# An Intelligent Decision Support Systems for Financial Fraud Detection Using Pythagorean Neutrosophic Bonferroni Mean Approach with Machine Learning Models

**Umidjon Matyakubov[1], Ranokhon Sharofutdinova[2], Aleksey Ilyin[3], Rustem Shichiyakh[4], K. Shankar[5], E. Laxmi Lydia[6,*]**

[1]Department of Economics, Mamun University, Khiva, 220900, Uzbekistan

[2]Department of Social and Humanitarian Sciences, Fergana State University, Fergana, 150100, Uzbekistan

[3]Kursk Branch, Financial University under the Government of the Russian Federation, Moscow, 125167, Russia

[4]Department of Management, Kuban State Agrarian University named after I.T. Trubilin, Krasnodar, 350044, Russia

[5]Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, 602105, India

[6]Department of Information Technology, VR Siddhartha Engineering College (A), Siddhartha Academy of Higher Education (Deemed to be University), Vijayawada, India

Emails: matyakubov_umidjon@mamunedu.uz; ranoxon3511@gmail.com; aeilin@fa.ru; shichiyax.r@kubsau.ru; drkshankar@ieee.org; elaxmi2002@yahoo.com

## Abstract

Neutrosophy has developed as a generalization to fuzzy logic and is being employed in the research field in many areas such as set theory, logic, and others. Neutrosophic Logic is one of the neonate study regions and its intention is assessed to have the percentage of truth in a subset T, the percentage of falsity in a subset F, and the percentage of indeterminacy in a subset I. Recently, financial fraud has become a highly major issue, which results in severe consequences across firm sectors and affects people's everyday lives. Therefore, financial fraud recognition is critical for the prevention of the regularly overwhelming effects of financial fraud. It includes differentiating fraudulent financial data from accurate data and permitting decision-makers to progress suitable plans to reduce the effect of fraud. Over the past few years, Artificial intelligence (AI), mainly machine learning (ML) systems, turned out to be the highest thriving model in fraud detection. This study presents a novel Intelligent Decision Support System for Financial Fraud Detection Using Pythagorean Neutrosophic Bonferroni Mean (IDSSFFD-PNBM) model. The main intention of the IDSSFFD-PNBM algorithm is to enrich the detection model for financial fraud using advanced optimization models. Initially, the z-score normalization is applied in the data normalization stage for converting input data into a beneficial format. Besides, the proposed IDSSFFD-PNBM designs a grasshopper optimization algorithm (GOA) for the selection of feature processes to enhance the efficiency and performance of the model. For the detection and classification procedure, the pythagorean neutrosophic bonferroni mean (PNBM) model has been employed. Additionally, the firefly optimization algorithm (FFOA)-based hyperparameter range method has been done to heighten the recognition outcomes of the PNBM system. The experimental evaluation of the IDSSFFD-PNBM technique takes place using a benchmark dataset. The experimental results indicated an enhanced performance of the IDSSFFD-PNBM technique compared to recent approaches

**Keywords:** Neutrosophic Logic; Financial Fraud Detection; Fuzzy Logic; Pythagorean Neutrosophic Bonferroni Mean; Machine Learning



DOI: https://doi.org/10.54216/IJNS.250418

## 1. Introduction

A highly effective tool for managing uncertainty in decision-making is the Neutrosophic Set (NS) [1]. An effective device to represent vagueness and uncertainty in decision-making is the NS that are classical set generality, intuitionistic fuzzy set (IFS), and fuzzy set by increasing 3 classes of falsehood, truth, and uncertainty of a confirmed statement [2]. It employs several decision-making procedures. However, to modify NS with more real complicated cases, INS and CNS are projected [3]. Financial fraud is a widespread issue with far-reaching implications for both the finance industry and everyday life. Fraud can lessen confidence destabilize economies, and industries, and affect individual expenses [4]. Recently, financial fraud activities like money laundering, and credit card fraud, have increased deliberately. These actions cause the loss of personal and/or enterprise assets [5]. Nevertheless, financial fraud recognition is not an easy task owing to the complicated trading systems and transactions involved. For instance, money laundering is described as the process of utilizing trades to transfer goods or money with the intent of obscuring the true source of funds [6]. Generally, the prices, quality, or quantity of goods on an invoice of money laundering is fake deliberately.

Despite multiple efforts to decrease financial fraudulent actions, it persists and affects society and the economy undesirably, huge volumes of money are lost to fraud day by day. Various fraud detection methods were presented a long time ago [7]. Financial fraud recognition is vital for preventing the frequently devastating financial fraud concerns [8]. It includes distinguishing fraudulent economic data from authentic data, consequently disclosing fraudulent activities or behavior and allowing decision-makers to advance proper approaches to reduce the effect of fraud. Most conventional approaches are manual, and it is not only expensive, inaccurate, and time-consuming, but also unfeasible [9]. Advancements in Machine Learning (ML), data mining, and Artificial Intelligence (AI) are leveraged to identify fraudulent activities within the economic sector. Both supervised and unsupervised approaches are applied to forecast fraud activities [10]. Classification models are widely used as a primary approach for identifying fraudulent financial transactions.

This study presents a novel Intelligent Decision Support System for Financial Fraud Detection Using Pythagorean Neutrosophic Bonferroni Mean (IDSSFFD-PNBM) model. Initially, the z-score normalization is applied in the data normalization stage for converting input data into a beneficial format. Besides, the proposed IDSSFFD-PNBM designs a grasshopper optimization algorithm (GOA) for the selection of feature processes. For the detection and classification procedure, the pythagorean neutrosophic bonferroni mean (PNBM) model has been employed. Additionally, the firefly optimization algorithm (FFOA)-based hyperparameter range method has been done to heighten the recognition outcomes of the PNBM system. The experimental evaluation of the IDSSFFD-PNBM technique takes place using a benchmark database.

