Comprehensive Framework for Financial Transaction Fraud Detection via Dimensionality Reduction with an Explainable Artificial Intelligence Approach

Lyudmila Chernikova[1,*], Svetlana Dreving[1], Olga Borisova[1], Tatiana Tazikhina[1]

[1]Department of Corporate Finance and Corporate Governance, Financial University under the Government of the Russian Federation, Moscow, 125167, Russia

Emails: lichernikova@fa.ru; srdreving@fa.ru; olvborisova@fa.ru; ttazihina@fa.ru

Abstract

The expanding growth of financial transactions has resulted in the development of fraud systems. These progressions have considerably improved overall productivity, improved corporate management, and reduced operational costs. With the expanded utilization of automated financial transaction, organization and businesses have progressed to digital platform, convert their financial operation. Still, such a change in addition revealed financial systems to new threats, mainly through fraudulent activity and cybercrime. The large datasets, incorporated with the limits of conventional fraud detection techniques, provide a chance to accept Artificial Intelligence (AI) methods. The fraud detection problem is addressed by using Explainable AI (XAI) to give specialists with explained AI predictions over different explanation models. This paper proposes a Financial Transaction Fraud Detection via Dimensionality Reduction with an Explainable Artificial Intelligence Approach (FTFD-DRXAIA) technique. The aim is to develop an effective and intelligent system for accurate fraud detection in financial transaction utilizing progressive deep learning (DL) methods. Initially, the min-max method is used for data pre-processing to convert raw data into an appropriate format. Furthermore, the recursive feature elimination (RFE) system is applied for feature selection. For financial fraud detection process, the Elman recurrent neural network (ERNN) has been utilized. Moreover, the wildebeest herd optimization (WHO) method fine-tunes the ERNN model's hyperparameters, resulting in improved classification performance. Finally, the XAI technique applies LIME and SHAP to interpret complex AI models, enabling auditors and analysts to detect suspicious transaction patterns with greater clarity and confidence. The experimental outcome of FTFD-DRXAIA system is examined under the financial fraud detection database. The comparison analysis of FTFD-DRXAIA algorithm demonstrated an optimum precision value of 98.96% over recent methods.

Keywords: Financial Transaction; Fraud Detection; Dimensionality Reduction; Explainable Artificial Intelligence; Wildebeest Herd Optimization; Min-Max

1. Introduction

Financial fraud is an important concern today, affecting not just the economic sector then the broader marketplace but digital payment systems are deployed. A credit card is a great financial instrument; vast numbers of people use it every single day [1]. The credit card sector deals with many fraudulent transactions daily due to its expanded usage. Recently, the ability to gather information from fraud cases has fueled the development of fraud detection [2]. With this, financial institutions continue to carry out thorough investigations to reduce and detect fraud. Nevertheless, fraud remains complex because of its continually altering approaches and different behaviours [3]. A leading field that is the emphasis of broad study is bank-related fraud. Bank account fraud varies from other

DOI: https://doi.org/10.54216/IJNS.270202

financial deceptions in its techniques, effects, and recognition difficulties [4]. In contrast to credit card fraud, but unauthorized transactions are quickly recognized because of rare spending patterns, bank account fraud can manifest itself in refined methods such as account takeovers, unauthorized funds transfers, or even notice theft resulting in the development of new accounts [5]. The significances for the victim are long lasting, either emotionally or financially. Understanding and decreasing these attacks requires detailed investigation, supported by robust and varied datasets [6].

The developing complexity of fraud has permitted scholars to produce sophisticated fraud detection methods, precisely over the combination of Artificial Intelligence (AI) and Machine Learning (ML) [7]. By examining a widespread financial database, ML methodologies might expose concealed patterns, recognise abnormalities, and classify suspicious transactions with extraordinary precision [8]. The ML usage in fraud detection has led to significant progress through many fields like cybersecurity, banking, and healthcare [9]. The crucial issues in ML-driven fraud detection are those of interpretability and explanation. Many standard ML approaches operate in a 'black box' way to some level, restraining the easier kind of their decision-making methods by regulatory authorities and financial analysts [10]. Currently, explainable AI (XAI) is presented as a method to handle interpretability concerns of ML-based approaches of fraud detection [11]. Moreover, XAI develops fraud detection clarity, thus maintaining data security and privacy [12].

1.2. Addressing Dynamic Threats in Modern Digital Payment Systems

Financial systems encounter numerous fraudulent activities due to the rapid growth of digital payment platforms that evolve in complexity and scale. Conventional detection methods often failed to identify subtle or recently emerging fraud patterns in real time. The requirement for adaptable, intelligent solutions has become crucial for securing financial transactions. Provide work for XAI and feature reduction increases detection accuracy but preserving interpretability. This direction empowers institutions to perform decisively against fraud but understanding the reasoning behind model predictions.

1.3. Key Contributions and Novel Aspects of the FTFD-DRXAIA Technique

This paper presents a Financial Transaction Fraud Detection via Dimensionality Reduction with Explainable Artificial Intelligence Approach (FTFD-DRXAIA). The purpose is to progress an effective and smart structure for accurate fraud detection in financial transactions exploiting cutting-edge DL approaches. To begin with, the min-max method has been utilized to convert raw data into suitable format. Additionally, the recursive feature elimination (RFE) model was executed for FS. For financial fraud detection process, the Elman recurrent neural network (ERNN) can be deployed. In addition, the wildebeest herd optimization (WHO) adjusts the ERNN classifier hyperparameters, resulting in better classification solution. Lastly, the XAI executes LIME and SHAP to interpret complex AI approaches. The experimental investigation of FTFD-DRXAIA system has been inspected concerning the financial fraud detection database.

The presented FTFD-DRXAIA methodology undergoes min-max standardization for rescaling the input data within a fixed range.

The RFE approach to methodically remove less crucial attributes and retain the most informative ones.

The FTFD-DRXAIA approach employs the ERNN technique for capturing temporal dependency in transaction sequences, allowing the method to learn behavioural patterns over time that are significant for detecting anomalies linked to fraudulent activity, thereby improving detection accuracy and providing a more dynamic comprehension of growing financial fraud tactics.

The FTFD-DRXAIA method implements the WHO model for fine-tuning the hyperparameters of the ERNN model by balancing exploration and exploitation in the search space, resulting in optimal parameter selection that enhances classification performance, mitigates error rates, and ensures the model adapts effectively to complex fraud detection scenarios in financial transaction data.

An integrating XAI methods like SHAP and LIME for proposing interpretable and transparent understandings into predictions of the model.

