models identify anomalous transaction behaviors that may correspond to previously unseen fraud cases.[6,11]

Let P_{XGB} represent the fraud probability generated by the XGBoost classifier and S_{AE} represent the anomaly score obtained from the autoencoder reconstruction error. The input vector to the meta-classifier is defined as

$$z = [P_{XGB}, S_{AE}] \quad (7)$$

The final fraud probability predicted by the hybrid model is obtained using the logistic regression function

$$P_{Hybrid} = \sigma(w_1 P_{XGB} + w_2 S_{AE} + b) \quad (8)$$

where,

w_1 and w_2 represent the weights learned by the meta-classifier and $\sigma(\cdot)$ represents the sigmoid activation function. The logistic regression model learns these parameters during training by minimizing the classification loss over validation

data. This formulation enables the meta-classifier to automatically determine the relative importance of supervised fraud probabilities and unsupervised anomaly scores when making final predictions.

By learning this adaptive fusion strategy, the hybrid framework can dynamically adjust the contribution of each model depending on the transaction characteristics. Similar meta-learning–based ensemble approaches have been successfully applied in fraud detection systems to improve robustness and decision reliability.[9,10]

2.6 Meta classifier fusion strategy

To address the limitations of fixed fusion approaches, logistic regression meta-classifier-based adaptive fusion of model outputs was proposed. Substantial improvement in ensemble learning have been achieved by learning optimal fusion strategies directly from data rather than depending on rule-based specifications.[4,6] In the proposed model, the meta-classifier is applied to the outputs of both supervised and unsupervised learning components, namely, the fraud probability estimated using the XGBoost classifier and the anomaly scores computed from the reconstruction errors of the autoencoder. The meta-classifier then combines these two different outputs to yield the eventual fraud probability of each transaction. This learning model enables automatic validation data-driven computation of the weights of supervised and unsupervised data inputs to the model, thus making it irrelevant to include static fusion weights in supervised learning models. In addition, it should be noted that a validation data-driven decision threshold was considered to optimize the F1-score in this research. This was aimed at achieving a good balance between precision and recall in fraud probability estimation.[8,12] Logistic regression was selected for meta-classifier fusion because of its interpretability, probabilistic calibration, and stability, rather than as a novel modeling contribution.

2.7 System architecture of the proposed hybrid fraud detection framework

The proposed hybrid credit card fraud detection framework follows a multi-stage machine learning architecture designed to integrate supervised classification and unsupervised anomaly detection in a unified system. Hybrid architectures combining classification and anomaly detection models have been widely explored in security and anomaly detection systems to improve detection robustness and adaptability across evolving attack patterns.[18,19] The architecture begins with the input of raw transaction data containing various attributes such as transaction step, transaction amount, customer age group, gender, and transaction category. These features represent the behavioral and transactional characteristics that can indicate normal or fraudulent activity patterns.

Before model training, the transaction data pass through a preprocessing module that performs several essential transformations. Categorical attributes such as age group, gender, and transaction category are converted into numerical representations using one-hot encoding to make them suitable for machine learning models. In addition, numerical attributes such as transaction amount and transaction step are normalized through feature scaling to ensure that all features contribute equally during model learning. Similar preprocessing and feature engineering strategies have been widely applied in intelligent security systems and anomaly detection frameworks to improve classification accuracy and system stability.[19,20] The preprocessing stage produces a 26-dimensional feature vector that is used as the input for the learning models. Importantly, the same preprocessing pipeline is applied during both training and inference to maintain consistency and prevent data leakage.

Following preprocessing, the framework employs two parallel machine learning branches. The first branch consists of a supervised learning model implemented using the XGBoost classifier. This model is trained using labeled transaction data and learns complex nonlinear relationships between transaction features and fraud labels. The output of this branch is a probability score representing the likelihood that a given transaction is fraudulent.

The second branch consists of an unsupervised anomaly detection model based on an autoencoder neural network. The autoencoder learns the normal structure of legitimate transaction data through an encoder–decoder architecture that attempts to reconstruct the input features. During inference, the reconstruction error between the original transaction and the reconstructed output is computed. Transactions with higher reconstruction errors are interpreted as anomalies and are assigned higher anomaly scores.

Table 1: Model configuration details.

Component	Configuration detail
XGBoost Classifier	Binary logistic objective; trained using default hyperparameters of the XGBoost library; probability-based output used for fusion; class imbalance handled through data resampling
Autoencoder	Fully connected symmetric architecture; input dimension = 26 features; trained on normal transactions only; loss function = mean squared error; optimizer = Adam; reconstruction error used as anomaly score
Autoencoder Preprocessing	Feature scaling applied prior to training; reconstruction scores normalized using Min–Max scaling
Meta-Classifier	Logistic regression; input features = XGBoost fraud probability and autoencoder anomaly score; trained on validation data
Decision Threshold	Selected based on F1-score maximization on validation set