## 2. Related Works

Alhabib et al. [11] developed a method to utilize an association of KNN and Random Forest (RF) techniques for credit card fraud detection, describing its methodology and architecture. This paper examined the outcomes, determining the precision, scalability, and robustness of the presented method. Ismail and Haq [12] developed a structure that applies ML models and data analytics to precisely recognize anomalies, patterns, and fraudulent activity signs. They utilized exploratory data analysis methods to recognize occurrences of imbalanced data and missing values. The choice of RF Classifier depends on its capability to dependably take intricate designs and effectively tackle multicollinearity problems. The isolation forest method produced higher precision. Ghanim and Awad [13] projected an innovative selection of models for unsupervised anomaly detection (AD) utilizing a grouping of reinforcement learning (RL) and time series forest (TSF) methods that dynamically select an AD model. This paper employs a real-world series database to represent the efficacy of the developed method.

In [14], an innovative method is introduced for fraud detection. This method utilizes the deep features extracted from CNN as inputs to several ML techniques, therefore substantially contributing to enhancing fraud detection efficiency and accuracy. Latipov et al. [15] developed an Optimizing Financial Fraud Detection utilizing BO and Variable Selection with Neutrosophic Vague Soft Set (OFFDBO-VSNVS) methodology. The GWO-based feature selection efficiently decreases dimensions. For the detection and classification of financial fraud, the NVS method is utilized. Finally, the BO technique alters the hyper-parameter values of the NVS model optimum and results in better classification results.

Theodorakopoulos et al. [16] introduce extensive research on employing ML models for real-world credit card fraud detection. This investigation estimates several ML techniques, for their efficiency in recognizing fraudulent actions. This work highlights the significance of real-world examination, managing imbalanced datasets, and adaptive learning. Ding et al. [17] developed a PSO-XGBoost fraud detection structure and utilized explainable AI to make the predictions. Classical approaches, comprising NB, SVM, BP Neural Network, and LR, determine modest precision.



DOI: https://doi.org/10.54216/IJNS.250418

## 3. The Proposed Method

In this study, we have presented a novel IDSSFFD-PNBM model. The main intention of the IDSSFFD-PNBM algorithm is to enrich the detection model for financial fraud using advanced optimization models. Fig. 1 represents the entire flow of the IDSSFFD-PNBM algorithm.

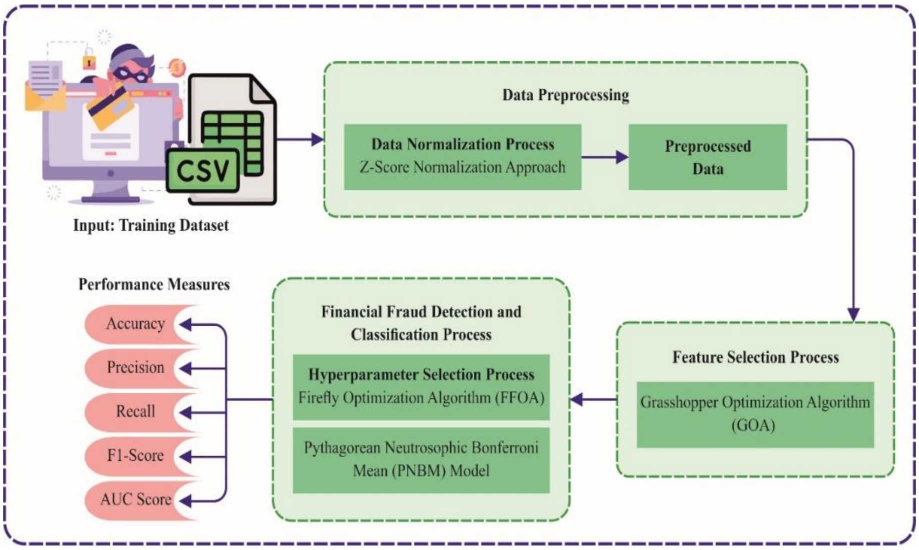


**Figure 1.** Overall flow of IDSSFFD-PNBM algorithm

### A. Z-score Normalization

Primarily, the z-score normalization is applied in the data normalization stage for converting input data into a beneficial format. Z-score normalization is a data pre-processing model, employed to normalize features by centering them over a standard deviation of one and a mean of zero [18]. This is chiefly beneficial in financial fraud recognition, where datasets frequently cover features with fluctuating units and scales. By regularizing the data, the Z-score certifies that all the features donates similarly to the method, averting supremacy by features with greater magnitudes. It aids models, particularly distance-based ones such as kNN or clustering, to do more precisely. In addition, this normalization helps in classifying anomalies, as great Z-scores might imply latent fraudulent activity. Overall, it improves the reliability and consistency of fraud recognition methods.

### B. GOA-based Feature Selection

Besides, the proposed IDSSFFD-PNBM designs GOA for the selection of feature processes to enhance the efficiency and performance of the model. The GOA is a population-based and nature-influenced meta-heuristic stimulated by the food resource searching features of grasshoppers naturally, like their distinguishing swarm behavior whereas moving and jumping in dissimilar vegetation's search for survival [19]. During this GOA modeling procedure, grasshoppers are established to look for food resources and are named $search\ agents$, however, the food positions are the optimal places for $search\ agents$ in the colony or swarm.

$$Y_i^c = D\left(\sum_{\substack{j=1 \\ j\neq 1}}^{M} D\,\frac{Of_c - pf_c}{2} R\left(\left|Y_j^c - Y_i^c\right|\right)\frac{Y_j - Y_i}{c_{ij}}\right) + \hat{Z}_c \tag{1}$$

Whereas $Y_j$ and $Y_i$ represents $j^{th}$ and $i^{th}$ grasshoppers' places, correspondingly; $c_{ij}$ signifies as far away$i^{th}$ and $j^{th}$ grasshoppers are; $O_{fc}$ and $p_{fc}$ signify the maximal and minimal boundaries in the $D^{th}$ dimension, correspondingly; $M$ represents the grasshopper's population, $Z_c$ signifies the size of the $D^{th}$ dimension in the most responsible solution established thus far, and is a coefficient to decrease the repulsion, attraction, and comfort places. The function $R$, which calculates the efficiency of the social services, is described by Eq. (2).

$$R(s) = ga^{-\frac{s}{p}} - a^{-s} \tag{2}$$

Whereas $p$ signifies the attraction level length and $g$ symbolizes the attraction force intensity. To reach equilibrium amongst the exploitative and exploratory stages, the variable c, recognized as the decreasing coefficient, must be consistently decreased to the iteration amounts. The decreasing coefficient (D)was measured utilizing Eq. (3).