The proposed methods uniquely integrates ERNN with WHO for optimized fraud detection but mixing XAI models like LIME and SHAP. This novel integration develops detection accuracy and offers clear understandings into model decisions, addressing both solution and explainability in financial fraud analysis efficiently.

DOI: https://doi.org/10.54216/IJNS.270202

2. Literature Survey of Financial Fraud Detection

Rahmati [13] examined a fraud detection that incorporates federated learning (FL) and adaptive graph neural networks (GNNs) to address these restrictions. FL permits various financial organizations to collectively train fraud detection approaches without exchanging customer data, then tackling privacy issues. Mazumder et al. [14] proposed a new method for detecting anomalies in financial transactions through CNNs, a kind of DL technique recognized. The CNN architecture extracts sophisticated patterns that differentiate standard from abnormal action. This method encompasses an inclusive process, data normalization, feature learning models, and the formation of multi-channel representation to apply CNN. The authors [15] examined the synergistic combination of AI and blockchain (BC) to optimize detection. By using BC's tamper-proof ledger and AI's innovative pattern recognition abilities, financial institutions can attain real-time, clear, and precise detection of fraudulent activities.

Awosika et al. [16] presented a new method applying FL and XAI to overcome the problems. FL allows financial organizations to collectively train a method to recognize fake transactions, so upholding data secrecy. Li et al. [17] presented SEFraud, a new graph-based self-explainable fraud detection technique that concurrently addresses fraud detection and results in interpretability. Specifically, SE Fraud initially employs tailored, varied graph transformer systems with edge and learnable feature masks for learning meaningful representations from the beneficial heterogeneous transactions. Castellanos [18] explored the incorporation of AI and BC technologies as a groundbreaking technique for fraud detection and anomaly prevention in dynamic management systems. AI-based approaches, such as ML, DL, and predictive analytics, play a vital role in detecting fraudulent patterns practically. Yazdinejad et al. [19] designed and implemented a secure, smart fuzzy BC system. This system employs an innovative fuzzy DL method, an enhanced adaptive neuro-fuzzy inference system (ANFIS)-enabled threat recognition, a fuzzy control system (FCS), and fuzzy matching (FM) for detecting network threats. Utilizing metaheuristic algorithms for optimizing the threat recognition error functions in ANFIS.

2.1. Limitations and research gaps in advanced financial fraud detection techniques addressing challenges in scalability, privacy, interpretability, and real-world applicability

The limitations of the existing studies comprise issues with data representativeness, high computational needs, and restricted model interpretability. Various techniques rely on synthetic or centralized datasets that might not completely reflect the real-time diversity financial transactions, mitigating their generalizability across diverse environments. Scalability remains a significant threat as some models demand substantial computational resources and may encounter difficulty in maintaining accuracy when handling massive data volumes. Although the FL model assists with privacy concerns, it encounters threats in handling heterogeneous and dynamic data distributions across diverse institutions. Additionally, many models lack sufficient transparency, which affects user trust and regulatory compliance, which is significant for deployment in financial systems.

Research gap while addressing it:

- FL models often neglect the communication overhead and synchronization issues that arise in real-world networks, limiting their effectiveness and applicability in large-scale, multi-institution collaborations.
- Graph-based models such as GNNs and graph transformers are efficient in capturing associations but tend to be sensitive to noisy or incomplete data and are less adaptive to rapidly growing fraud patterns.
- The incorporation of BC and AI for fraud detection is capable but lacks a thorough evaluation of real-time scalability and integration threats in operational financial environments.
- Explainable AI methods improve transparency but are inconsistently applied across diverse fraud detection frameworks, restricting their usefulness for auditors and compliance officers who need clear, interpretable explanations.
- Most existing studies concentrate on a limited set of transaction types and fail to generalize across various financial instruments, which restricts the robustness of fraud detection systems in diverse markets.

3. System Design and Techniques

This study proposes an FTFD-DRXAIA model for financial fraud detection. The study intends to advance an efficient and intelligent method for precise fraud detection in financial transactions employing progressive DL methodologies. It comprises various processes involved in data pre-processing, feature selection, fraud detection, parameter tuning, and identifying suspicious transaction patterns. Fig. 1 epitomizes the overall procedure of the FTFD-DRXAIA technique.