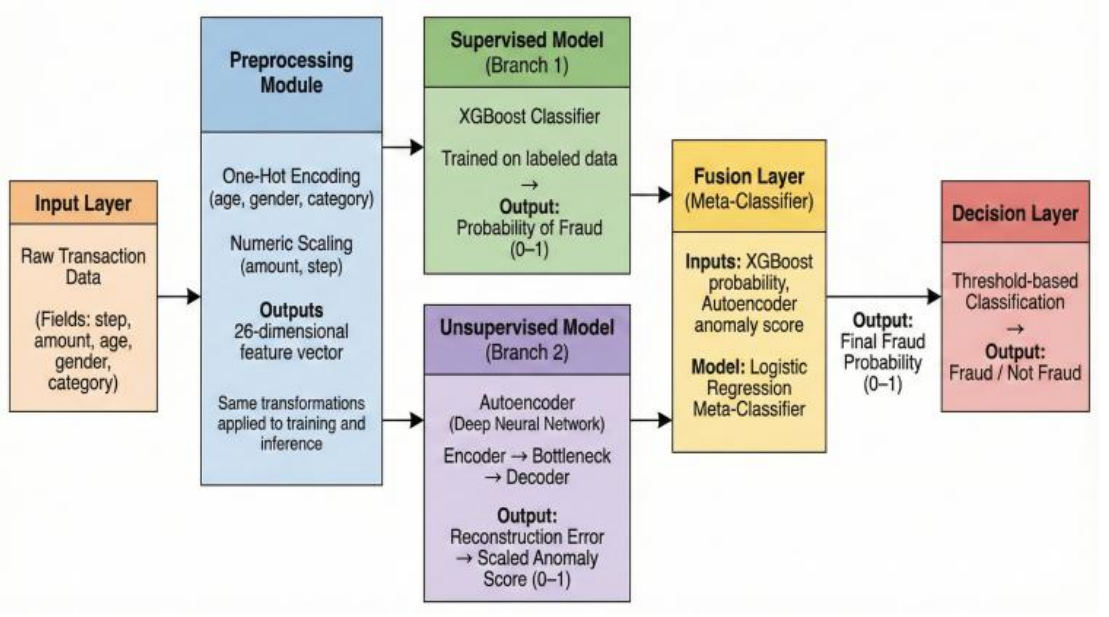

Fig. 1: Architecture of the proposed hybrid credit card fraud detection framework integrating supervised XGBoost classification, autoencoder based anomaly detection, and logistic regression meta-classifier fusion.

Similar anomaly detection approaches have been used in network security and traffic monitoring systems to identify unusual behavioral patterns within large-scale data streams.[21]

As illustrated in Fig. 1, the outputs generated by the supervised and unsupervised models are subsequently combined through a fusion layer implemented using a logistic regression meta-classifier. This meta-classifier receives two inputs: the fraud probability predicted by the XGBoost classifier and the anomaly score produced by the autoencoder. By learning from validation data, the logistic regression model determines how much importance should be assigned to each signal when making the final prediction. The meta-classifier generates a final fraud probability that represents the combined decision of the supervised and unsupervised components. This probability is then evaluated against a predefined decision threshold to classify the transaction as either fraudulent or legitimate. The hybrid architecture therefore leverages the strengths of both learning paradigms: the supervised model effectively detects known fraud patterns learned from labeled data, while the unsupervised autoencoder identifies unusual transaction behavior that may correspond to previously unseen fraud scenarios.

By combining these complementary approaches through meta-learning, the proposed architecture improves the robustness and adaptability of the fraud detection system. This design enables the framework to maintain strong detection performance while reducing the limitations associated with using a single learning method.

2.8 Evaluation metrics

In assessing the model performance, a number of metrics were incorporated into the test environment. The accuracy was calculated but not utilized to determine the performance because of the imbalance in classes. Precision is a performance evaluation metric that measures the correctness of positive predictions made by the model. Moreover, for a threshold-independent assessment, the receiver operating characteristic area under the curve and average precision were employed. The average precision metric is very much suitable for datasets that are imbalanced, as it provides overall precision and recall for all the thresholds.

2.9 System implementation and deployment

The proposed framework was implemented utilizing the following Python-based machine learning libraries: scikit-learn, XGBoost, and TensorFlow/Keras. All preprocessing steps, models trained, and fusion components were serialized for reproducibility and consistency in deployment.

First, to demonstrate practical applicability, the final hybrid model was deployed as a web-based application using Streamlit. That application supports both manual transaction inputs and batch CSV uploads, not perform preprocessing, model inference, hybrid fusion, and result visualization in real time. Deployment-focused design considerations are informed by best practices reported in recent applied fraud detection systems.[1,4]

3. Results and analysis

3.1 Evaluation strategy and metrics

This section describes an in-depth experimental assessment and investigation of the proposed hybrid credit card fraud detection system. A comprehensive analysis of traditional supervised, unsupervised, and hybrid models is conducted to highlight the effectiveness of the proposed meta-model, which integrates all individual models. Special attention will be given to the evaluation metrics for imbalanced datasets,

practicability, and improvement in performance compared with those of individual models.

3.1.1 Accuracy

The accuracy of a system is essentially the ratio of positively predicted results to the total number of observations performed. Equation (9) shows how accuracy is calculated using the following parameters, where the numerator accounts for all the predictions for which the model is correct and the denominator denotes all predictions that were made.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (9)$$

where;
TP = True positive
TN = True negative
FP = False positive
FN = False negative
Here, true positive refers to the case in which the object to be detected is actually fraud, and the system positively classifies it as fraud whereas, true negative is the case in which the class is not fraud, and the prototype accurately classifies it as not fraud.

It terms of false detections and scenarios we have parameters such as false positive and negative respectively. False Positive is the case when the model positively classified it to be a fraud but in actual it was not a fraud class and false negative is the case when the model classifies the object as not a fraud class; in practice, it belongs to the fraud class.