DOI: https://doi.org/10.54216/IJNS.250418

$$D = D_{\max} - S_i \frac{D_{\max} - D_{\min}}{N_i} \tag{3}$$

Whereas $D_{\min}$ and $D_{\max}$ symbolize the lowest and highest values, individually; $Si$ signifies the running iteration, and $Ni$ characterizes the maximum iteration counts.

The fitness function (FF) reveals the accuracy of the classifier and the sum of preferred features. It exploits the classifier accuracy and lessens the chosen feature set dimension. So, the below-mentioned FF is employed for evaluating individual solutions, as exposed in Eq. (4).

$$Fitness = \alpha * ErrorRate + (1-\alpha) * \frac{\#SF}{\#All\_F} \tag{4}$$

Here, $ErrorRate$ symbolizes the classification rate of error by employing the chosen features. $ErrorRate$ is intended as the percentage of incorrect, which classifies to the amount of classifications set within the range of 0 and 1. $\#SF$ and $\#All\_F$ refers to the quantity of selected and total amount of features in an original data. $\alpha$ is employed for controlling the significance of classifier excellence and sub-set length. $\alpha$ value is 0.9.

### C. PNBM-based Classification Process

For the detection and classification procedure, the PNBM model has been employed. Neutrosophic Sets (NS) are a general method of the IFL philosophy [20]. In the philosophy of Neutrosophy, there are no limitations for indeterminacy, falsity, and truth and having a unit true intermission value for all component NS. These values are self-determining of one another. Occasionally, intuitionistic fuzzy logic (IFL) is insufficient for resolving certain real-world difficulties, for example, engineering obstacles. Therefore, mathematically, neutrosophy components have become significant in modelling these difficulties. Researches were accompanied by various mathematics fields and other associated sciences, particularly computer science.

Definition3. Let $E$ stand the world of discourse and $A \subseteq$ E. $A = \{(x, T(x), I(x), F(x)): x \in E\}$ is a NS or single-value NS ($SVNS$), while $T_A, I_A, F_A: A \to ]^-0, 1^+[$ represents the function of truth-membership, falsity-membership, and indeterminacy-membership, correspondingly. Now, $-0 \le T_A(x) + I_A(x) + F_A(x) \le 3^+$.

Definition4. Over SVNS $A$ in $E$, the threefold $\langle T_A, I_A, F_A \rangle$ is named the single-value Neutrosophic Number ($SVNN$).

Definition5. Let $n = \langle T_n, I_n, F_n \rangle$ remain $SVNN$, later the $n$ function of the score is provided as shown:

$$s_n = \frac{1 + T_n - 2I_n - F_n}{2} \tag{5}$$

Whereas $s_n \in [-1,1]$.

Definition6. Let $n = \langle T_n, I_n, F_n \rangle$ stand $SVNN$, after $n$ function of precision is specified as demonstrated:

$$h_n = \frac{2 + T_n - I_n - F_n}{3} \tag{6}$$

Now $h_n \in [0,1]$.

Definition7. Let $n_1$ and $n_2$ be dual $SVNNs$. Formerly, the grade of double $SVNNs$ is described as represented:

(i) If $s_{n_1} > s_{n_2}$, then $n_1 > n_2$,

(ii) If $s_{n_1} = s_{n_2}$ and $h_{n_1} \ge h_{n_2}$, then $n_1 \ge n_2$.

The underlying concept related to the Pythagorean Neutrosophic set (PNS) is considered in this part [21].

Considered $X$ as a non-empty or universe set. The PNS with $\psi$ and $\kappa$ as subject to membership

$$A = \{\langle x, \psi_A \zeta_A, k_A(x) \rangle | x \epsilon X\} \tag{7}$$

During Eq. (7),$\kappa_A$,$\psi_{At}$and$\sigma_A$ represents false, truth, and indeterminacy memberships correspondingly.

$$0 \le \psi^2 + \kappa^2 \le 1 \tag{8}$$

$$0 \le \psi^2 + \zeta^2 + \kappa^2 \le 2 \tag{9}$$

Take into account $x_1 = (\psi_{x_1}, \zeta_{x_1}, \kappa_{x_1})_t x_2 = (\psi_{x_2}, \sigma_{x_2}, \kappa_{x_2})$ and $x = (\psi_x, \zeta_x, \kappa_x)$ are PNSs, then the operation instructions are provided under:

$$x_1 \oplus x_2 = \left(\sqrt{\psi_{\chi_1}^2 + \psi_{\chi_2}^2 - \psi_{\chi_1}^2 \psi_{\chi_2}^2} t \zeta_{\chi_1} \zeta_{\chi_2} \kappa_{\chi_1} \kappa_{\chi_2}\right) \tag{10}$$

$$x_1 \otimes x_2 = \left(\psi_{\chi_1} \psi_{\chi_2}, \zeta_{\chi_1} + \zeta_{\chi_2} - \zeta_{\chi_1} \zeta_{\chi_2} \sqrt{\kappa_{\chi_1}^2 + \kappa_{\chi_2}^2 - \kappa_{\chi_1}^2 \kappa_{\chi_2}^2}\right) \tag{11}$$



DOI: https://doi.org/10.54216/IJNS.250418

$$\mu x = \left(\sqrt{1-(1-\psi_{\chi}^{2})^{\mu}}, \zeta^{\mu}, \kappa_{\chi}^{\mu}\right) where \mu \in \Re \; and \; \mu \geq 0 \quad (12)$$

$$\chi^{\mu} = (\psi_{\chi t}^{\mu} 1-\left(1-\zeta_{\chi}\right)^{\mu t}\sqrt{1-(1-\kappa_{\chi}^{2})^{\mu}}) \; where \; \mu \in \Re \; and \; \mu \geq 0 \quad (13)$$