DOI: https://doi.org/10.54216/IJNS.270202

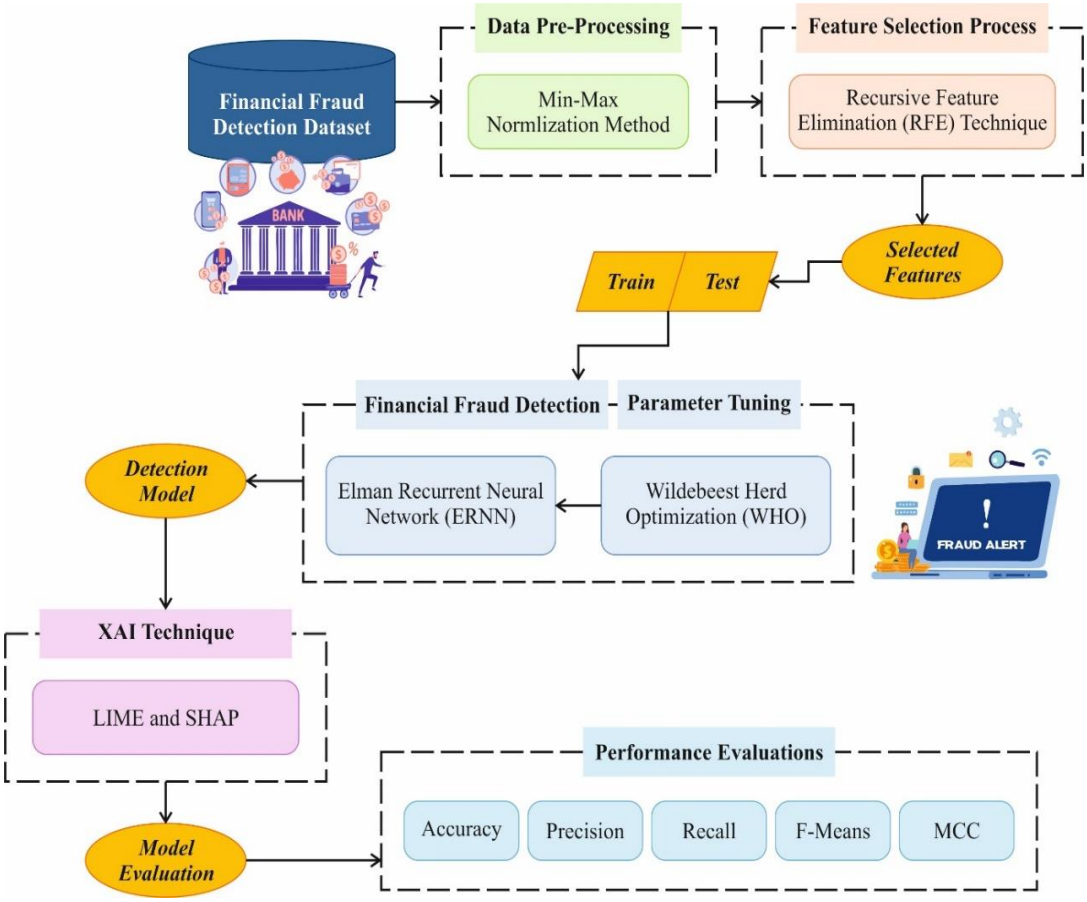

Figure 1. Overall flow of proposed method

A. Min-max Method

Primarily, the data pre-processing utilizes the min-max method to transfer raw data into appropriate patterns [20]. This method is selected for its competence in scaling data to a secure range, frequently between zero and one, preserving the original data. It is computationally effectual and easy to device, making it ideal for larger-scale financial database. This technique maintains the relative relationships among values, which benefits algorithms sensitive to feature magnitude, unlike normalization. It also helps in developing convergence speed and stability of various ML models.

A systematic Min-Max normalization model is used to equalize the range of values through each feature in the dataset, which guarantees that every feature contributes correspondingly to the training of the model. Therefore, the possibility of feature dominance is reduced, while features with greater numerical scales would be excessively affected by the model's learning.

$$x_{norm} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{1}$$

Whereas x signifies the new value of the attribute, $x_{\min}$ and $x_{\max}$ represent the minimum and maximum values of the attribute throughout the dataset. In this process, all features were normalized to the interval $[0,1]$. This model allowed us to support consistency through feature scales, which is essential for the stability and convergence of these DL methods.

B. RFE-based Feature Reduction Model

Moreover, mainly the RFE model to choose the optimum features from the dataset [21] implements the process of FS. This model is selected for its efficiency in iteratively removing less crucial features based on model performance, resulting in a refined and relevant feature subset. This model considers feature interactions by utilizing a predictive model, thus enhancing accuracy and mitigating overfitting, unlike simple filter models. This model is valuable when dealing with high-dimensional data, as it systematically detects the most informative

DOI: https://doi.org/10.54216/IJNS.270202

features. Compared to embedded methods, RFE presents more control over the feature selection process and can be incorporated with diverse estimators. The robust model interpretability and computational efficiency make it highly efficient for optimizing performance in complex datasets. Fig. 2 indicates the RFE method.

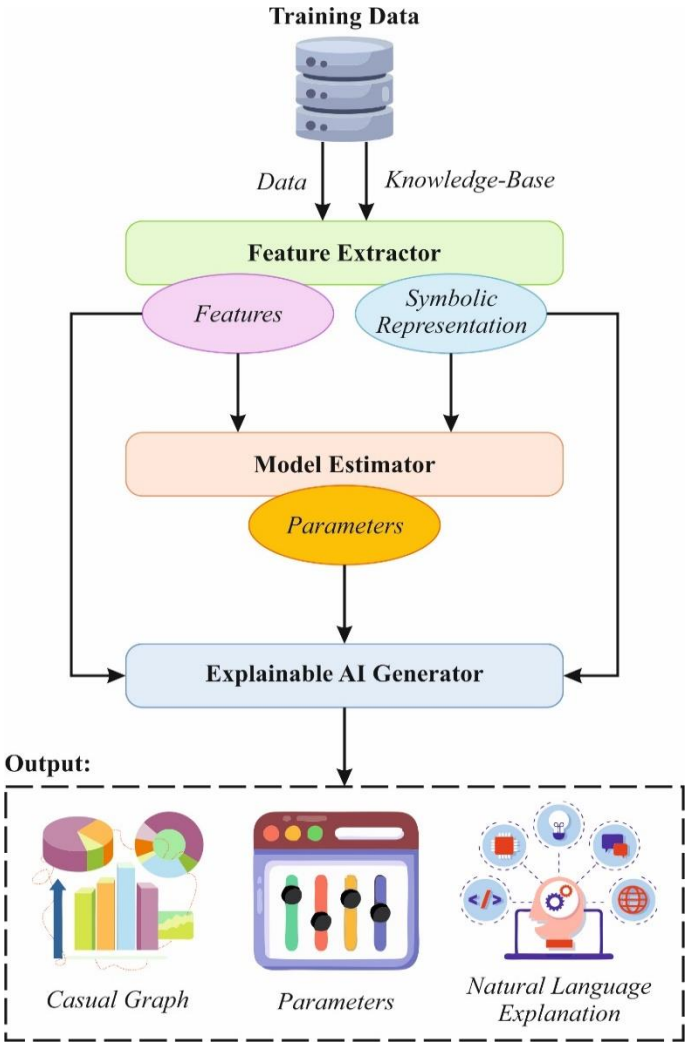

Figure 2. Architecture of the RFE method

RFE is the most generally employed technique for the selection of features. It functions by penetrating a feature subset starting with every feature in the training database and effectively eliminating attributes until the preferred amount is kept. It is attained by appropriating an assumed algorithmic method and next positioning the feature by significance to remove the smallest significant features and improve the technique. The selection procedure is carried out until an exact number of features is kept, and its exact steps are given as follows:

Algorithm of RFE
S1: Importing the training set with each input feature X_0 and output label Y_0. Set feature subsets $S = [1,2, \dots, m]$ and the positions of feature $R = [\,]$
S2: Training a method according to the set of training and compute the prognostic precision.
S3: Compute the feature significance and position them. Please find out the smallest significant feature f and reject it from the subset of feature $S = S - S(f)$. Upgrade the position of feature $R = [S(f), R]$.
S4: Do the 2nd and 3rd steps until the feature subset S is null.
S5: Outputting the position of feature R.

DOI: https://doi.org/10.54216/IJNS.270202

C. Financial Fraud Detection using ERNN Method

The ERNN approach is used for financial fraud detection [22]. This method has been chosen for its robust capacity to model temporal and consecutive needs in transaction data. Compared to other RNN alternates, ERNN existing a simpler manner with faster training but preserving effectual learning of temporal features. Its capability to capture dynamic patterns rises detection accuracy and diminishes false positives, making it an ideal decision for fraud detection.

