SvedbergOpen
DISSEMINATION OF KNOWLEDGE
International Journal of Artificial Intelligence and Machine Learning
Publisher's Home Page: https://www.svedbergopen.com/
Research Paper Open Access

Ensemble Learning Framework for Financial Fraud Detection Using Temporal Transaction Pattern Analytics

Dr. S. Kavitha[1] , G. Krishnamoorthy[2] , Dr. K.Kiruthika Devi[3] , Dr. N. Nagarani[4] , Elangovan Muniyandy[5] , Dr. R.Mohan Kumar[6]

[1]Computer Science Mother Teresa Women's University, Sakthi College of Arts and Science for Women, Dindigul . Mail id: kavithapari40@gmail.com
[2]Assistant professor, Department of Artificial intelligence and data science, R.V.S College of Engineering, Dindigul. Email id : pnggkrishnamoorthy@gmail.com
[3]Associate professor, Sri Venkateswara College of Engineering Sriperumbudur, Email id : kiruthika@svce.ac.in, Orcid : 0000-0001-7149-6064
[4]Assitant Professor, Electronics and Communication Engineering, Velammal College of Engineering and Technology, Madurai. Email id: yazhni6@gmail.com
[5]Department of Biosciences, Saveetha School of Engineering. Saveetha Institute of Medical and Technical Sciences, Chennai - 602 105, Email id: muniyandy.e@gmail.com
[6]Professor, Department Bio medical Engineering , PSR Engineering College . Sivakasi, Sevalpatti, Appayanaickenpatti, Tamil Nadu 626140 email id : rmohankumar@psr.edu.in , Orcid: https://orcid.org/0009-0004-4914-2269
Corresponding author mail id : kavithapari40@gmail.com

Abstract
Financial fraud has increased with the rapid growth of digital payment platforms, making accurate fraud detection a major challenge. Traditional machine learning (ML) methods often struggle to adapt to evolving fraud patterns, effectively capture temporal transaction behaviors, and accurately distinguish fraudulent transactions. To address these limitations, research proposes an Ensemble Learning Framework (ELF) for intelligent financial fraud detection using temporal transaction pattern analytics. The framework utilizes a financial fraud transaction analytics dataset that includes 15,000 instances of financial transactions and 44 features of temporal transactions comprising time-stamped transaction records, customer information, transaction attributes, and fraud labels. Initially, the preprocessing stage employs K-Nearest Neighbors (KNN) imputation for missing value handling, followed by Isolation Forest and Local Outlier Factor (LOF) for outlier detection. Subsequently, Neighborhood Components Analysis (NCA) is employed to extract highly discriminative transaction features while preserving class separability. Finally, the framework integrates the proposed Dynamic Artificial Rabbits Optimization-Driven Intelligent Light Gradient Boosting Machine (DARO-ILightGBM), where the Dynamic Artificial Rabbits Optimization algorithm performs adaptive hyperparameter optimization, while the Intelligent Light Gradient Boosting Machine Model accurately classifies fraudulent transactions. Experimental evaluation demonstrates that the proposed framework achieves a superior Receiver Operating Characteristic–Area under the Curve (ROC-AUC) of 0.9765, Precision–Recall Area under the Curve (PR-AUC) of 0.9677, F1-score of 0.9720, Precision of 0.9740, and Recall of 0.9685using Python 3.11. Overall, the proposed ELF effectively captures temporal transaction patterns and provides a scalable solution for intelligent financial fraud detection in modern digital banking environments.

Keywords: Financial Fraud Detection, Ensemble Learning, Temporal Transaction Analytics, Fraudulent Transactions.

1. Introduction

The detection of fraudulent transactions has become an important area of research in today's financial system owing to the increase in the number of electronic transactions, digital banking, and financial technology. This has made the process of doing transactions easy, but at the same time, it has posed complicated security challenges [1]. This technique has been used in many fields, such as spamming on social networks, intrusions into the network, identity checks, and the identification of fraudulent transactions, with the main goal of detecting suspicious actions and lowering possible financial risks [2]. Financial fraud can be defined as acts of deception carried out by people to occur some illegal money, namely payment fraud, money laundering, insurance fraud, financial account manipulation, and related financial crimes [3]. While digital payment methods and technology-enabled financial services made transactions more convenient and accessible, it also created highly sophisticated means for the fraudster to exploit any loopholes in the financial systems [4]. Occupational and transactional fraud caused significant financial losses, and businesses were severely affected by such practices because of ineffective means of detecting these crimes. Given the growing complexity of financial networks and customers' transaction behaviors, it is crucial to create effective means for detecting fraud [5].

Detection of financial fraud is considered to be a very difficult problem due to the dynamic nature of fraud, their huge class imbalances, and ever-changing behaviour. In practical applications for financial fraud detection, only about 1% of all transactions can be fraudulent, which makes traditional methods of fraud detection unsuitable for such a minority class [6]. At the same time, fraudsters change their methods from time to time to deceive the security system, thus requiring intelligent systems that can learn transactional behaviour and find new anomalies. The growing effect of financial fraud on a global scale led to serious economic losses and even damaged financial infrastructures [7]. Fraud in banking and digital payments involves different forms of illegal acts such as identity fraud, credit card fraud, phishing schemes, fraudulently accessing bank accounts, and transaction manipulations [8]. Involvement in such types of fraud leads to high financial expenses for organizations due to investigation cost and compensation to customers. As a result, certain measures developed to reduce fraud, such as anti-money laundering laws and compliance with financial security practices [9]. It is necessary to develop efficient methods of financial fraud detection to prevent organizations from losses and ensure financial stability within the global financial system [10]. High-quality transaction data and labels are required for the framework, and it performs poorly on unknown fraud schemes.