3.1.2 Precision

Precision in deep learning is a performance evaluation metric that essentially evaluates the quality and correctness of the accuracy parameter i.e., positive classifications by the model.

$$Precision = \frac{TP}{TP+FP} \quad (10)$$

Equation (10) shows how the precision of a model is calculated on the basis of true positives and false positives. Here we seek to determine the actual correctness of a model; hence, we consider only positive classification scenarios where the prediction is always correct. However, it also has a major drawback because it does not include the negatives at all, which might cause the model to miss certain correct predictions (i.e., low recall).

3.1.3 Recall

Within this performance evaluation parameter, we check in actuality how many cases did the model actually classified out of all the positive cases positively. It ranges from 0 to 1. This essentially measures the model's ability to capture all the relevant instances of the positive class.

$$Recall = \frac{TP}{TP+FN} \quad (11)$$

Equation (11) answers our question, "Out of all the actual fraud, how many did our model find?" If a model has higher recall, then we can safely say that the model is classifying most of the positive classes; hence, maximum fraud in the field of transaction is being successfully detected.

However, if the recall alone is too high, then the model is classifying every object as fraud, thus making the recall of the model 100% but reducing the precision in its classification, which accounts for a failure in the model's classification.

3.1.4 F1 score

This parameter is solely based on the values of precision and recall of the particular model, as it is a harmonic mean of the precision and recall of the model. It ranges between 0 (worst) and 1 (best). This metric is the one that gives us a trade-off between the precision and recall of a particular model. As the harmonic mean is observed to punish the extreme high resulting values more, this is preferred over the arithmetic mean process. As a result, both the precision and the recall must be above the mark to achieve a reasonably high F1 score.

$$F1\ score = 2 \times \frac{Precision \cdot Recall}{Precision+Recall} \quad (12)$$

Equation (12) shows how mathematically the F1 score is calculated using the precision and recall metric values. It is especially used in cases where a particular model has an imbalanced dataset or cases where the model needs to have a proper balance between precision and recall.

3.1.5 ROC–AUC (receiver operating characteristic – area under the curve)

A receiver operating characteristic (ROC) curve is a tool used in the assessment of the discriminatory powers possessed by a binary classification model at various levels. This graph helps in plotting the true positive rate against the false positive rate. The area under the ROC curve is a numerical value, that reflects the degree of separability between the classes.

$$TPR = \frac{TP}{TP+FN}$$

$$FPR = \frac{FP}{FP+TN}$$

$$ROC-AUC = \int_0^1 TPR(FPR)d(FPR) \quad (13)$$

Equation (13) above expresses the area under the ROC curve and always lies between 0 and 1. An ROC-AUC of 1 represents excellent discriminant power. On the other hand, 0.5 represents random classification. Turning to the problem of credit card fraud detection, the ROC-AUC is used to measure how well the model separates the fraudulent class from the legitimate transactions, regardless of the classification threshold. The higher the ROC-AUC score is, the more likely the model is to predict a higher probability of fraud than legitimate transactions; thus, this serves as a good

measure of imbalance in this problem.

3.1.6 Average precision

The average precision (AP) is a method for measuring the performance, it is obtained by using the precision-recall curve and is widely utilized for highly imbalanced classification tasks. The average precision is the weighted sum of the precision values obtained at different points on the recall curve.

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

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

$$AP = \sum_n (R_n\text{-}R_{n-1}).P_n \quad (14)$$

In equation (14), P_n and R_n, represent the precision and recall at the nth threshold, respectively. The average precision are between 0 and 1, and higher values represent the superior performance of the model. AP is particularly useful in fraud detection models since it highlights the detection of the minority class, which in this case represents fraud, and punishes false-positive predictions. When the AP is high, it implies that the model is achieving high precision with increasingly correct detection of fraudulent transactions, rendering it a very credible metric.

3.2 Results of the supervised learning models

Three supervised classifiers-logistic regression, random forest, and XGBoost-were trained and tested on the same pre-processed dataset.

3.2.1 Logistic regression

A baseline supervised logistic regression model was used because it is quite simple to understand. Although it converged well and performed satisfactorily, its linear decision boundary restrained it from handling nonlinear patterns in fraudulent activities. It therefore resulted in lower recall and F1 measure being achieved by it as opposed to methods that are based on ensembles. This is in line with previous studies on fraud detection in that linear methods perform poorly when faced with high-dimensional data as is inherent in transaction data.[6,7]

3.2.2 Random forest

The random forest classifier performed better than the logistic regression model in handling nonlinear relationships between features. The ensemble model helped to achieve lower variance and higher recall values; although, it had a slightly higher FP rate than the XGBoost models did. This pattern was expected on the basis of the results in [1] and [13], where RF models favored recall over precision in fraud detection problems with class imbalance issues.

3.2.3 XGBoost

Among these, the best performance was recorded by XGBoost for a classifier. The model achieved high precision with strong recall and an F1-score, reflecting that XGBoost learned the fraudulent transaction patterns correctly. With highly promising ROC-AUC and average precision values, XGBoost clearly maintained great discriminative power across a wide range of thresholds.

This superior performance of XGBoost was due to its powerful gradient boosting framework, regularization mechanisms, and ability to model complex feature interactions. In fact, the results presented here are in good agreement with recent related works that position XGBoost as a state-of-the-art approach for financial fraud detection.[2,4,12] Therefore, XGBoost was chosen as the supervised backbone of the hybrid framework. Confusion matrix can be given as follows.