The ERNN is chosen because of its capability to take contextual relationships and temporal dependences in the consecutive data. When equated to easier methods, ERNN is more appropriate for perceiving basic, non-linear patterns in higher-dimensional data. Besides, ERNN helps in mitigating overfitting by providing dynamic memory devices that permits the method to focus on related spatial attributes. When equated to CNN or RNN, ERNN gives an enhanced accuracy as well as performance in responsibilities by taking rich and context-aware representations.

ERNN is one of the famous ML techniques that might take off dynamic models owing to feedback links between nodes. This link of feedback might be employed to note the timing data for input and output methods. Every ERNN layer covers neurons, which use weights according to the comprehensive amount of input items, create a non-linear function, and convey that information to the next layer. The input layer formulation is given below:

$$X_{it}(k) = \sum_{i=1}^{n} X_{it}\,(k-1) \tag{2}$$

X_{it} denotes an input at $the\ tth$ time utilizing n neuron count. The hidden layer (HL) was described as follows.

$$ner_{jt}(k) = \Sigma_{i=1}^{n} W_{ij} X_{it}(k-1) + \Sigma_{j=1}^{p} C_j r_{jt} \tag{3}$$

W_{ij} represents the weight of the connection between the input layer and HL, $and\ C_j$ describes the weight of the connection among the recurrent layer and HL. The output of HL is shown below:

$$Z_{jt}(k) = f(net_{jk}(k) = \Sigma_{i=1}^{n} W_j x_{it}(k-1) + \Sigma_{j=1}^{p} C_j R_{jt}(k) \tag{4}$$

The recurrent layer is computed below:

$$R_{jt}(k) = Z_{jt}(k-1) \tag{5}$$

The output layer is analyzed as established:

$$Y_t(k) = f(\Sigma_{j=1}^{p} V_j Z_{jt}(k) \tag{6}$$

D. WHO-based Hyperparameter Tuning Technique

Then, the WHO technique fine-tunes the parameter outcomes of the ERNN classifier, resulting in enhanced performance of classification [23]. This model is selected as it effectively balances exploration and exploitation, inspired by natural herd behaviours. This approach outperforms in avoiding local optima and converging faster towards global solutions, compared to conventional optimization methods. Its population-based approach enables simultaneous evaluation of multiple candidate solutions, enhancing search diversity and robustness. The adaptability and simplicity also make it appropriate for tuning intrinsic methods such as ERNN, improving classification accuracy and stability overall. WHO provides an effective and reliable method for optimizing hyperparameters, outperforming many conventional techniques in terms of convergence speed and solution quality.

The WHO model is applied for parameter tuning to confirm that the top parameters are chosen for heightened precision. The hyperparameters of the projected structure are enhanced by utilizing the WHO model. The vital justification for selecting this model is that it compares with other types of metaheuristic models by using this metaheuristic model to increase the efficacy of the projected methodologies. The Wildebeests' habit acts as a WHO methodology. Wildebeests are sociable, active animals that hunt for food resources. Males compete with sex tasks to appeal to females in the mating process. The population is arbitrarily initialized as candidates at the beginning of the WHO model. The lesser $(X_{\min})$ and higher $(X_{\max})$ borders restrict the population.

$$X_i \in [X_{\min}, X_{\max}] \tag{7}$$

Where, $I = 1,2 \ldots N$.

Then, the wildebeest utilized the milling model of movement. Continue to search for the optimum place, although deliberating a fixed number (n) at random mobility. The competitor in place X had utilized an arbitrary stage Z_n routinely searching for random stages. The length is modifiable and dependent on the dimension of the contestants' arbitrary steps. Therefore, the localized stage Z_n is created utilizing the following equation:

DOI: https://doi.org/10.54216/IJNS.270202

$$Z_n = X_i + \varepsilon \times \theta \times v \tag{8}$$

Now, arbitrary uniform values between zero and one are depicted by θ, an arbitrary random unit vector is denoted $by\ v$, the rate of learning is shown as ε, and the i-th candidate number can signify X_i. Once assessing a fixed number (n), the wildebeest modified its position to attain a perfect arbitrary location.

$$X_i = \alpha_1 \times Z_n^* + \beta_1 \times (X_i - Z_n^*) \tag{9}$$

Now, the local drive of candidates is trained by α_1 and β_1 leader variables. Modelling the behaviour of a wildebeest swarm is the final stage. It is simulated once another contestant is in a spot with proper food resources.

$$X_i = \alpha_2 \times X_i + \beta_2 \times X_h \tag{10}$$

Where X_h represents an arbitrary candidate and α_2 and β_2 indicate leader variables for directing the local movement of the crew.

$$X_i = X_i + \theta \times (X_{\max} - X_{\min}) \times V \tag{11}$$

Now, the randomly generated unit vector is represented as V. Pretending busy regions is another phrase employed in the model. At the same time, the grassland has huge efficiency. To accomplish a task, the finest contender employs the succeeding Eq. (12) for destroying other contenders.

$$if\,(\|X^* - X_i\|) < \eta, (\|X^* - X_i\|) > 1 \tag{12}$$

$$X_i = X^* + \varepsilon \times \hat{n} \tag{13}$$

Here $\hat{n}$ represents available section counts near points of ideal solution, and η depicts a threshold to preclude jamming in the site. The social memory of the swarm pretends in the final phase to offer better placements.

$$X = X^* + 0.1 \times \hat{v} \tag{14}$$

Table 1 specifies the hyperparameters of the WHO model and highlights the strategic setup required for achieving optimal tuning of the ERNN model. Each parameter within the algorithm is carefully adjusted for balancing the global search capability and local refinement, allowing the method to navigate intrinsic solution landscapes effectively. The tuning setup ensures that the optimization process avoids premature convergence and maintains diversity among candidate solutions, which is significant in handling dynamic and imbalanced data like financial fraud. The configuration reflects a trade-off between computational cost and convergence speed, tailored particularly for enhancing the robustness and prediction accuracy in real-world financial datasets.