1.1 Research Aim

To develop a smart ensemble learning methodology for financial fraud detection using time-series electronic payment transactions. The approach enhances preprocessing and optimization-based classification to accurately distinguish fraudulent and legitimate transactions, reduce false positives, improve prediction accuracy, and ensure scalable, reliable decision support.

1.2 Research Organization

The research is structured into five sections. In Section 1, the introduction and background of financial fraud detection are provided. Section 2 examines the existing literature on fraud detection in the financial sector. Section 3 discusses the methodology of ELF with the DARO-ILightGBM Model, preprocessing, and feature extraction. Section 4 includes experimental results and performance analysis. The conclusion of the research is discussed in Section 5.

2. Related Work

The existing research section examined previous research by assessing their purpose, methodology, findings, and weaknesses to determine the gaps in the area of Financial Fraud Detection.Table 1 summarizes objectives, methods, results, and limitations of existing studies.

Table 1: Comparative analysis of existing financial fraud detection research methodologies and limitations

Ref.	Objective	Method	Result	Limitation
[11]	Detect fraudulent healthcare insurance transactions using sequential transaction analysis.	Sequence analysis integrated with a rule engine to generate frequent patterns and confidence scores.	Achieved a 96.5% fraud detection rate.	Performance depended on historical transaction patterns, reducing adaptability to emerging fraud types.

[12]	Improve fraud classification for imbalanced transaction data.	Optimized Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM) model combined with the Synthetic Minority Over-sampling Technique (SMOTE).	Achieved 83% precision.	Required extensive hyperparameter tuning and incurred high computational cost.
[13]	Identify financial fraud by analyzing sparse transaction behaviors.	Combined ML, Deep Learning (DL), and Natural Language Processing (NLP) techniques on the Fraud Natural Language Processing (NLP) dataset.	Achieved an F1-score of 97.5%.	Limited ability to generalize across different fraud scenarios.
[14]	Improve fraud detection performance for imbalanced financial transaction data.	Applied Generative Adversarial Network (GAN) based data augmentation with Recurrent Neural Network (RNN), LSTM, and Gated Recurrent Unit (GRU) models.	The GAN-GRU model achieved high detection sensitivity.	Detection performance depended on the quality of synthetic data and required substantial computational resources.
[15]	Detect Internet financial fraud using graph-based transaction modeling.	Transaction Amount-based Structural Context to Vector Representation (TA-Struc2Vec) was employed to extract structural relationships and transaction amount features.	Achieved an F1-score of 95.7%.	Required high-quality graph structures and involved considerable computational overhead.
[16]	Enhance financial fraud detection using quantum graph learning techniques.	Implemented a Quantum Graph Neural Network (QGNN) integrated with Variational Quantum Circuits (VQC).	Achieved an AUC of 85%.	Relied on quantum computing infrastructure and faced scalability challenges.
[17]	Detect evolving financial fraud through transaction network analysis.	Developed Fraud Graph Neural Network-Reinforcement Learning (GNN-RL) by integrating Temporal-Spatial-Semantic Graph Convolution (TSSGC),Deep Q-Network (DQN), and federated learning.	Achieved an F1-score of 97.3% while reducing false positives by 31%.	Incurred high communication overhead and depended on distributed data availability.
[18]	Predict financial fraud in listed companies using machine learning techniques.	Applied feature selection with a hybrid Logistic Regression- Extreme Gradient Boosting (LR-XGBOOST) classifier.	Achieved high classification accuracy.	Limited transaction diversity affected adaptability to evolving fraud patterns.
[19]	Improve fraud detection accuracy and model generalization across financial applications.	Proposed a Transformer-based Distributed Knowledge Distillation (DKD) framework with multi-attention mechanisms.	Achieved 81.48% precision.	Knowledge distillation increased model complexity and training overhead.
[20]	Detect fraudulent financial	Developed an XGBoost–Deep Neural Network (XGBoost-	Achieved 0.78 precision.	Computational complexity increased

	transactions while supporting regulatory compliance.	DNN) ensemble using transaction, temporal, and identity features.		when processing large-scale and low-quality datasets.

The ELF overcomes the above-mentioned limitations by means of NCA, DARO, and ILightGBM for discriminative feature extraction, adaptation and optimization, as well as accurate fraud identification. Consequently, this method significantly increases the quality of detection, generalization, and identification of fraudulent transactions.

3. Methodology

ELF is proposed for intelligent financial fraud detection through analysis of temporal transaction patterns. Financial transaction data collection with time-stamped transactions, customer details, transaction features, and fraudulent labels is used. The preprocessing phase consists of KNN imputation for dealing with missing values, followed by isolation forest and LOF algorithms for outlier detection. The NCA method executes discriminative feature extraction. The ELF finally uses the DARO-ILightGBM for adaptive optimization-based classification, hyperparameter tuning, and intelligent classification of fraud with better detection performance. The methodology workflow for financial fraud detection is evaluated in Figure 1.