Assuming $p_t q \geq 0$ with $\chi_i = (\psi_i(x), \zeta_i \kappa_i$ whereas $(i = 1,2,3, \dots, n)$ refers to PNS, formerly PNBM specified under

$$PNBM \; (x_1, x_2, \dots, x_n)^{p,q} = \left(\frac{1}{n(n-1)} \oplus_{\substack{i,j,=1 \\ i \neq j}}^{n} \left(x_i^p \otimes x_j^p\right)\right)^{\frac{1}{p+q}} \quad (14)$$

Stage 1: Generate the Direct-Relation Matrix, $X^k$

The matrix was created by PNS to describe the direct relationship according to the alternative decision-making and evaluates the choice as the matrix of non-negative, $X^k = [x_{ij}^k]_{n \times n}$, while $1 \leq k \leq m$. The representation of $\chi_{ij} = \left(\psi_{\mathrm{ij}}, \zeta_{ij}, \kappa_{ij}\right)$ identifies the amount the decision-making regarded that situations $i$ impact condition $j$, with the diagonal element becoming 0. This marking has been calculated by seven semantical procedures ranging between $no \; effect$ to $high \; effect$ based on PNS's linguistic-variable. As a result, there is a $m$ dissimilar matrix $X^k = \{X^1, X^2, \dots, X^m\}$ according to each $DM$.

Stage 2: Get the Aggregative Direct-Relation Matrix $A$.

According to Eq. (14), PN-NWBM incorporates the direct-relation matrix $X^k = \{X^1, X^2, \dots, X^m\}$ into an integrated decision matrix $A =$ whereas $a_{ij} = \left(\psi_{ij}, \zeta_{ij}, \kappa_{ij}\right)$ .

Step 3: Deneutrosophication into crisp-matrix $B$.

The aggregate matrix $A = [a_{ij}]_{m \times n}$ into the matrix of crisp, $B$ is deneutrosophicate in Eq. (15).

$$B = \frac{\psi_A(x) + \zeta_A(x) + \kappa_A(x)}{3} \quad (15)$$

Stage 4: Normalization of the matrix to normalized matrix $Z$

Eq. (10) has been applied to standardize the procedure of matrix.

$$Z = \frac{B}{s} \quad (16)$$

Here $s = \max \sum_{j=1}^{n} b_{ij}$ and all components in the matrix $Z$ follow $0 \leq z_{ij} < 1$.

Stage 5: Constructing the Total-Influence Matrix $T$

Constructing the influence matrix utilizing the succeeding expression.

$$T = Z(l - Z)^{-1} \quad (17)$$

During Eq. (17), $l$ signifies the matrix of identity.

Stage 6: Calculate the number of Columns and Rows

The vector $C$ and $R$ represents the number of rows and the complete-influence matrix $T$ can be described in Eq. (18) and (19).

$$R = [\tilde{r}_l]_{n \times 1} = [\sum_{j=1}^{n} r_{ij}]_{n \times 1} \quad (18)$$

$$C = [\tilde{c}_l]_{1 \times n} = [\sum_{i=1}^{n} r_{ij}]_{1 \times n}^{T} \quad (19)$$

Now $t_{ij}$ denotes the matrix component $T$.

$R + C$ and $R$- $C$ values define relationship and significance values, consistently.

Stage 7: Network Relationship Map (NRM), the Value of Threshold

The graph dataset design is $(R + C_t R - C)$. $R + C$ and $R - C$ are marked on the horizontal and vertical axis.

**D. Parameter Optimizer using FFOA**



DOI: https://doi.org/10.54216/IJNS.250418

Additionally, the FFOA-based hyperparameter range method has been used to heighten the recognition outcomes of the PNBM system. The FFOA is a meta-heuristic optimizer approach, which mimics the fireflies' behaviour, while the movement and attractiveness of fireflies are controlled by their brilliance [22]. The main concept overdue the Firefly Model is that all fireflies appeal to optimistic fireflies, through the brilliance being relative to the objective function being fine-tuned.

Here, the Firefly Method has been applied to enhance the hyperparameters of the CNN method, containing the neuron counts in the batch size, rate of learning, dense layers, and filter counts within the convolution layers. The optimizer procedure included utilizing 10 fireflies over 20 iterations, taking into consideration efficient exploration of the hyperparameter area to discover optimum values. The stages of the Firefly Method are as shown:

1. Initialization

Initialize a firefly population with arbitrary locations in the search space. All fireflies characterize possible solutions. During the study, the primary firefly count was 10, with 20 iterations.

2. Attractiveness

The fireflies' light intensity $I$ at a certain position $x$ can be established by the objective function$f(x)$. The fireflies' attractiveness $\beta$ is provided by Eq. (20):

$$\beta(r) = \beta_0 e^{-\gamma r^2} \tag{20}$$

Whereas $\beta_0$ refers to maximal attractiveness, $\gamma$ denotes the coefficient of the absorption of light, and $r$ signifies distance amongst dual fireflies. Here, $\gamma$ was fixed to 1.0.