Table 1: Hyperparameters used in WHO-based optimization

Parameter	Description	Value / Range
POPU_SIZE	Wildebeest count (solutions) in the herd	30-100
MAX_ITER	Number of optimization iterations (generations)	50-200
TRANSFER_COEFFICIENT	Regulates the exploration and exploitation transition	Linearly increases [0.1–1]
ACCELERATION_FACTOR	Determines movement towards optimal solutions	0.1 – 1.0
COLLISION_COEFFICIENT	Handles avoidance behaviour during local search	0.1 – 0.9
LOW_BOUND	Minimum limit for the hyperparameter search space	Model parameter dependent
UPP_BOUND	Maximum limit for the hyperparameter search space	Model parameter dependent
DIMENSIONALITY	Hyperparameters to be tuned	Relies on the target model

DOI: https://doi.org/10.54216/IJNS.270202

Finally, the finest optimum value of this model is set to the classifier parameters. The WHO model invents a fitness function (FF) to attain an amended classifier outcome. It describes a positive value to signify the enhanced candidate solution. The classifier error reduction was evaluated as FF, as provided in Eq. (15).

$$fitness(x_i) = ClassifierErrorRate(x_i) = \frac{no.of\ misclassified\ instances}{Total\ no.of\ instances} \times 100 \quad (15)$$

E. XAI Model using LIME and SHAP

Finally, the XAI model applies LIME, and SHAP are employed to understand composite AI methods. The dual primary explainability-related methods employed were LIME and SHAP [24]. Both techniques function on a diverse level and tackle other issues of explainability; therefore, they should be used in integration to enhance stability. At first, the LIME functions on a distinct point, and it clarifies a single assessment by forming a local estimate of the method near that exact example. It adjusts an input in a small quantity and detects the fault produced by this modification while constructing a simple and more interpretable technique (linear regression method). The primary standard is to demonstrate an original method by utilizing an easy form for an exact forecast. Therefore, it clarifies why the technique created a precise assessment of this data point. From the calculated viewpoint, LIME intends to estimate a method f, close to a particular example $x \in R^d$ utilizing a local surrogated method $g \in G$, G signifying an interpretable method. Here, the LIME addresses the succeeding optimizer issue:

$$\underset{g \in G}{\text{argminL}}(f, g, \pi_x), +\Omega(g) \quad (16)$$

While f is a new model, g denotes an interpretable surrogated method; $\pi_x(z)$ represents a proximity measure among the perturbed sample z and instance x ($z \in Z$); $L(f, g, \pi_x)$ specifies a function of local fidelity loss, $\Omega(g)$ means complexity penalties of g.

Whereas SHAP functions on a global level by seeing every prediction completely. It uses the model itself, along with the last forecasts, to evaluate the Shapley values that were employed for producing those predictions. The values of Shapley were initially presented to equally distribute complete rewards amongst player in the supportive game according to their distinct contributions. The main indication behind SHAP is that every feature is assessed by its influence on each probable feature subset. A group of $\mathcal{F}$ features and model functions f, the Shapley values for an exact feature i were

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!\,(|F| - |S| - 1)!}{|F|!} [f(S \cup \{i\} - f(S)] \quad (17)$$

Here, S denotes a feature's subset, $f(S)$ signifies the output of the model in S, and $|F|$ represents the total no. of attributes. Within the sum, the fraction indicates a weighting factor. Whereas the sub-set S seems right earlier, including the feature i. So, these fraction supports the term goal of weighting and guarantees that all contributors is equally weighted without giving more significance to exact features accidentally.

4. Model Assessment and Results

The performance analysis of the FTFD-DRXAIA model is inspected in terms of the financial fraud detection database [25]. The database was artificially created using the PaySim simulation tool, which replicates mobile financial transactions based on a month's worth of actual transaction records from a mobile payment provider in an African nation. Data confidentiality is maintained through authentic transaction and a global corporation operating in 14 different countries shares the records. The dataset spans 30 days with 744 hourly steps and incorporates multiple transaction types, including fraudulent activities, for research purposes. Ethical considerations were addressed by anonymizing data and nullifying fraudulent transaction balances to protect user privacy while enabling fraud detection studies. It comprises 8000 samples in terms of 2 classes, like isFraud_Yes and isFraud_No, with each class having 4000 instances, as shown below in Table 2. The no. of features is 11, but only eight features are chosen.

DOI: https://doi.org/10.54216/IJNS.270202

Table 2: Details of the database

Classes	No. of Instances
"isFraud_Yes"	4000
"isFraud_No"	4000
Total Instances	8000

Fig. 3 describes the correlation matrix made by the FTFD-DRXAIA model. The outputs indicate that the FTFD-DRXAIA methodology accurately identifies and detects both classes.

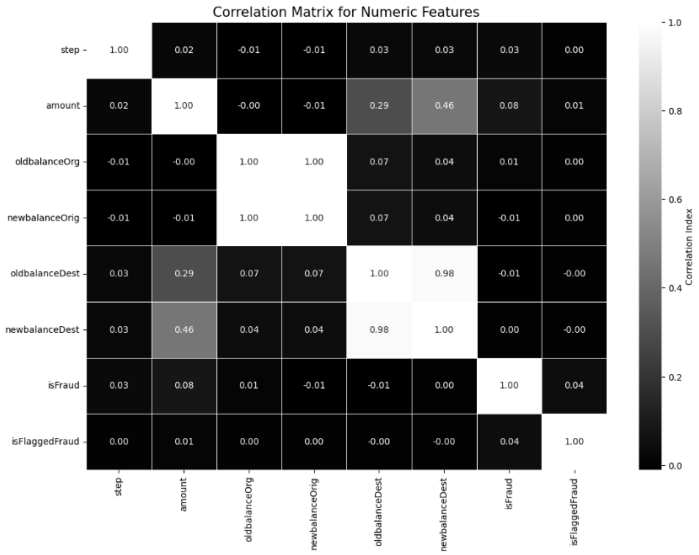

Figure 3. Correlation matrix of the FTFD-DRXAIA model for numeric features

Fig. 4 illustrates the amount of every kind of transaction, which signifies the distribution of dissimilar types of financial transactions. The central part corresponds to CASH_OUT transactions, which make up 35.2% of the total. At the same time, the PAYMENT class comprises 33.8%. CASH_IN transactions explain 22.0%, representing the part of funds placed or added to accounts. TRANSFER transactions make up 8.4%, presenting moderate usage of funds being moved between accounts. At last, DEBIT transactions accounted for only 0.7%, indicating the lowest usage compared with other forms.

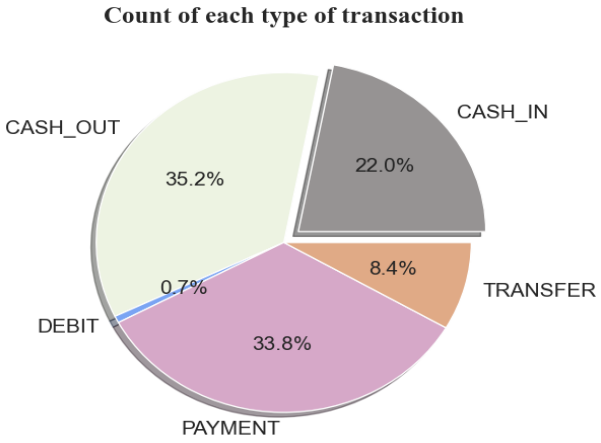

Figure 4. Count of each type of transaction

DOI: https://doi.org/10.54216/IJNS.270202

Fig. 5 presents the classifier results of the FTFD-DRXAIA technique. Figs. 5a-5c display the confusion matrices at 70:30. Fig. 5b shows the PR examination, signifying maximal performance through each class. Finally, Fig. 5d describes the ROC inspection, determining efficient outputs with better values of ROC for dissimilar classes.