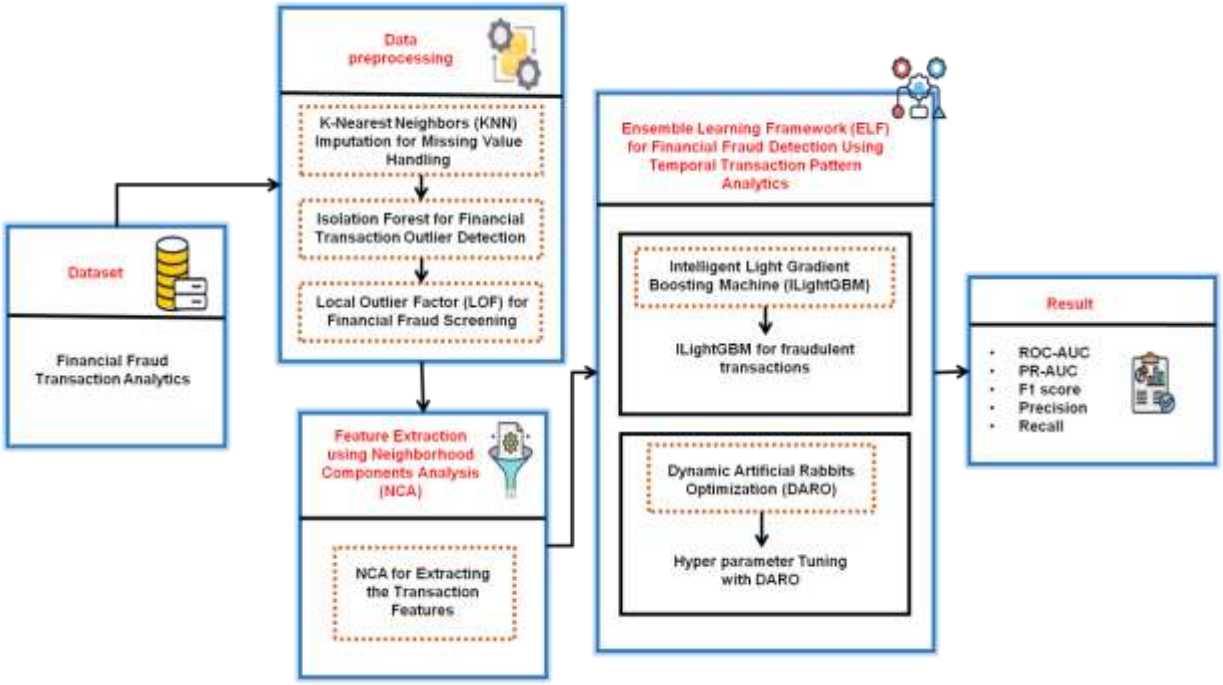

Figure 1: Methodology Workflow for the Financial Fraud Detection

3.1 Data Collection

The set of financial transaction data was acquired from the Kaggle dataset (https://www.kaggle.com/datasets/colabsss/financial-fraud-transaction-analytics). It contains 15,000 instances of financial transactions along with 44 features representing temporal transaction analytics, customer details, account details, transaction details, authentication information, and a label for fraud. The dataset provides comprehensive descriptions of legal and illegal financial transactions for temporal fraud detection. 80% of the records are kept aside for training purposes, while 20% are kept aside for performance testing of the suggested ELF.

3.2 Data Preprocessing Methods for Improving the Financial Transaction Data

Data preprocessing refers to the process through which financial transaction data that is initially raw is cleaned and transformed to enhance its quality and consistency. The KNN imputation is employed first for addressing

any cases of missing values, while the detection of anomalies in the transactions is done using the Isolation Forest and the LOF methods.

➢ **K-Nearest Neighbors (KNN) Imputation for Missing Value Handling**

Missing value imputation is performed by making estimations about the missing transactions based on information obtained from similar neighbors. This process requires that incomplete transaction data be provided as an input and results in a complete dataset for effective detection of financial fraud, as expressed in Equations (1) and (2).

$$s(b,a) = \frac{Cov(b,a)}{\sqrt{Var(b)}\sqrt{Var(a)}} \quad (1)$$

$$Cov(b,a) = \frac{\sum(b_j - \overline{b})(a_j - \overline{a})}{m} \quad (2)$$

Where $s(b,a)$ represents thesimilarity score of two financial transaction records. $b$denotes thetransaction record having the missing attribute, and$a$ refers to the neighboring transaction. $Cov(b,a)$ stands for the covariance between transaction record$b$ and a. $\sqrt{Var(b)}$represents the selected transaction attributes. $\sqrt{Var(a)}$ refers to the neighboring selected transaction attributes. In Equation (2), $b_j$represents thevalue of the j^{th} feature in transaction record b.a_j refers to the value of the neighboring transaction record. $\overline{b}$ denotes the mean value of all selected features in a transaction record. $\overline{a}$ stands for the mean value of all selected features in a transaction record. j denotes each transaction feature (transaction hour), and m refers to the $\sum$ summation of total transactions.

➢ **Isolation Forest for Financial Transaction Outlier Detection**

Isolation forest finds anomalous transaction records by isolating individual data points using random splits. It takes the imputed transaction dataset as an input, creates the isolation tree to identify the anomalies, and generates a filtered dataset by excluding the outliers, as expressed in Equations (3) and (4).

$$t(y,m) = 2^{\frac{F(g(y))}{D(m)}} \quad (3)$$

$$D(m) = 2G(m-1) - \frac{2(m-1)}{m} \quad (4)$$

Where $t(y,m)$ represents the anomaly score, which is used to detect suspicious financial transactions. $y$refers to the financial transaction (transaction ID, customer ID, account ID). m denotes the total number of financial transactions used to build an Isolation Forest. $F\big(g(y)\big)$ stands for the expected path length that to be needed to isolate financial transaction y in the Isolation Trees. $D(m)$ refers to thenormalization constant that normalizes the path length depending on the size of the dataset. In Equation (4), $2G(m-1)$ refers to theharmonic number to estimate the average search depth for $m-1$ transaction records.