The confusion matrix presented in Fig. 2 illustrates the classification performance of the XGBoost model on the test dataset by showing the distribution of true positives (TP), true negatives (TN), false positives (FP), and false negatives

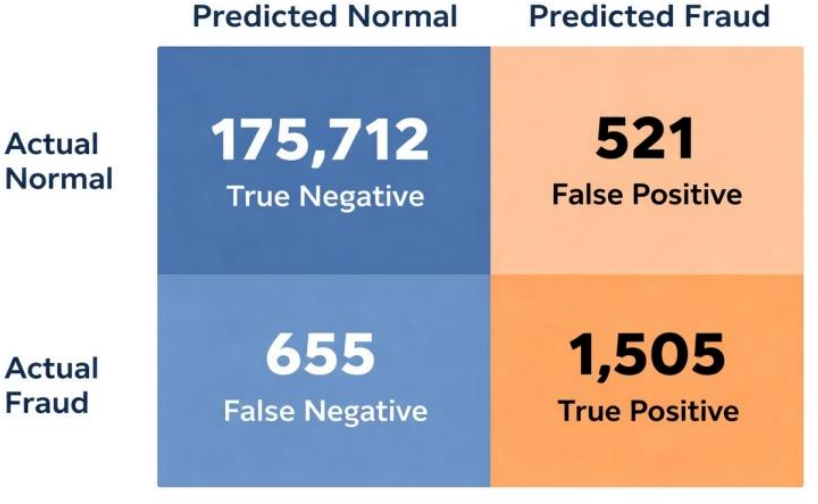

Fig. 2: Confusion matrix of XGBoost.

represent fraudulent transactions that the model fails to (FN). True positives correspond to fraudulent transactions that are correctly identified as fraud, while true negatives represent legitimate transactions correctly classified as nonfraud. False positives occur when legitimate transactions are incorrectly classified as fraud, and false negatives detect. The model correctly classified 175,712 normal transactions and 1,505 fraudulent transactions, while 521 normal transactions were incorrectly flagged as fraud and 655 fraud cases were missed. These results indicate that the XGBoost model achieves strong classification performance with high accuracy and precision in detecting fraudulent transactions.

Based on the values obtained from the confusion matrix, evaluation metrics can be calculated to assess the performance of the XGBoost model. Precision measures the proportion of predicted fraud transactions that are actually fraudulent and is computed using Equation (10). Recall measures the ability of the model to correctly identify fraudulent transactions among all actual fraud cases and is calculated using Equation (11). The F1-score combines precision and recall into a single metric to evaluate the balance between detection accuracy and completeness. Additionally, the receiver operating characteristic–area under the curve (ROC–AUC) and average precision (AP) metrics are used to evaluate the model's performance across different classification thresholds, which is particularly important for highly imbalanced datasets such as credit card transactions.

Using these metrics, the XGBoost classifier achieved a precision of 0.743, recall of 0.697, and an F1-score of 0.719 for the fraud class. In addition, the model obtained an Average Precision (AP) score of 0.79 based on probability outputs across classification thresholds, indicating strong performance in detecting fraudulent transactions under class imbalance conditions. These results demonstrate that XGBoost provides a strong baseline model for fraud detection and justify its selection as the primary supervised component of the proposed hybrid framework.

Table 2: Performance Metrics of XGBoost model.

Metric	Formula	Value
Precision	TP / (TP + FP)	0.743
Recall	TP / (TP + FN)	0.697
F1-score	2PR / (P + R)	0.719
Accuracy	(TP + TN) / Total	0.9934 (99.34%)

3.3 Results of unsupervised learning by autoencoder

The autoencoder was trained on mostly nonfraudulent examples to learn a reduced representation of normal behavior. The reconstruction error was used as an anomaly score (AE score), where higher values represent greater deviations from normal samples.

3.3.1 Standalone performance of the autoencoder

When tested alone, the autoencoder performed well in the anomaly detection task, especially when the anomalous transactions whose pattern deviated greatly from the usual spending habits were pointed out. The results revealed that the average reconstruction error of the anomalous transactions is significantly greater than that of the normal transactions, indicating that the autoencoder can learn the structure of the data.

However, when evaluated as an individual classifier, it performed well in terms of both high recall and low precision values. This indicates that the model successfully identified many fraudulent patterns; however, a considerable number of normal transactions were misclassified as fraudulent. Such behavior is consistent with findings in the anomaly detection literature, where anomaly detection systems tend to over-detect normal behaviors as anomalies.[7,14]

These findings indicate that the autoencoder model performs well in the task of detecting anomalies, but it is not sufficient for accurate fraud classification in a real-world environment.

3.4 Obtaining initial hybrid scores from a weighted fusion approach

An initial method of hybrid scoring that integrates elements from unsupervised and supervised models was assessed via weighted average combinations. By definition, the hybrid score is derived by linearly combining XGBoost's predicted probabilities with those of the autoencoder model's predicted anomalies.

3.4.1 Initial hybrid results using weighted fusion

The determination of weights was conducted through maximizing the average precision on the validation set(s), which had been created prior to the empirical studies. In this initial analysis the optimal weights are very heavily concentrated toward the supervised nature of the models being used, with the fusion weight assigned to the XGBoost model being in excess of approximately 0.95. This suggests that the performance of the supervised models provided a more reliable source of information than the unsupervised models within the evaluated dataset did.