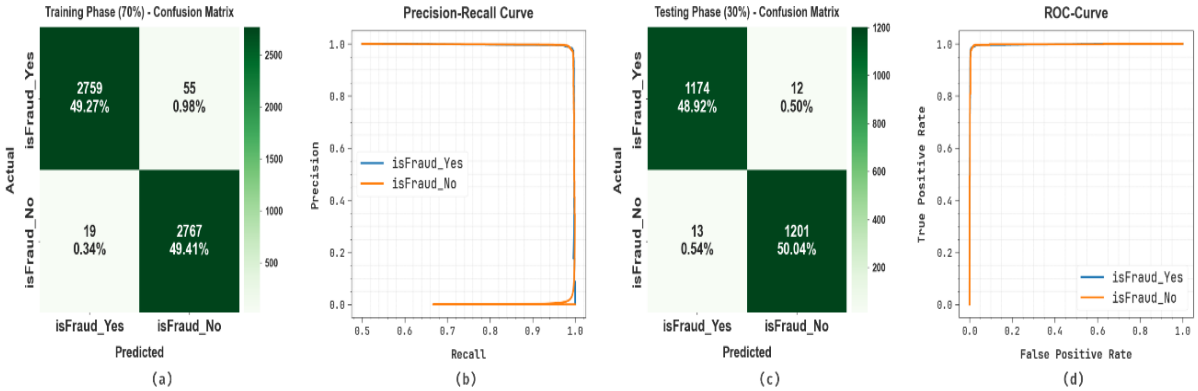

Figure 5. Classifier outcome (a-c) 70:30 confusion matrices and (b-d) curves of PR and ROC

Table 3 and Fig. 6 portray the financial fraud detection of FTFD-DRXAIA methodology in terms of 70:30. On 70% TRPHE, the FTFD-DRXAIA technique attains an average $accu_y$ of 98.68%, $prec_n$ of 98.68%, $reca_l$ of 98.68%, F_{Means} of 98.68%, and MCC of 97.37%. Furthermore, on 30% TSPHE, the FTFD-DRXAIA technique reaches an average $accu_y$ of 98.96%, $prec_n$ of 98.96%, $reca_l$ of 98.96%, F_{Means} of 98.96%, and MCC of 97.92%.

Table 3: Financial fraud detection of FTFD-DRXAIA model at 70:30

Classes	$Accu_y$	$Prec_n$	$Reca_l$	F_{Means}	MCC
TRPHE (70%)					
isFraud_Yes	98.05	99.32	98.05	98.68	97.37
isFraud_No	99.32	98.05	99.32	98.68	97.37
Average	98.68	98.68	98.68	98.68	97.37
TSPHE (30%)					
isFraud_Yes	98.99	98.90	98.99	98.95	97.92
isFraud_No	98.93	99.01	98.93	98.97	97.92
Average	98.96	98.96	98.96	98.96	97.92

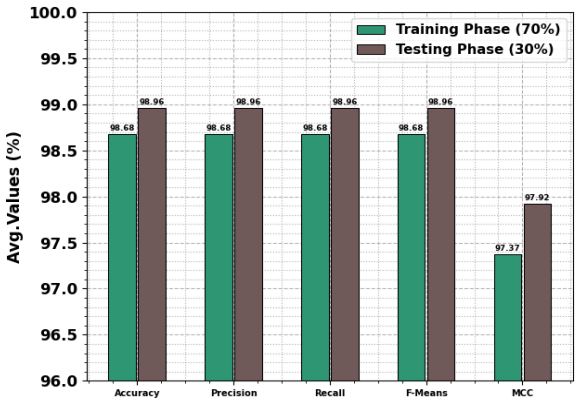

Figure 6. Average values of FTFD-DRXAIA model (a) 70% and (b) 30%

DOI: https://doi.org/10.54216/IJNS.270202

Fig. 7 exemplified the training (TRAIN) $accu_y$ and validation (VALID) $accu_y$ of an FTFD-DRXAIA method over **50 epochs**. At the initial stage, either TRAIN or VALID $accu_y$ rises quickly, indicating effective learning of data patterns. Around a certain epoch, VALID $accu_y$ slightly surpasses TRAIN, showing robust generalization without overfitting. As training progresses, the close alignment of both curves suggests the model is well regularized and capable of retaining valuable features across both seen and unseen data.

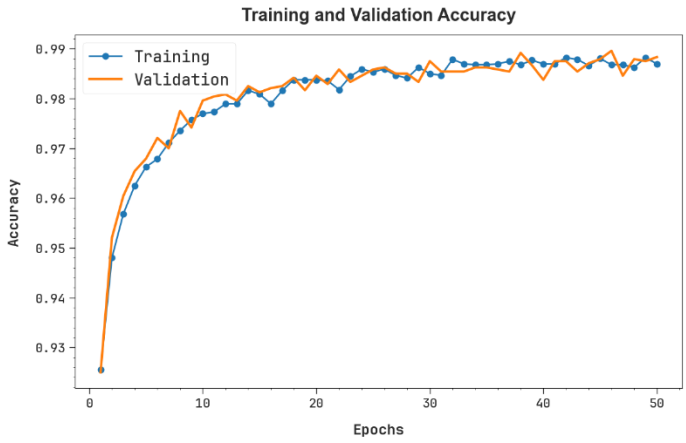

Figure 7. $Accu_y$ Curve of FTFD-DRXAIA model

Fig. 8 portrays the TRAIN and VALID losses of the FTFD-DRXAIA approach over 50 epochs. In the early stages, either TRAIN or VALID loss is lower, exhibiting an initial grasp of the data. As training progresses, both losses decrease steadily, reflecting effective learning. The close alignment of TRAIN and VALID loss curves indicates robust generalization and minimal overfitting.