➢ **Local Outlier Factor (LOF)for Financial Fraud Screening**

The LOF detects outliers from transaction data through density deviation analysis. It takes the dataset filtered using Isolation Forest as its input, analyzes density deviation within neighborhoods, and generates a transaction dataset containing reliable data for financial fraud detection, as expressed in Equations (5) and (6).

$$G_v^j = \begin{cases} 1, if\ LOF_l(v_j) \geq \delta^v \\ 0, if\ LOF_l(v_j) < \delta^v \end{cases} \quad (5)$$

$$G_v^j = \begin{cases} 1, if\ LOF_l(U_j) \geq \delta^S \\ 0, if\ LOF_l(U_j) < \delta^S \end{cases} \quad (6)$$

Where G_v^jrepresents theoutlier indicator for customer data in financial fraud detection, $v_j$refers to the transaction data associated with customer ID, and LOF_l denotes thelocal Outlier Factor value for transaction data. l stands for the number of nearest transaction data considered for density calculation. δ^vrepresents thevalue of the threshold used to detect outliers from transaction data; 1 refers to the transaction data being classified as an outlier, and 0 denotes the transaction data being considered to be a normal transaction. In Equation (6),U_j represents the financial transaction data point (transaction datetime). δ^Srefers to thethreshold for separating outliers from the normal transactions.

3.3 Extracting Transaction Features using Neighborhood Components Analysis (NCA)

Feature extraction involves converting transaction data into a high-dimensional form to create a discriminative feature space. NCA takes the processed transaction features as input and learns their importance based on maximizing class separability to produce an optimal set of features for financial fraud detection, as expressed in Equations (7) and (8).

$$q_{ji} = \frac{-e(y_j, y_i)}{\sum_{l \notin j} -e(y_j, y_l)} \quad (7)$$

$$K_{NCA} = -log\left(\frac{\sum_{l \notin D_j} exp(-e(y_j, y_i))}{\sum_{l \notin D_j} exp(-e(y_j, y_l))}\right) \quad (8)$$

Where q_{ji}represents theprobability of choosing transaction record i to be the neighbor of transaction record. y_j refers to the feature vector for the j^{th}financial transaction (day of week, transaction amount). y_i denotes the feature vector for the i^{th} neighboring financial transaction. e stands for the distance between transaction feature vectors. $\sum_{l \notin j}$ refers to the summation across all transactions but not the current transaction j. y_l denotes the feature vector of the l^{th} financial transaction analyzed during neighborhood similarity calculation. In Equation (8), K_{NCA}represents the NCA objective function for learning discriminatory transaction features. $\sum_{l \notin D_j}$refers to theset of neighboring transaction records for fraud category as transaction j.$exp(\cdot)$ denotes the exponential function that converts distances between features to similarities. $log(\cdot)$ stands for the logarithmic function for optimizing the separability of transaction features.

All the discriminative features of transactions derived by NCA are sent to the ELF. The input of the features in ELF is the DARO method, where adaptive hyperparameter optimization is done; after that, the optimized features and hyperparameters are sent to the ILightGBM for financial fraud detection.

3.4 Ensemble Learning Framework (ELF) for Financial Fraud Detection Using Temporal Transaction Pattern Analytics

The ELF framework makes use of the transaction features generated by NCA to feed into the fraud classification process. The process of dynamic adaptive hyperparameter optimization is done using DARO, and the ILightGBM classifier identifies fraud through optimized features. The framework yields effective results for fraudulent transaction detection in finance through accurate classification.

➢ Intelligent Light Gradient Boosting Machine (ILightGBM) for fraudulent transactions

Light Gradient Boosting Machine (LightGBM) is a gradient boosting ML method for effective classification; nevertheless, its performance can be hampered by poor choices of hyperparameters and insufficient adaptation to temporal fraud features. ILightGBM compensates for these shortcomings with the help of DARO-based hyperparameters that provide accurate classification of fraudulent and non-fraudulent transactions, as expressed in Equation (9).

$$\overline{y}_j = \frac{y - y_{min}}{y_{max} - y_{min}} \quad (9)$$

Where $\overline{y}_j$ represents thenormalized transaction feature that serves as an input into the ILightGBM algorithm. y refers to the raw amount of the corresponding transaction feature (transaction amount). y_{min} denotes the minimal quantity of the corresponding transaction information variable. y_{max} stands for the maximal value of the corresponding transaction descriptive variable. j represents the index of the normalized transaction feature (browsertype).

➢ Hyperparameter Tuning with Dynamic Artificial Rabbits Optimization (DARO)

The Artificial Rabbits Optimization (ARO) is one of the nature-inspired optimization algorithms used to solve optimization problems, but it can experience slow and premature convergence in a high-dimensional search space. The DARO solves the problems faced by the traditional method by using adaptive search methods to generate optimal hyperparameters for the ILightGBM fraud classification method. The center-driven strategy process is expressed in Equations (10) and (11).

- **Center-driven strategy**

Fraud detection is enhanced through a center-driven strategy that balances exploration and exploitation using the best hyperparameter solution, a random solution, and a newly generated solution.

$$Y_d = \frac{b.Y_{best} + c.Y_s + d.Y_{new}}{b + c + d} \quad (10)$$

$$Y_{new} = kc + rand.(vc - kc) \quad (11)$$

Where Y_d represents the financial fraud detection problem, $b$refers to thebest fraud detection, and c denotes the random exploration of fraud detection. d represents the new fraud detection strategy. Y_{best} refers to the best hyperparameters for maximum financial fraud detection. Y_s denotes the random hyperparameters for exploring fraud detection. Y_{new} represents the new hyperparameters for fraud detection. In Equation (11), kc and vc refer to the lower and upper limits for the search for fraud detection. $rand$ denotes the random search for fraud detection. The Gaussian Random Walk process is expressed in Equations (12) and (13).