3. Calculation of Distance

The distance $r_{ij}$ amongst dual fireflies $i$ and $j$ at locations $x_i$ and $\chi_j$ is computed utilizing the Euclidean distance, as in Eq. (21):

$$r_{ij} = \|x_{\mathrm{i}-x_j}\| \tag{21}$$

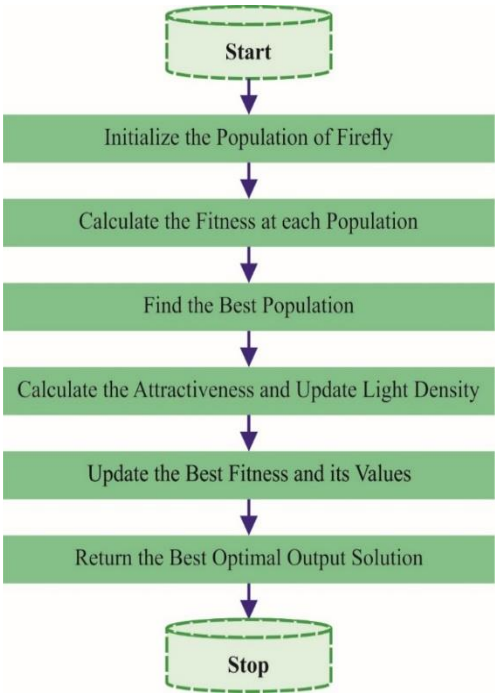


**Figure 2.** Flowchart of FFOA

4. Movement

A firefly $i$ moves near a more appealing (optimistic) firefly $j$. The movements are established by Eq. (22):

$$x_i = x_i + \beta_0 e^{-\gamma r_{ij}^2}(x_j - x_i) + \alpha(rand - 0.5) \tag{22}$$

Whereas $\alpha$ denotes the randomized parameter and $rand$ symbolizes a randomly generated number uniformly distributed amongst (0, 1).



DOI: https://doi.org/10.54216/IJNS.250418

The $\beta_0$ the values applied in the study were fixed to 0.2, and the $\alpha$ value became 0.5. The term $\beta_0 e^{-\gamma r_i^2} i(x_j - x_i)$ guarantees that fireflies move near all others depending on their attractiveness, which is robust for brighter and nearer fireflies. The factor of exponential $e^{-\gamma r_i^2} i$ guarantees that the impact reduces with distance, making different fireflies less appealing. The term $\alpha(rand - 0.5)$ presents stochastic behavior, stopping the model from becoming trapped in local optima and improving the exploration of the search space. Fig. 2 depicts the flowchart of FFOA. The FFOA originates an FF for accomplishing an enhanced outcome of classification. It defines a positive numeral to imply the better efficiency of the candidate solution. Here, the reduction of the classifier rate of error has been measured as FF. Its formulation is expressed in Eq. (23).

$$fitness(x_i) = ClassifierErrorRate(x_i) = \frac{no.of\ misclassified\ samples}{Total\ no.of\ samples} * 100 \quad (23)$$

## 4. Experimental Validation

In this section, the experimental validation of the IDSSFFD-PNBM method is examined under the financial fraud detection dataset [23]. The dataset contains 200 instances under dual classes such as isFraud_Yes and isFraud_No as exposed in Table 1. The total number of features is 10 such as step, type, amount, nameOrig, oldbalnceOrg, newbalceOrig, nameDest, oldbalanceDest, isFlaggedFraud, and newbalanceDest. But only seven features are selected. They are step, oldbalnceOrg, newbalceOrig, nameDest, oldbalanceDest, newbalanceDest, isFlaggedFraud.

**Table 1:** Details of Dataset

| Class labels | No. of Instances |
|---|---|
| isFraud_Yes | 100 |
| isFraud_No | 100 |
| Total Instances | 200 |

Fig. 3 established the classifier results of the IDSSFFD-PNBM methodology. Figs. 3a-3b displays the confusion matrices with correct recognition and classification of each class below 70%TRPH and 30%TSPH. Fig. 3c demonstrates the PR curve, indicating superior outcomes across every class. At the same time, Fig. 3d illustrates the ROC values, signifying proficient outcomes with better ROC analysis for different classes.

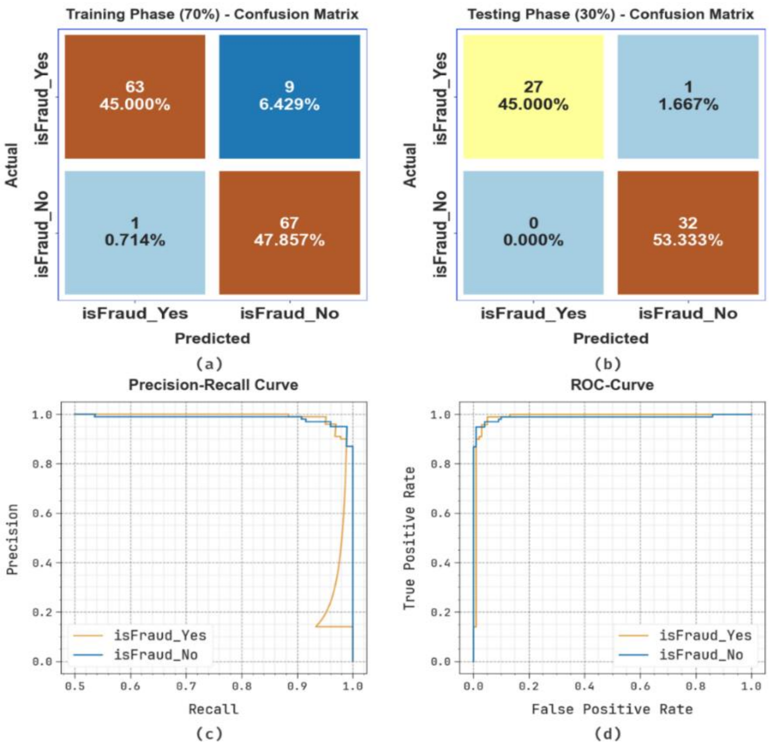


**Figure 3.** Classifier result (a-b) 70%TRPH and 30%TSPH of Confusion matrix and (c-d) curves of PR and ROC



DOI: https://doi.org/10.54216/IJNS.250418

In Table 2 and Fig. 4, the financial fraud detection of the IDSSFFD-PNBM algorithm under 70%TRPH and 30%TSPH is illustrated. The results showed that the IDSSFFD-PNBM approach reached the effectual detection of dual-class labels. Based on 70%TRPH, the IDSSFFD-PNBM system reaches average $accu_y$ of 92.86%, $prec_n$ of 93.30%, $reca_l$ of 93.01%, $F1_{score}$ of 92.85%, and $AUC_{score}$ of 93.01%. Followed by, with 30%TSPH, the IDSSFFD-PNBM approach attains average $accu_y$ of 98.33%, $prec_n$ of 98.48%, $reca_l$ of 98.21%, $F1_{score}$ of 98.32%, and $AUC_{score}$ of 98.21%.