Figure 8. Loss curve of the FTFD-DRXAIA model

The comparison analysis of the FTFD-DRXAIA method with other recent techniques is provided in Table 4 and Fig. 9 [26-28]. According to $accu_y$, the FTFD-DRXAIA methodology has maximum $accu_y$ of 98.96% while the logistic regression (LR), decision tree (DT), neural network (NN), autoencoder (AE), naïve bayes (NB), GCN, and SVM techniques have the least $accu_y$ of 95.20%, 96.50%, 98.30%, 97.56%, 95.11%, 92.90%, and 96.23%.

DOI: https://doi.org/10.54216/IJNS.270202

Table 4: Comparative analysis of FTFD-DRXAIA model with existing technique

Models	$Accu_y$	$Prec_n$	$Reca_l$	F_{Means}
LR	95.20	90.53	98.45	94.96
DT	96.50	98.00	93.51	95.77
NN	98.30	91.84	92.63	93.11
AE	97.56	91.30	92.29	91.51
NB	95.11	94.14	93.37	92.38
GCN	92.90	98.20	90.41	91.11
SVM	96.23	95.98	90.45	95.84
FTFD-DRXAIA	98.96	98.96	98.96	98.96

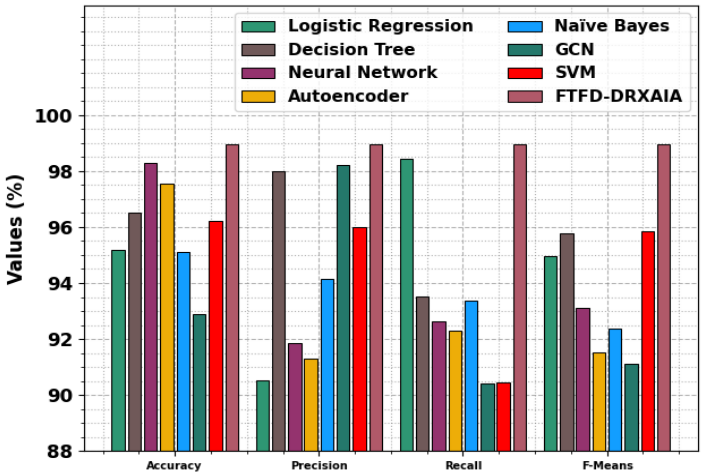

Figure 9. Comparative analysis of the FTFD-DRXAIA model with the existing technique

Finally, depending on F_{Means}, the FTFD-DRXAIA model has a greater F_{Means} of 98.96% whereas the LR, DT, NN, AE, NB, GCN, and SVM techniques have a minimum F_{Means} of 94.96%, 95.77%, 93.11%, 91.51%, 92.38%, 91.11%, and 95.84%, respectively.

Table 5 and Fig. 10 specifies the computational time (CT) study of the FTFD-DRXAIA approach with recent models. The -DRXAIA approach achieved the fastest CT of 4.21 seconds, making it the most efficient among all evaluated models. Compared to LR with 14.07 seconds and SVM with 13.19 seconds, the FTFD-DRXAIA model reduces CT by approximately 70% and 68%, respectively. Similarly, it is significantly faster than NN at 6.31 seconds and DTs at 7.85 seconds. This substantial loss in CT highpoints the FTFD-DRXAIA capability of the model to deliver rapid results, which is highly beneficial in real-time B2B decision-making environments.

DOI: https://doi.org/10.54216/IJNS.270202

Table 5: CT evaluation of FTFD-DRXAIA approach with existing models

Models	**CT (sec)**
LR	14.07
DT	7.85
NN	6.31
AE	10.62
NB	9.44
GCN	9.54
SVM	13.19
FTFD-DRXAIA	4.21

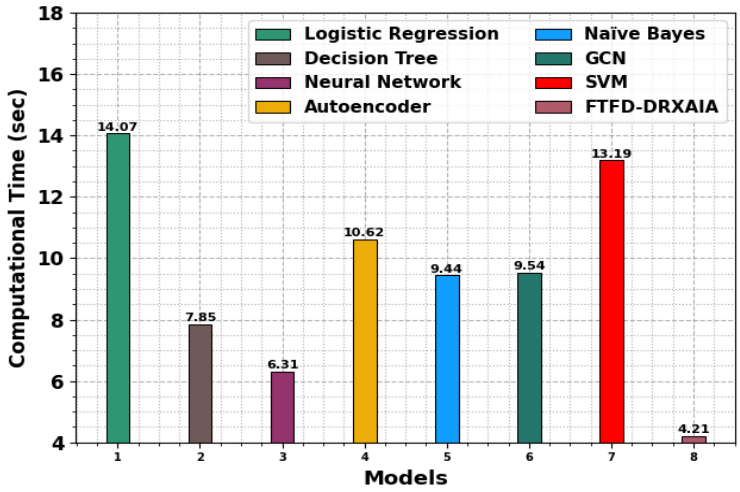

Figure 10. CT evaluation of FTFD-DRXAIA approach with existing models

5. Conclusion

This article proposes the FTFD-DRXAIA approach for financial fraud detection. The aim is to develop an intelligent and efficient system for precise fraud detection. Initially, the data pre-processing applies the min-max model to transform raw data into an appropriate format. Moreover, the RFE model is used for the FS process. The ERNN method is employed for financial fraud detection. The WHO technique fine-tunes the ERNN model's hyperparameters, resulting in enhanced classification performance. Finally, the XAI model applies LIME and SHAP to understand composite AI methods, allowing auditors and analysts to recognize suspicious transaction patterns with greater confidence and clarity. The experimental analysis of the FTFD-DRXAIA approach is inspected regarding the financial fraud detection database. The comparison study of the FTFD-DRXAIA approach showed a higher precision value of 98.96% over recent techniques. The limits of the FTFD-DRXAIA method contain the dependence on a synthetic database that might not capture all complexities of real-world financial transactions, potentially affecting generalizability. Furthermore, the performance across diverse geographic regions and varying transaction behaviours remains unexamined. The computational cost and resource needs for large-scale deployment could also pose threats in resource-constrained environments. Future work may concentrate on integrating multi-source real transaction data, enhancing model adaptability to growing fraud patterns, and optimizing algorithms for faster processing. Integrating real-time streaming data and improving interpretability for non-technical users will also be explored. The FTFD-DRXAIA method depicts strong potential for real-world use in financial institution to successfully identify and stop fraud in dynamic transaction sceneries.

Funding: "This research received no external funding"

Conflicts of Interest: "The authors declare no conflict of interest."

DOI: https://doi.org/10.54216/IJNS.270202

References

[1] A. M. Smith and L. J. Johnson, "A novel approach to financial fraud detection using ensemble learning techniques," *Journal of Financial Technology*, vol. 6, no. 2, pp. 123–134, 2023.