- **Gaussian Random Walk**

Fraud detection is improved using the Gaussian Random Walk technique, which enhances local search by performing random walks around the optimal hyperparameter set. This enables DARO to escape local optima and identify optimal hyperparameters for more accurate fraud detection.

$$Y_{next} = Gaussian(Y_{best}, \sigma) + (rand.Y_{best} - rand.Y_j) \quad (12)$$

$$\sigma = \frac{In(GF_t)}{GF_t}(Y_j - Y_s) \quad (13)$$

Where Y_{next} represents the hyperparameters updated for precise fraud detection; $Gaussian(Y_{best}, \sigma)$ refers to the best fraud detection, and Y_{best}denotes the best fraud detection. σ represents the step size for optimizing fraud detection. $rand$ refers to the exploration of randomness for optimization. Y_j denotes the current fraud detection. In Equation (13), $In(GF_t)$ represents thelogarithmic factor influencing the convergence of the optimization, Y_j refers to thecurrent solution for fraud detection hyperparameters, and GF_t stands for the optimization iteration factor at iteration t, used to regulate the convergence behavior of the DARO algorithm. Y_s denotes the random solution for fraud detection.

4. Result

The proposed ELF showed better performance in detecting financial fraud owing to its capacity to extract time-dependent patterns of the transactions and classify fraudulent transactions accurately against genuine transactions. Utilization of KNN Imputation, Isolation Forest, LOF, NCA, and DARO-ILightGBM algorithms led to enhanced ROC-AUC, PR-AUC, F1 score, Precision, and Recall, with reduced false positives. The proposed framework performed much better than other available frameworks of ensemble learning and boosting for fraud detection.

4.1 Data Exploratory Analysis

The relationship between customer and transaction characteristics to detect financial fraud is depicted in Figure 2. This method aims to identify linear relationships and the various attributes used to differentiate fraudulent transactions from genuine ones. The features under consideration include Customer Age, which falls within the range of 20 to 80 years; Account Balance, with values up to 250,000; Monthly Transaction, which has up to 150 counts; and Distance from Home, which covers 400 kilometers.

Figure 2: Relationship between Customer and Transaction Characteristics

The financial value for fraud risks through high-risk transactions is described in Figure 3(a). The main idea behind this is to distinguish between valid and fake signatures based on continuous variables ranging from 0.0 to 1.0. The important variables are transaction amount, transactions last 1hour normalized to 5 intervals, failedlogin attempts, and previous declined transactions.Analysis of the time gap between consecutive transactions that enables fraud based on the speed of such transactions is outlined in Figure 3(b). To show the probability density functions of the amount of time. The X-axis indicates Time since the Last Transaction measured from 0 to more than 500 minutes, with very high probability densities at zero intervals up to 0.0065.

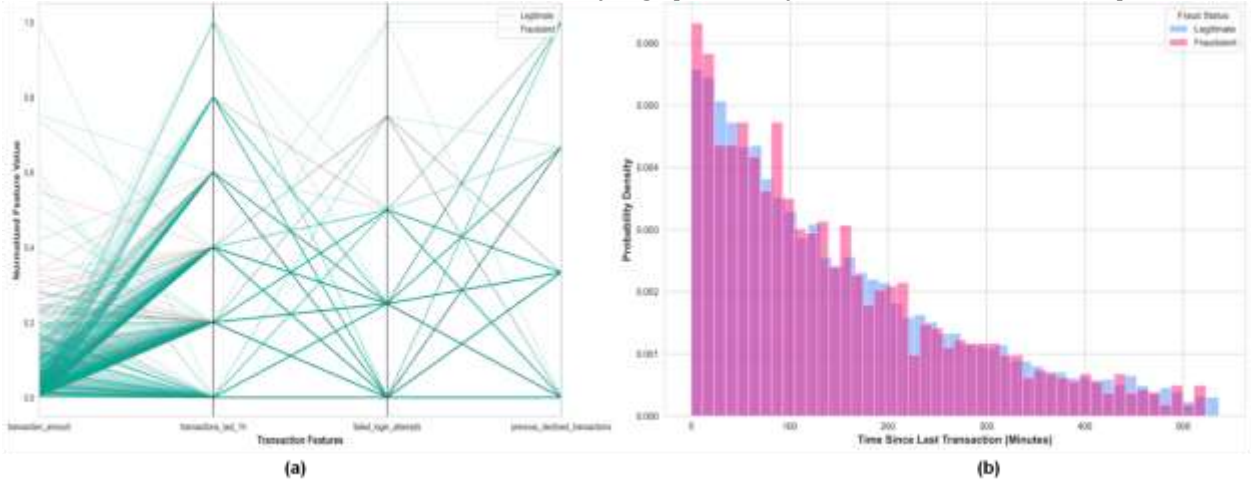

Figure 3: Visualization of (a) High-Risk Financial Transactions, and (b) Analysis of Time Intervals in Financial Transactions

The Amount Deviation Ratio Distribution across various transaction mediums for detecting any anomalies in terms of transaction surges is presented in Figure 4to detect behavior anomalies between fraud and valid transactions. Numeric features include Amount Deviation Ratios ranging from 0 to more than 100 for the transaction mediums, which include Automated Teller Machine (ATM), Application Programming Interface (API), Mobile App, Web Banking, and Point of Sale (POS) Terminal.