The analysis also indicated a limited degree of improvement across the various metrics as a result of this hybridization approach compared to that of the standalone performance of the XGBoost model. In some instances, the hybrid XGBoost model performed poorly as a result of the low F1-score. Similar limitations exhibited by fixed-weights using hybrid fusion approaches have been reported in previous works.[8,13] The evolution of this static fusion approach is constrained by the limited flexibility of its approach to accommodate variations in transaction patterns.

Fig. 3 shows Receiver Operating Characteristic (ROC) and precision–recall (PR) curves comparing the performance of the supervised, unsupervised, and hybrid fraud detection models on the test dataset. The ROC curve illustrates class separability across thresholds, whereas the PR curve highlights performance under class imbalance.

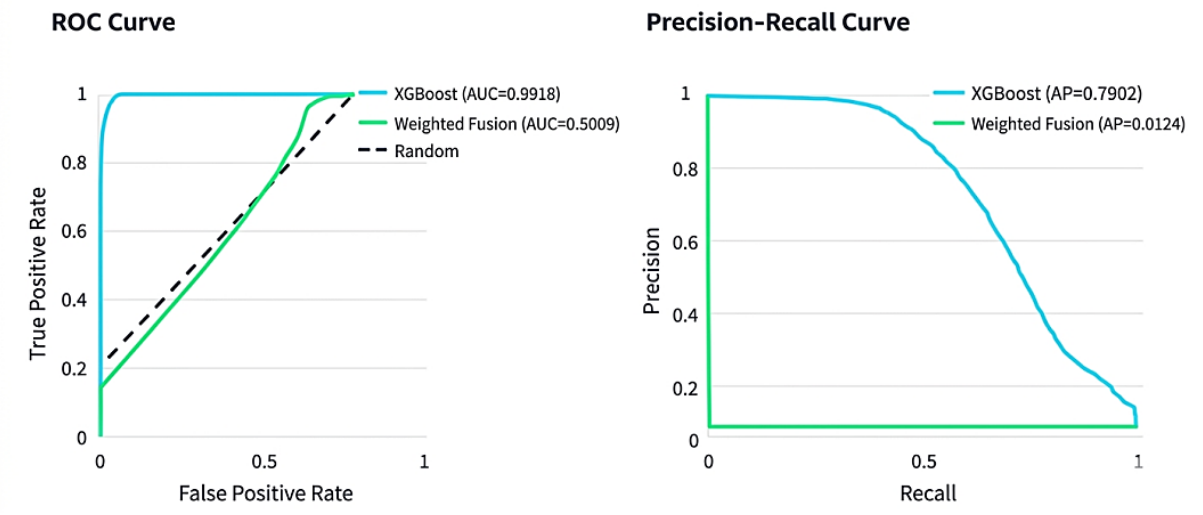

Fig. 3: Receiver Operating Characteristic (ROC) and precision-recall curves.

3.5 Corrected autoencoder fusion results

Suboptimal fusion was caused by incorrect scaling of the reconstruction scores from the autoencoders. Weighted fusion was then performed with the normalized reconstruction scores from the AEs.

3.5.1 Improved fusion after score calibration

With the proper scaling of the AE scores, the hybrid model was able to achieve better agreement between precision and recall. There was a large change in the optimal fusion weight, with smaller weights (0.05-0.30) given to the autoencoder component. This finding that, although the information in the anomalies is important, it has to play a supplemental, not primary, role.

These results are true to the observations in [14] and [17], where the emphasis is on calibrating the anomaly scores carefully prior to use in a hybrid fraud detection approach. The low fusion weights associated with the autoencoder demonstrate that its contribution is supplementary and not primary, which is confirmed by its inherent ability to produce false positives as a sole contributor.

3.6 Final hybrid results using metaclassifier fusion

To overcome the limitations of fixed-weight fusion methods, a meta-classifier based on logistic regression was proposed as the final hybrid scheme.

3.6.1 Meta classifier training

The meta-classifier was then trained on two input variables: the fraud probability derived from the XGBoost supervised learning classifier and the reconstruction anomaly score derived from the autoencoder. The target output for the meta-classifier corresponded to the binary fraud label. This was achieved on the validation set, allowing the meta-classifier to learn the best possible combination of both supervised and unsupervised information. This learning enabled the logistic regression meta-classifier to automatically weigh the importance of the input signals for an improved fraud detection accuracy on any transaction pattern without human intervention.

3.6.2 Test set performance

Compared with all the other approaches, the meta-classifier-based hybrid model resulted in the best overall performance. The meta-classifier-based hybrid approach enhanced the recall without leading to a significant increase in the number of false positives. Consequently, the F1-score improved.

The hybrid approach maintained a high value for the ROC-AUC, meaning a indicating strong separation between classes, and scored well on average precision, ensuring good robustness to strong class imbalance. These experiments show that meta-learning is a more flexible and robust fusion function than weighting based on a human judgment, as already suggested in [2,4] and [6]. While the single XGBoost model has better performance in terms of average precision, the hybrid meta-classifier is much more stable and robust for recall through thresholds. This trade-off is important in fraud detection systems where missing only a few cases of fraud is usually preferred over minor improvements in precision.

Fig. 4 illustrates the confusion matrix for our proposed hybrid fraud detection framework that employs a logistic regression meta-classifier. The hybrid system accurately detected 175,535 normal transactions as well as 1,550 fraudulent transactions. Meanwhile, 698 normal transactions were wrongly classified as fraud and 610 fraudulent ones were misclassified as normal. Overall, this shows that the hybrid framework enhances recall and the ability to detect fraud by integrating supervised and anomaly-based signals.