**Table 2:** Financial fraud detection of IDSSFFD-PNBM method under 70%TRPH and 30%TSPH

| Classes | $Accu_y$ | $Prec_n$ | $Reca_l$ | $F1_{score}$ | $AUC_{score}$ |
|---|---|---|---|---|---|
| TRPH (70%) | | | | | |
| isFraud_Yes | 92.86 | 98.44 | 87.50 | 92.65 | 93.01 |
| isFraud_No | 92.86 | 88.16 | 98.53 | 93.06 | 93.01 |
| Average | 92.86 | 93.30 | 93.01 | 92.85 | 93.01 |
| TSPH (30%) | | | | | |
| isFraud_Yes | 98.33 | 100.00 | 96.43 | 98.18 | 98.21 |
| isFraud_No | 98.33 | 96.97 | 100.00 | 98.46 | 98.21 |
| Average | 98.33 | 98.48 | 98.21 | 98.32 | 98.21 |

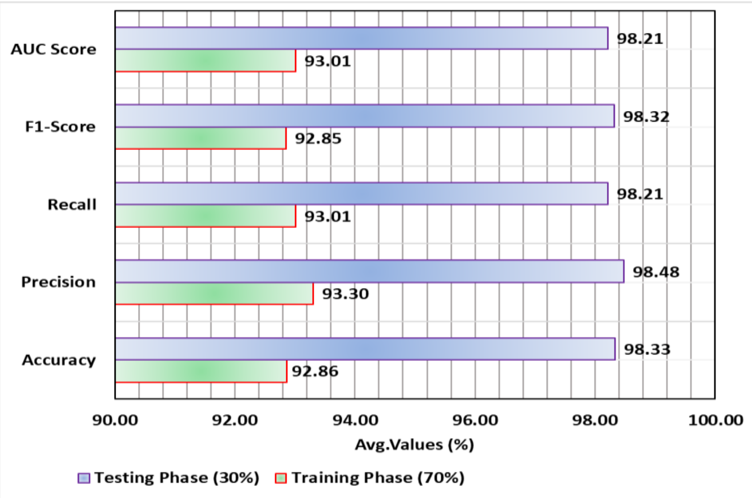


**Figure 4.** Average of IDSSFFD-PNBM method below 70%TRPH and 30%TSPH

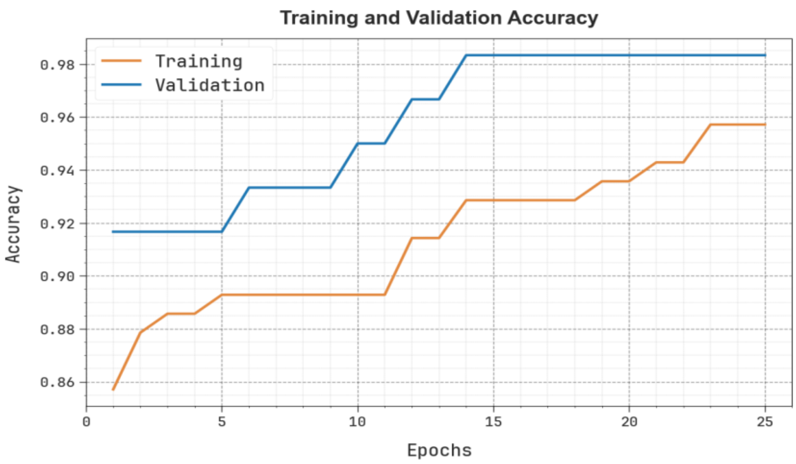


**Figure 5.** $Accu_y$ Analysis of the IDSSFFD-PNBM technique



DOI: https://doi.org/10.54216/IJNS.250418

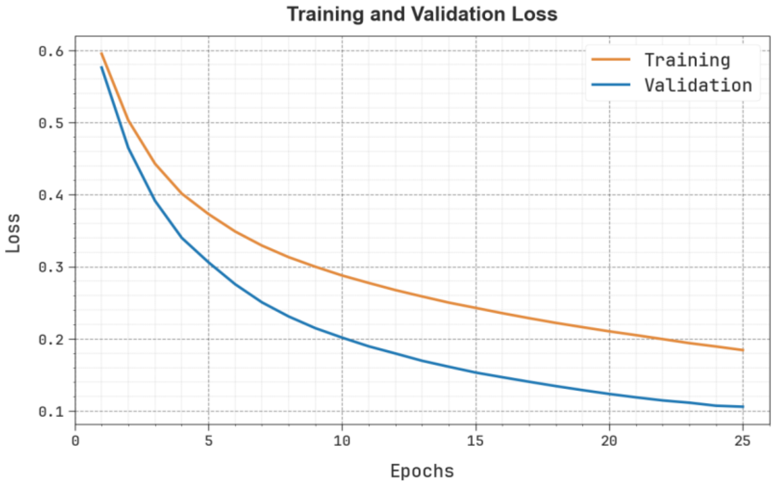


**Figure 6.** Loss graph of IDSSFFD-PNBM method

In Fig. 5, the training (TRA) $accu_y$ and validation (VAL) $accu_y$ results of the IDSSFFD-PNBM methodology are illustrated. The $accu_y$ analysis are computed within the range of 0-25 epochs. The figure highlighting that the TRA and VAL $accu_y$ analysis exhibitions an increasing trend which informed the capacity of the IDSSFFD-PNBM algorithm with maximal performance across multiple iterations. Simultaneously, the TRA and VAL $accu_y$ leftovers closer under the epochs, which indicates inferior overfitting and demonstrates the higher performance of the IDSSFFD-PNBM methodology, assuring reliable prediction on hidden samples.