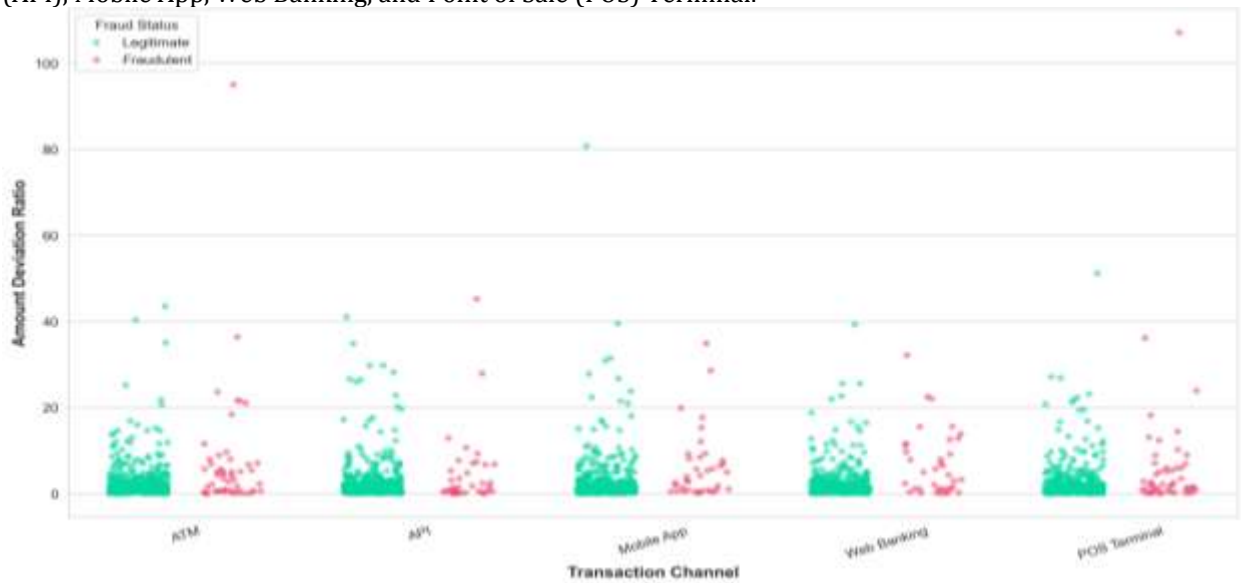

Figure 4: Analyzing theAnomalies Fraud Behavior and Valid Transactions

4.2 Evaluation Metrics

The evaluation metrics are used to measure the efficiency of the ELF in the detection of financial fraud. The ROC-AUC metric is applied to test the discriminating power of fraudulent and non-fraudulent transactions. The PR-AUC is applied to measure the performance of fraud detection in the presence of class imbalance. The F1-score is an efficient balance between the precision and recall metrics.Precision measures the number of correct fraud detections without triggering many false alarms. Recall is used to measure the ability to detect fraudulent transactions.

4.3 Performance Evaluation

The performance of the existing model known as XGB+DNN Ensemble [20] and the proposed DARO-ILightGBMmodel is evaluated on the existing IEEE-CIS fraud detectiondataset used for fraud detection in Table 1 and Figure 5. While the performance of the XGB+DNN Ensemble [20] model is good, the proposed DARO-ILightGBM model has performed better in terms of ROC-AUC(0.9689), PR-AUC(0.9574), F1-score(0.9610), Precision(0.9635), and Recall 0.9580.

Table 1: Comparison Performance of Existing and Proposed Methods for Financial Fraud Detection

Model	ROC-AUC	PR-AUC	F1-Score	Precision	Recall
XGB+DNN Ensemble [20]	0.9638	0.6582	0.74	0.78	0.69
DARO-ILightGBM [Proposed]	0.9689	0.9574	0.9610	0.9635	0.9580

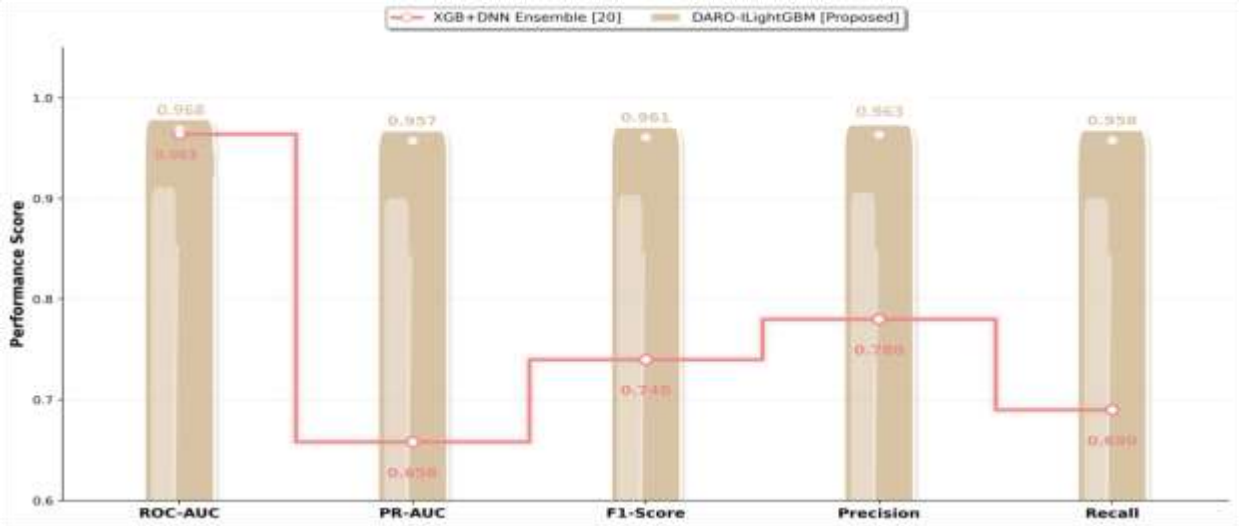

Figure 5: Comparative Analysis of Baseline and Proposed Models for Financial Fraud Detection