3.7 Comparative analysis

A clear advantage of the final hybrid approach is evident from a comparative analysis across all the models evaluated. While the performance of the supervised learning models is very high in identifying known fraud patterns learned from historical labeled data, there are significant limitations in detecting novel and/or unseen fraudulent behaviors. As

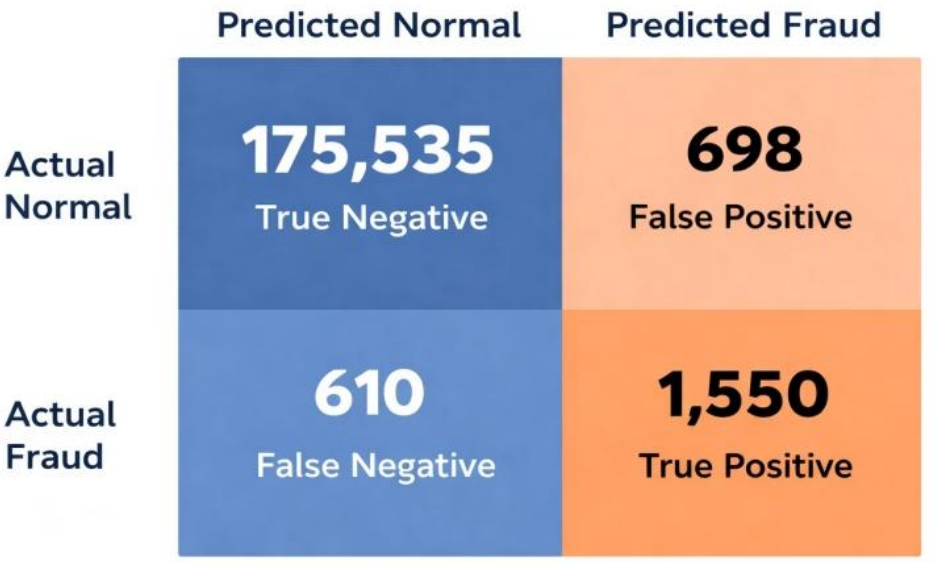

Fig. 4 Confusion matrix of Hybrid model (XGB+AE+Meta-classifier).

Table 3: Comparative performance of the proposed hybrid fraud detection framework

Method Name	AP	F1(Fraud Class)	Precision	Recall
XGBoost only	0.790	0.718	0.742	0.697
Hybrid 1: XGB+IF(weighted)	0.7908	0.718	0.742	0.697
Hybrid 2: XGB+AE(broken scaling)	0.012	0.026	0.01	0.97
Hybrid 3: XGB+AE(fixed scaling)	0.79(test)	0.331	0.209	0.793
Final Hybrid 4: XGB+AE+Meta-classifier	0.724 (AP)	0.7033	0.6895	0.718

expected, unsupervised models effectively identify anomalous transaction patterns but tend to yield a high number of false positives since they are sensitive to rare yet legitimate behaviours. Initial weighted hybrid fusion mainly provides marginal performance improvements and is sensitive to score calibration and the selection of weights for fusion. By learning how to trust appropriately and combine the outputs of the supervised and unsupervised models under varying transaction conditions, it achieved the most balanced and robust performance. This progressive evaluation justifies the design choices that were undertaken in the proposed framework and confirms that recent studies in the literature integrate the concept of complementary learning paradigms into fraud detection.[8,11,16] The results suggest that, overall, the hybrid methods do not outperform strong supervised methods across the board, but they provide improved recall and robustness, which is very important in high-risk fraud detection situations.

Table 3 presents the performance comparison of different fraud detection models and hybrid configurations using standard classification metrics, including Average Precision (AP), F1-score (Fraud class), Precision, and Recall. The baseline XGBoost model achieved an AP of 0.790, F1-score of 0.718, precision of 0.742, and recall of 0.697, indicating a balanced performance between detecting fraudulent transactions and minimizing false alarms.

Hybrid 1 (XGB + Isolation Forest with weighted fusion) slightly improved the anomaly ranking performance, increasing AP from 0.790 to 0.7908, representing a marginal improvement of 0.0008 (0.1%). However, the F1-score (0.718), precision (0.742), and recall (0.697) remained unchanged compared to the baseline model. This suggests that the weighted fusion strategy slightly enhanced anomaly ranking while maintaining the overall classification behavior of the supervised model.

In contrast, Hybrid 2 (XGB + Autoencoder with broken scaling) resulted in a dramatic degradation in performance. The AP dropped sharply from 0.790 to 0.01, representing a 98.7% decrease, while the F1-score decreased from 0.718 to 0.02, and precision declined from 0.742 to 0.01. Meanwhile, recall increased substantially from 0.697 to 0.97, indicating that the model flagged almost all transactions as fraudulent. This imbalance resulted in an extremely high number of false positives, demonstrating that improper feature scaling severely disrupts the model's ability to distinguish fraudulent and legitimate transactions.

Hybrid 3 (XGB + Autoencoder with fixed scaling) significantly improved performance compared to Hybrid 2. The AP increased from 0.01 to 0.79, representing an improvement of 0.78 points, while the F1-score rose to 0.331, precision to 0.209, and recall to 0.793. Compared with the baseline XGBoost model, this configuration increased recall by 0.096 (13.8%), but reduced precision by 0.533, which consequently lowered the F1-score. These results highlight the importance of proper feature scaling in autoencoder-based anomaly detection systems.