In Fig. 6, the TRA loss (TRALOS) and VAL loss (VALLOS) curve of the IDSSFFD-PNBM technique is shown. The values of loss are computed across an interval of 0-25 epochs. The continuous reduction in values of loss besides assurances of the maximum performance of the IDSSFFD-PNBM methodology and tuning the prediction results over time.

Table 3 and Fig. 7 present the comparative outcomes of the IDSSFFD-PNBM approach with existing methods [24, 25]. The results imply that the proposed IDSSFFD-PNBM system has achieved better performance $accu_y$ of 98.33%, $prec_n$ of 98.48%, $reca_l$ of 98.21%, and $F1_{score}$ of 98.32%. Whereas, the existing techniques DT, LR, SVM, RF, XGBoost, DeepWalk, and Node2Vec methodologies have gained the worst performance below different metrics.

**Table 3:** Comparative results of IDSSFFD-PNBM technique with existing classifiers

| Classifier | $Accu_y$ | $Prec_n$ | $Reca_l$ | $F1_{score}$ |
|---|---|---|---|---|
| Decision Tree | 82.00 | 90.30 | 82.03 | 87.78 |
| Logistic Regression | 72.13 | 87.70 | 81.46 | 88.11 |
| SVM Method | 97.62 | 79.26 | 95.90 | 94.81 |
| Random Forest | 87.42 | 96.99 | 85.10 | 86.22 |
| XGBoost | 98.01 | 97.83 | 86.34 | 96.83 |
| DeepWalk | 97.23 | 88.40 | 94.95 | 88.36 |
| Node2Vec | 95.88 | 97.32 | 93.10 | 97.21 |
| IDSSFFD-PNBM | 98.33 | 98.48 | 98.21 | 98.32 |



DOI: https://doi.org/10.54216/IJNS.250418

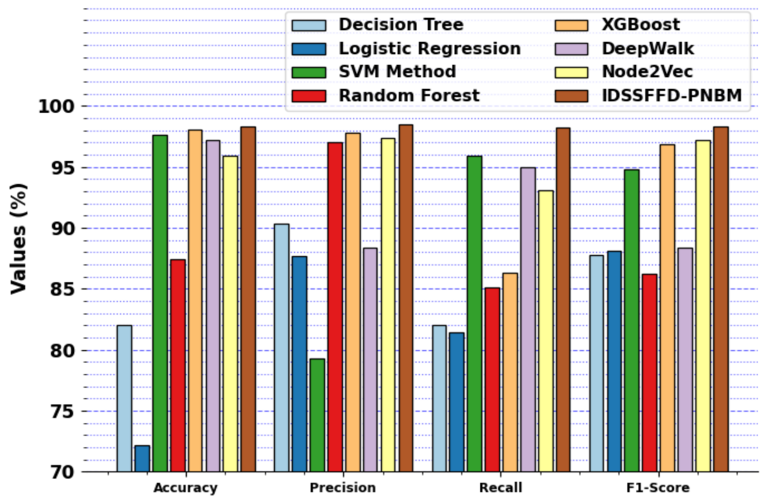


**Figure 7.** Comparative outcomes of IDSSFFD-PNBM methodology with existing classifiers

In Table 4 and Fig. 8, the comparative outcomes of the IDSSFFD-PNBM model are specified in terms of computational time (CT). Based on CT, the proposed IDSSFFD-PNBM method obtains a lesser CT of 6.77sec whereas the DT, LR, SVM, RF, XGBoost, DeepWalk, and Node2Vec techniques attain greater CT values of 10.38sec, 13.91sec, 15.51sec, 11.21sec, 8.08sec, 17.76sec, and 17.92sec, respectively.

**Table 4:** CT result of IDSSFFD-PNBM model with existing classifiers

| Classifier | Computational Time (Sec) |
|---|---|
| Decision Tree | 10.38 |
| Logistic Regression | 13.91 |
| SVM Method | 15.51 |
| Random Forest | 11.21 |
| XGBoost | 8.08 |
| DeepWalk | 17.76 |
| Node2Vec | 17.92 |
| IDSSFFD-PNBM | 6.77 |

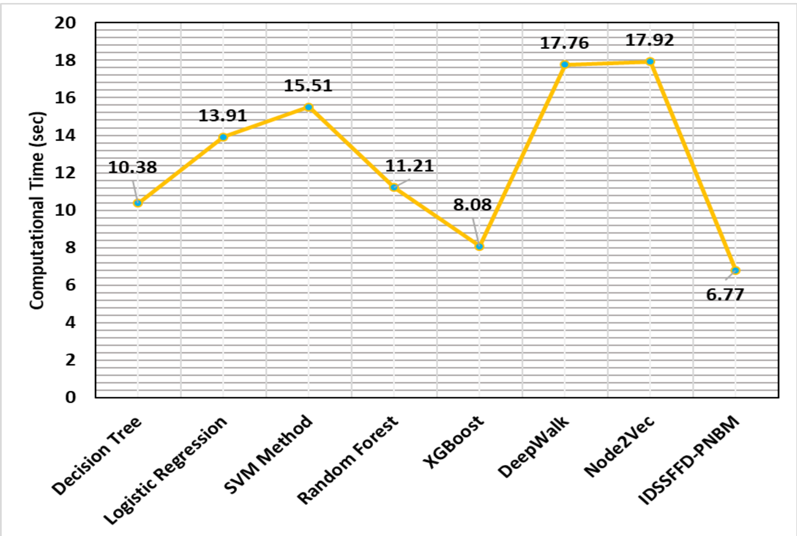


**Figure 8.** CT outcome of IDSSFFD-PNBM technique with existing classifiers



DOI: https://doi.org/10.54216/IJNS.250418

## 5. Conclusion

In this paper, we have presented a novel IDSSFFD-PNBM model. The main intention of the IDSSFFD-PNBM algorithm is to enrich the detection model for financial fraud using an advanced optimization model. Initially, the z-score normalization is applied in the data normalization stage for converting input data into a beneficial format. Besides, the proposed IDSSFFD-PNBM designs GOA for the selection of feature processes to enhance the efficiency and performance of the model. In addition, the PNBM model has been employed for the detection and classification procedure. Additionally, the FFOA-based hyperparameter range method has been used to heighten the recognition outcomes of the PNBM system. The experimental evaluation of the IDSSFFD-PNBM technique takes place using a benchmark dataset. The experimental results indicated an enhanced performance of the IDSSFFD-PNBM technique compared to recent approaches.