The retrained performance comparison between the existing and proposed methods of ELF for financial fraud detection is provided in Table 2 and Figure 6. The existing XGB+DNN Ensemble [20] model andproposed DARO-ILightGBMmodel were retrained on the financialfraudtransactionanalytics dataset under the same conditions, which include the dataset, preprocessing, feature extraction, and split. For the purpose of evaluating the effectiveness of existing and proposed approaches based on the ELF framework, training and evaluation are carried out using the proposed financialfraudtransactionanalytics dataset.While the XGB+DNN Ensemble gave similar performance, the LightGBM method performed better in the classification of fraud due to its superior boosting capabilities. The proposed method DARO-ILightGBM outperformed both the current methods in terms of ROC-AUC (0.9765), PR-AUC (0.9677), F1-score (0.9720), Precision (0.9740,) and Recall (0.9685).

Table 2: Retrained Results of Existing and Proposed Methods for Financial Fraud Detection

Model	PR-AUC	ROC-AUC	F1-Score	Recall	Precision
XGB+DNN Ensemble	0.7184	0.9740	0.83	0.75	0.86
LightGBM	0.9560	0.9750	0.9610	0.9580	0.9630
DARO-ILightGBM [Proposed]	0.9677	0.9765	0.9720	0.9685	0.9740

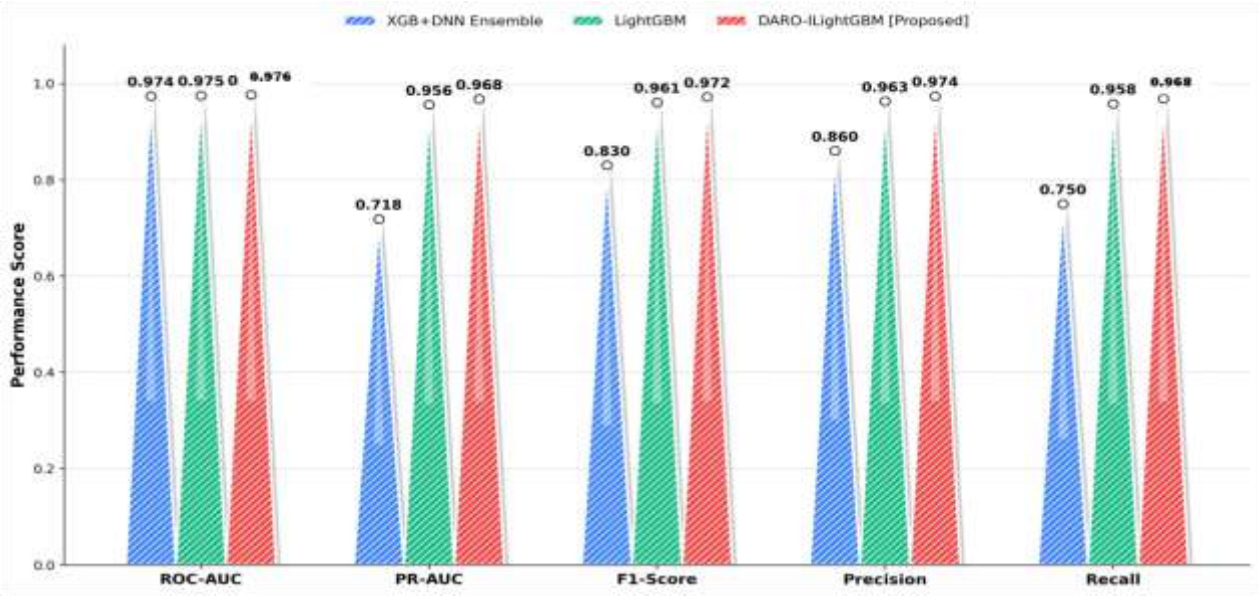

Figure6: Performance Evaluation of Existing and Proposed Methods for Financial Fraud Detection

4.4 Discussion

Enhance financial fraud detection by properly detecting fraudulent transactions using temporal analysis of transaction patterns. Theexisting model of the XGBoost-DNN ensemble [20] considered transaction, temporal, and identity features to detect fraud and comply with regulations; however, the complexity of the computation was affected by the scale of the data and different qualities of data, resulting in poor performance. However, the proposed ELF is able to solve these challenges through robust preproxcessing, discriminative feature extraction using SLCA, and hyperparameter tuning using DARO-ILightGBM for fraud detection.

5. Conclusion

The suggested ELF succeeded in realizing the purpose of improving financial fraud detection by temporal transaction analytics. A dataset of financial fraud transaction analytics with 15,000 transaction records and 44 temporal transaction features was used. Preprocessing incorporated KNN imputation, Isolation Forest, and LOF, while NCA extracted discriminative features. The suggested DARO-ILightGBM model scored 0.9765 for ROC-AUC, 0.9677 for PR-AUC, 0.9720 for F1 score, 0.9740 for Precision, and 0.9685 for Recall, achieving better performance in detecting financial fraud.The presented framework was limited by the fact that computational costs grow with increasing transaction volumes. Generalization in other domains needs to be validated.Future research pays attention to real-time fraud detection through the use of streaming transaction analysis, and Artificial Intelligence (AI)integrated to improve model transparency and interpretability.