Finally, the proposed hybrid model (XGB + Autoencoder + Meta-classifier) achieved an AP of 0.724, F1-score of 0.7033, precision of 0.6895, and recall of 0.7176. Compared with the baseline XGBoost model, the recall improved from 0.697 to 0.7176 (+0.0206 or 3% improvement), while precision decreased slightly by 0.0525, and the F1-score decreased marginally by 0.0147. Although the baseline model achieved a slightly higher AP, the hybrid framework provided a more balanced detection capability by integrating both supervised and unsupervised signals, improving robustness in identifying diverse fraud patterns.

3.8 Practical implications

Apart from the numerical results, the final hybrid yielded prospective usability regarding successful implementation in a dedicated web-based application. The tool enables real-time inference, batch inference, explanation ability via probability scores and component contributions. This type of evaluation, aimed at utilizing research models in real-world finance systems, was stressed in [1,4].

4. Conclusion

In this study, a hybrid credit card fraud detection approach combining supervised and unsupervised machine learning algorithms is applied to address the limitations and drawbacks of using a single model for handling imbalanced credit card transaction data. The proposed framework combines data preprocessing, classification, anomaly detection, and meta-level fusion in a practical pipeline rather than focusing solely on developing new algorithms. Initially, several supervised models were evaluated, and among them the XGBoost model demonstrated strong capability in learning complex fraud patterns from labeled data, achieving an Average Precision (AP) of 0.79, an F1-score of 0.719, precision of 0.743, and recall of 0.697. In parallel, an unsupervised autoencoder-based approach was implemented to learn the patterns of normal financial transactions and detect anomalous behavior. While these models showed promising results individually, their limitation such as reduced sensitivity in supervised models and high false positive rates in unsupervised anomaly detection highlighted the need for a combined strategy. To address these limitations, a hybrid framework based on a meta-classifier was developed to combine supervised fraud probabilities with unsupervised anomaly scores for adaptive final predictions. In this strategy, traditional fixed heuristic weights were replaced by weights learned from validation data through the meta-classifier, enabling a more data-driven and flexible fusion of model outputs. The final hybrid model achieved an AP of 0.724, an F1-score of 0.703, a precision of 0.689, and a recall of 0.718, demonstrating improved recall stability and greater robustness in detecting diverse fraud patterns, even though the standalone XGBoost model slightly outperformed it in certain metrics. Furthermore, the implementation of a real-time web-based system based on the proposed framework demonstrated the feasibility of deploying such hybrid models in operational environments. The system supports real-time transaction monitoring, batch analysis, and visualization of fraud detection results, illustrating the practical applicability of the proposed approach in real-world financial systems. All experiments in this study were conducted using the BankSim synthetic dataset. While this dataset enables controlled and reproducible experimentation and captures several characteristics of real transaction data, further validation on additional and real-world financial datasets is necessary to evaluate the generalizability of the proposed framework. Future research will focus on evaluating the model using real financial transaction datasets, integrating explainable AI techniques to improve transparency, and exploring online or adaptive learning mechanisms to better capture the evolving nature of fraud patterns.

Credit Author Contribution Statement

Tejas Gandhi: Conceptualization; Methodology; Software implementation; Formal analysis; Investigation; Data curation; Writing-original draft; Visualization. **Pragati Gupta:** Methodology; Data Curation; Formal analysis; Validation; Writing-review & editing. **Chirag Gandhi:** Literature review; Validation; Writing-review & editing. **Sarvesh Gagare:** Investigation; Formal analysis; Visualization; Writing-review & editing. **Vaishali Rajput:** Supervision; Conceptualization; Methodology; Validation; Writing-review & editing. **Rohini Chavan:** Supervision; Conceptualization; Methodology; Validation; Writing-review & editing. All authors have read and agreed to the published version of the manuscript.

Funding Declaration

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability Statement

The experimental data generated and analyzed during this study for system evaluation and testing, including model configuration details, validation results, and the application used to validate the hybrid model within the web-based analysis system, are available from the corresponding author upon reasonable request.

Conflict of Interest

There are no conflicts of interest.

Artificial Intelligence (AI) Use Disclosure

The authors confirm that there was no use of artificial intelligence (AI)-assisted technology for assisting in the writing or editing of the manuscript and no images were manipulated using AI.

Supporting Information

Not applicable.

References

[1] T. V. Jaswant, G. S. Manoj, V. Vamisdhar, A. Aravind, S. V. Reddy, J. C. Patni, N. B. Rohit Kumar, Credit card fraud detection using machine learning: A comprehensive review, 2024 3rd Edition of IEEE Delhi Section Flagship Conference (DELCON), New Delhi, India, 2024, 1-4, doi: 10.1109/DELCON64804.2024.10866830.

[2] M. Zanin, M. Romance, R. Criado, and S. Moral, Credit card fraud detection through parenclitic network analysis, Complexity, 2018, 5764370, doi: 10.1155/2018/5764370.

[3] C. Phua, V. Lee, K. Smith, R. Gayler, A comprehensive survey of data mining-based fraud detection research, School of Business Systems, Monash University, Australia, Technical Report, 2010.

[4] S. P. Maniraj, A. Saini, S. D. Sarkar, S. Ahmed, Credit card fraud detection using machine learning and data science, *International Journal of Engineering Research & Technology*, 2019, 8, 110-115 doi: 10.17577/IJERTV8IS090031.