[2] C. W. Lee, M. W. Fu, C. C. Wang, and M. I. Azis, "Evaluating machine learning algorithms for financial fraud detection: Insights from Indonesia," *Mathematics*, vol. 13, no. 4, p. 600, 2025.

[3] E. M. Al-dahasi, R. K. Alsheikh, F. A. Khan, and G. Jeon, "Optimizing fraud detection in financial transactions with machine learning and imbalance mitigation," *Expert Systems*, vol. 42, no. 2, p. e13682, 2025.

[4] C. V. Sai, D. Das, N. Elmitwally, O. Elezaj, and M. B. Islam, "Explainable AI-driven financial transaction fraud detection using machine learning and deep neural networks," *SSRN*, 2023. [Online]. Available: https://ssrn.com/abstract=4439980

[5] S. S. Taher, S. Y. Ameen, and J. A. Ahmed, "Advanced fraud detection in blockchain transactions: An ensemble learning and explainable AI approach," *Engineering Technology and Applied Science Research*, vol. 14, no. 1, pp. 12822–12830, 2024.

[6] S. Ahmadi, "Advancing fraud detection in banking: Real-time applications of explainable AI (XAI)," *Journal of Electrical Systems*, vol. 18, no. 4, pp. 141–150, 2022.

[7] M. R. Hasan, M. S. Gazi, and N. Gurung, "Explainable AI in credit card fraud detection: Interpretable models and transparent decision-making for enhanced trust and compliance in the USA," *Journal of Computer Science and Technology Studies*, vol. 6, no. 2, pp. 1–12, 2024.

[8] D. Cirqueira, M. Helfert, and M. Bezbradica, "Towards design principles for user-centric explainable AI in fraud detection," in *Proc. Int. Conf. Human-Computer Interaction*, Cham, Switzerland: Springer, Jul. 2021, pp. 21–40.

[9] P. Thanathamathee, S. Sawangarreerak, S. Chantamunee, and D. N. M. Nizam, "SHAP-instance weighted and anchor explainable AI: Enhancing XGBoost for financial fraud detection," *Emerging Science Journal*, vol. 8, pp. 2404–2430, 2024.

[10] K. Vuppula, "An advanced machine learning algorithm for fraud financial transaction detection," *Journal of Innovative Development in Pharmaceutical Technology*, vol. 4, no. 9, 2021.

[11] O. A. Bello, A. Ogundipe, D. Mohammed, F. Adebola, and O. A. Alonge, "AI-driven approaches for real-time fraud detection in US financial transactions: Challenges and opportunities," *European Journal of Computer Science and Information Technology*, vol. 11, no. 6, pp. 84–102, 2023.

[12] A. Ali *et al.*, "Financial fraud detection based on machine learning: A systematic literature review," *Applied Sciences*, vol. 12, no. 19, p. 9637, 2022.

[13] M. Rahmati, "Real-time financial fraud detection using adaptive graph neural networks and federated learning," *International Journal of Management Data Analytics*, vol. 5, no. 1, pp. 98–110, 2025.

[14] M. T. R. Mazumder, M. S. H. Shourov, I. Rasul, S. Akter, and M. K. Miah, "Anomaly detection in financial transactions using convolutional neural networks," *Journal of Economics, Finance and Accounting Studies*, vol. 7, no. 2, pp. 195–207, 2025.

[15] M. Kasedullah *et al.*, "Energizing blockchain and artificial intelligence for enhanced fraud detection in modern banking systems," 2025.

[16] T. Awosika, R. M. Shukla, and B. Pranggono, "Transparency and privacy: The role of explainable AI and federated learning in financial fraud detection," *IEEE Access*, vol. 12, pp. 64551–64560, 2024.

[17] K. Li *et al.*, "Sefraud: Graph-based self-explainable fraud detection via interpretative mask learning," in *Proc. 30th ACM SIGKDD Conf. Knowledge Discovery and Data Mining*, 2024, pp. 5329–5338.

[18] N. G. Castellanos, "Revolutionizing management with AI and blockchain for smarter anomaly detection and fraud prevention," *Journal of Computer Innovations and Applications*, vol. 1, no. 1, pp. 10–18, 2023.

[19] A. Yazdinejad, A. Dehghantanha, R. M. Parizi, G. Srivastava, and H. Karimipour, "Secure intelligent fuzzy blockchain framework: Effective threat detection in IoT networks," *Computers in Industry*, vol. 144, p. 103801, 2023.

[20] M. J. Hossain, K. Alam, M. F. Monir, M. M. Hoque, and T. Ahmed, "Explainable AI meets synthetic data: A deep learning framework for detecting network intrusion in NextG network infrastructure," *IEEE Access*, 2025.

[21] W. Liu, Y. Suzuki, and S. Du, "Ensemble learning algorithms based on easyensemble sampling for financial distress prediction," *Annals of Operations Research*, pp. 1–32, 2025.

[22] S. Nagadevi *et al.*, "An efficient privacy-preserving multilevel fusion-based feature engineering framework for UAV-enabled land cover classification using remote sensing images," *Scientific Reports*, vol. 15, no. 1, p. 23707, 2025.

[23] S. S. R. Bairaboina and S. R. Battula, "Ghost-ResNeXt: An effective deep learning based on mature and immature WBC classification," *Applied Sciences*, vol. 13, no. 6, p. 4054, 2023.

DOI: https://doi.org/10.54216/IJNS.270202

[24] T. Tasioulis, E. Bagkis, T. Kassandros, and K. Karatzas, "The quest for the best explanation: Comparing models and XAI methods in air quality modeling tasks," *Applied Sciences*, vol. 15, no. 13, p. 7390, 2025.

[25] Kaggle, "Financial fraud detection dataset," [Online]. Available: https://www.kaggle.com/datasets/sriharshaeedala/financial-fraud-detection-dataset, [Accessed: Aug. 13, 2025].

[26] P. Sundaravadivel *et al.*, "Optimizing credit card fraud detection with random forests and SMOTE," *Scientific Reports*, vol. 15, no. 1, p. 17851, 2025.

[27] L. H. Aros, L. X. B. Molano, F. Gutierrez-Portela, J. J. M. Hernandez, and M. S. R. Barrero, "Financial fraud detection through the application of machine learning techniques: A literature review," *Humanities and Social Sciences Communications*, vol. 11, no. 1, pp. 1–22, 2024.

[28] A. Asiri and K. Somasundaram, "Graph convolution network for fraud detection in bitcoin transactions," *Scientific Reports*, vol. 15, no. 1, p. 11076, 2025.

DOI: https://doi.org/10.54216/IJNS.270202