References

1. Talukder, M. A., Khalid, M., &Uddin, M. A. (2024). An integrated multistage ensemble machine learning model for fraudulent transaction detection. Journal of Big Data, 11(1), 168. https://doi.org/10.1186/s40537-024-00996-5
2. Olaniyan, J., Olaniyan, D., Obagbuwa, I. C., &Ngafeeson, M. (2025). Graph-temporal contrastive transformer for financial fraud detection using transaction behavior modeling. Algorithms, 18(12), 770. https://doi.org/10.3390/a18120770
3. Wang, J., Liu, J., Zheng, W., &Ge, Y. (2025). Temporal heterogeneous graph contrastive learning for fraud detection in credit card transactions. IEEE Access, 13, 145754–145771. https://doi.org/10.1109/ACCESS.2025.3599787
4. Wang, X., Guo, J., Luo, X., & Yu, H. (2024). DyHDGE: Dynamic heterogeneous transaction graph embedding for safety-centric fraud detection in financial scenarios. Journal of Safety Science and Resilience, 5(4), 486–497. https://doi.org/10.1016/j.jnlssr.2024.05.005

5. Tripathi, R. K., Parveen, P., Kanchana, K., Chikkegowda, K. G., Gundapaneni, S., &Dineshkumar, J. (2026). Automated fraud detection in financial services using hybrid autoencoders and LSTM. International Journal of Artificial Intelligence and Machine Learning, 6(4s), 412–425.https://doi.org/10.51483/IJAIML.6.4s.2026.412-425
6. Anitha, K., Vinoth, R., Deolia, V. K., Thivya, R. S., Balakrishnan, S. G., &Bhushan, S. (2026). Financial fraud detection using Isolation Forest and DBSCAN clustering techniques. International Journal of Artificial Intelligence and Machine Learning, 6(3s), 142–150. https://doi.org/10.51483/IJAIML.6.2s.2026.142-150
7. Alarfaj, F. K., &Shahzadi, S. (2024). Enhancing fraud detection in banking with deep learning: Graph neural networks and autoencoders for real-time credit card fraud prevention. IEEE Access, 13, 20633–20646. https://doi.org/10.1109/ACCESS.2024.3466288
8. Hsin, Y. Y., Dai, T. S., Ti, Y. W., Huang, M. C., Chiang, T. H., & Liu, L. C. (2022). Feature engineering and resampling strategies for fund transfer fraud with limited transaction data and a time-inhomogeneous modus operandi. IEEE Access, 10, 86101–86116. https://doi.org/10.1109/ACCESS.2022.3199425
9. Li, J., & Yang, D. (2023). Research on financial fraud detection models integrating multiple relational graphs. Systems, 11(11), 539. https://doi.org/10.3390/systems11110539
10. Abbassi, H., El Mendili, S., &Gahi, Y. (2025). Adaptive, privacy-enhanced real-time fraud detection in banking networks through federated learning and VAE-QLSTM fusion. Big Data and Cognitive Computing, 9(7), 185. https://doi.org/10.3390/bdcc9070185
11. Matloob, I., Khan, S. A., Rukaiya, R., Khattak, M. A. K., &Munir, A. (2022). A sequence mining-based novel architecture for detecting fraudulent transactions in healthcare systems. IEEE Access, 10, 48447–48463. https://doi.org/10.1109/ACCESS.2022.3170888
12. Jabeen, M., Ramzan, S., Raza, A., Fitriyani, N. L., Syafrudin, M., & Lee, S. W. (2025). Enhanced credit card fraud detection using deep hybrid CLST model. Mathematics, 13(12), 1950. https://doi.org/10.3390/math13121950
13. Boulieris, P., Pavlopoulos, J., Xenos, A., &Vassalos, V. (2024). Fraud detection with natural language processing. Machine Learning, 113(8), 5087–5108. https://doi.org/10.1007/s10994-023-06354-5
14. Mienye, I. D., & Swart, T. G. (2024). A hybrid deep learning approach with generative adversarial network for credit card fraud detection. Technologies, 12(10), 186. https://doi.org/10.3390/technologies12100186
15. Li, R., Liu, Z., Ma, Y., Yang, D., & Sun, S. (2022). Internet financial fraud detection based on graph learning. IEEE Transactions on Computational Social Systems, 10(3), 1394–1401. https://doi.org/10.1109/TCSS.2022.3189368
16. Innan, N., Sawaika, A., Dhor, A., Dutta, S., Thota, S., Gokal, H., Patel, N., Khan, M. A. Z., Theodonis, I., &Bennai, M. (2024). Financial fraud detection using quantum graph neural networks. Quantum Machine Intelligence, 6(1), 7. https://doi.org/10.1007/s42484-024-00143-6
17. Cui, Y., Han, X., Chen, J., Zhang, X., Yang, J. and Zhang, X., 2025. FraudGNN-RL: a graph neural network with reinforcement learning for adaptive financial fraud detection. IEEE Open Journal of the Computer Society, 6, pp. 426- 437. https://doi.org/10.1109/OJCS.2025.3543450
18. Zhao, Z., &Bai, T. (2022). Financial fraud detection and prediction in listed companies using SMOTE and machine learning algorithms. Entropy, 24(8), 1157. https://doi.org/10.3390/e24081157
19. Tang, Y., & Liu, Z. (2024). A distributed knowledge distillation framework for financial fraud detection based on transformer. IEEE Access, 12, 62899–62911. https://doi.org/10.1109/ACCESS.2024.3387841
20. Luong, N. D. A., &Xie, S. (2026). Explainable ensemble machine learning for financial transaction fraud detection: Insights from XGBoost and deep neural networks. The Journal of Finance and Data Science, 12, 100195. https://doi.org/10.1016/j.jfds.2026.100195