[5] J. Gao, Z. Zhou, J. Ai, B. Xia, and S. Coggeshall, Predicting credit card transaction fraud using machine learning algorithms, *Journal of Intelligent Learning Systems*, 2019, **11**, doi: 10.4236/jilsa.2019.113003.

[6] V. Ceronmani Sharmila, Kiran Kumar R., Sundaram R., Samyuktha D., Harish R., Credit card fraud detection using anomaly techniques, 2019 1st International Conference on Innovations in Information and Communication Technology (ICIICT), Chennai, India, 2019, 1-6, doi: 10.1109/ICIICT1.2019.8741421.

[7] Y. Devavarapu, R. R. Bedadhala, S. S. Shaik, C. R. K. Pendela, K. Ashesh, Credit card fraud detection using outlier analysis, Proc. 4th Int. Conf. on Intelligent Technologies (CONIT), Karnataka, India, Jun. 21–23, 2024.

[8] M. A. Islam, M. A. Uddin, S. Aryal, G. Stea, An ensemble learning approach for anomaly detection in credit card data with imbalanced and overlapped classes, *Journal of Information Security and Applications*, 2023, **78**, 103618, doi: 10.1016/j.jisa.2023.103618.

[9] S. K. Sen, S. Dash, Meta-learning algorithms for credit card fraud detection, *Universal Journal of Engineering Research and Development*, 2013, **6**, 16–20.

[10] S. J. Stolfo, D. W. Fan, W. Lee, A. L. Prodromidis, P. K. Chan, Credit card fraud detection using meta-learning: Issues and initial results, Proceedings of AAAI Workshop on AI Methods in Fraud and Risk Management, 1998, 83–90.

[11] J. Onyeama, Credit card fraud detection in the Nigerian financial sector: A comparison of unsupervised TensorFlow-based anomaly detection techniques, autoencoders, and PCA, arXiv, 2024, doi: 10.48550/arXiv.2407.08758.

[12] H. Du, H. Wang, A. Guo, A novel method for detecting credit card fraud problems, *PLoS ONE*, 2024, 19, e0294537, doi: 10.1371/journal.pone.0294537.

[13] F. Khaled Alarfaj, S. Shahzadi, Enhancing fraud detection in banking with deep learning: Graph neural networks and autoencoders for real-time credit card fraud prevention, *IEEE Access*, 2025, **13**, 20633-20646, doi: https://doi.org/10.1109/ACCESS.2024.3466288.

[14] B. Chugh, N. Malik, D. Gupta, B. S. Alkahtani, A probabilistic approach-driven credit card anomaly detection with CBLOF and isolation forest models, *Alexandria Engineering Journal*, 2025, **114**, 231–242, doi: 10.1016/j.aej.2024.11.054.

[15] B. Ingre, A. Yadav, A. K. Soni, Decision tree-based intrusion detection system for NSL-KDD dataset, Information and Communication Technology for Intelligent Systems (ICTIS 2017), Smart Innovation, Systems and Technologies, vol. 84, Springer, Singapore, 2018.

[16] G. Ketepalli, S. Tata, S. Vaheed and Y. M. Srikanth, Anomaly detection in credit card transaction using deep learning techniques, 2022 7th international conference on communication and electronics systems (ICCES), Coimbatore, India, 2022, 1207-1214, doi: 10.1109/ICCES54183.2022.9835921.

[17] K. K. Renganathan, J. Karuppiah, M. Pathinathan, S. Raghuraman, Credit card fraud detection with advanced graph-based machine learning techniques, *Indonesian Journal of Electrical Engineering and Computer Science*, 2024, 35, 1963–1975, doi: 10.11591/ijeecs.v35.i3.pp1963-1975.

[18] P. Kar, S. Banerjee, K. C. Mondal, G. Mahapatra, S. Chattopadhyay, A hybrid intrusion detection system for hierarchical filtration of anomalies. In: Satapathy, S., Joshi, A. (eds) Information and Communication Technology for Intelligent Systems, Smart Innovation, Systems and Technologies, Springer, Singapore, 2018, 106, https://doi.org/10.1007/978-981-13-1742-2_41.

[19] J. Jabez, S. Gowri, S. Vigneshwari, J. Albert Mayan, S. Srinivasulu, Anomaly detection by using CFS subset and neural network with WEKA Tools. In: Satapathy, S., Joshi, A. (eds) Information and Communication Technology for Intelligent Systems, Smart Innovation, Systems and Technologies, Springer, Singapore, 2019, 107, doi: 10.1007/978-981-13-1747-7_66.

[20] A. Dal Pozzolo, O. Caelen, R. A. Johnson, G. Bontempi, Calibrating probability with undersampling for unbalanced classification, Proceedings of the IEEE Symposium Series on Computational Intelligence (SSCI). 2015, doi: 10.1109/SSCI.2015.33.

[21] M. Ahmed, A. N. Mahmood, J. Hu, A survey of network anomaly detection techniques, *Journal of Network and Computer Applications*, 2016, **60**, 19-31, doi: 10.1016/j.jnca.2015.11.016.

maps and institutional affiliations.

Citation

T. V. Gandhi, P. P. Gupta, C. S. Gandhi, S. D. Gagare, V. Rajput, R. Chavan, Credit card fraud detection using hybrid XGBoost and autoencoder models, Journal of Information and Communications Technology: Algorithms, Systems and Applications, 2026, 2(1), 26303, doi: 10.64189/ict.26303.