**Funding:** "This research received no external funding"

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

## References

[1] F. Taş, S. Topal, and F. Smarandache, "Clustering neutrosophic data sets and neutrosophic valued metric spaces," *Symmetry*, vol. 10, no. 10, p. 430, 2018.

[2] A. H. Saheb and R. H. Buti, "A specific category of harmonic functions characterized by a generalized Komatu operator in conjunction with the (RK) integral operator and applications to neutrosophic complex field," *Full Length Article*, vol. 23, no. 3, pp. 44-4, 2024.

[3] A. H. Saheb and R. H. Buti, "On the characterization of harmonic functions with a generalized Komatu operator and its applications to neutrosophic complex field," *Mathematics and Computers in Simulation*, vol. 187, pp. 113-125, 2021.

[4] K. G. Al-Hashedi and P. Magalingam, "Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019," *Computer Science Review*, vol. 40, p. 100402, 2021.

[5] K. S. Rajasekaran, "Application of machine learning techniques in financial fraud detection," *Journal of Financial Technology*, vol. 5, no. 2, pp. 112-126, 2023.

[6] A. Ali, S. Abd Razak, S. H. Othman, T. A. Eisa, A. Al-Dhaqm, M. Nasser, T. Elhassan, H. Elshafie, and A. Saif, "Financial fraud detection based on machine learning: A systematic literature review," *Applied Sciences*, vol. 12, no. 19, p. 9637, 2022.

[7] D. Huang, D. Mu, L. Yang, and X. Cai, "CoDetect: Financial fraud detection with anomaly feature detection," *IEEE Access*, vol. 6, pp. 19161-19174, 2018.

[8] W. Hilal, S. A. Gadsden, and J. Yawney, "Financial fraud: A review of anomaly detection techniques and recent advances," *Expert Systems with Applications*, vol. 193, p. 116429, 2022.

[9] A. Mehbodniya, I. Alam, S. Pande, R. Neware, K. P. Rane, M. Shabaz, and M. V. Madhavan, "[Retracted] Financial fraud detection in healthcare using machine learning and deep learning techniques," *Security and Communication Networks*, vol. 2021, p. 9293877, 2021.

[10] K. G. Al-Hashedi and P. Magalingam, "Financial fraud detection applying data mining techniques: A comprehensive review from 2009 to 2019," *Computer Science Review*, vol. 40, p. 100402, 2021.

[11] A. A. Alhabib, A. F. Alasiri, M. B. Alharbi, S. Ahmad, and A. E. M. Eljialy, "Credit card fraud detection using random forest and K-nearest neighbors (KNN) algorithms," in *International Conference on Cognitive Computing and Cyber Physical Systems, Singapore*, 2025, pp. 383-395.

[12] M. M. Ismail and M. A. Haq, "Enhancing enterprise financial fraud detection using machine learning," *Engineering, Technology & Applied Science Research*, vol. 14, no. 4, pp. 14854-14861, 2024.

[13] J. Ghanim and M. Awad, "An unsupervised anomaly detection in electricity consumption using reinforcement learning and time series forest based framework," *Journal of Artificial Intelligence and Soft Computing Research*, vol. 15, no. 1, pp. 5-24, 2025.

[14] R. Ming, O. Abdelrahman, N. Innab, and M. H. K. Ibrahim, "Enhancing fraud detection in auto insurance and credit card transactions: A novel approach integrating CNNs and machine learning algorithms," *PeerJ Computer Science*, vol. 10, p. e2088, 2024.

[15] Z. A. Latipov, K. A. Naminova, I. S. Abdullayev, A. E. Ilyin, R. A. Shichiyakh, and E. L. Lydia, "Optimizing financial fraud detection: Understandings from variable selection with neutrosophic vague soft set," *International Journal of Neutrosophic Science*, vol. 3, pp. 219-19, 2025.

[16] L. Theodorakopoulos, A. Theodoropoulou, F. Zakka, and C. Halkiopoulos, "Credit card fraud detection with machine learning and big data analytics: A PySpark framework implementation," 2024.

[17] N. Ding, X. Ruan, H. Wang, and Y. Liu, "Automobile insurance fraud detection based on PSO-XGBoost model and interpretable machine learning method," *Insurance: Mathematics and Economics*, vol. 120, pp. 51-60, 2025.



DOI: https://doi.org/10.54216/IJNS.250418

[18] L. Peng, Z. Lu, T. Lei, and P. Jiang, "Dual-structure elements morphological filtering and local Z-score normalization for infrared small target detection against heavy clouds," *Remote Sensing*, vol. 16, no. 13, p. 2343, 2024.

[19] N. O. Ogunsola, C. Shin, A. I. Lawal, Y. K. Kim, and S. Cho, "Blasting-induced ground vibration modeling in tunnel excavation: A comparative study of ANN, hybrid ANNs, and empirical models," *Rock Mechanics and Rock Engineering*, pp. 1-27, 2024.

[20] F. Taş, S. Topal, and F. Smarandache, "Clustering neutrosophic data sets and neutrosophic valued metric spaces," *Symmetry*, vol. 10, no. 10, p. 430, 2018.

[21] M. N. S. Swamy, "A survey on advanced decision-making techniques based on fuzzy and neutrosophic set theories," *Journal of Intelligent & Fuzzy Systems*, vol. 38, no. 4, pp. 4749-4762, 2020.

[22] A. A. R. K. Bsoul, "Human activity recognition using graph structures and deep neural networks," *Computers*, vol. 14, no. 1, p. 9, 2024.

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

[24] B. Li, J. Yen, and S. Wang, "Uncovering financial statement fraud: A machine learning approach with key financial indicators and real-world applications," *IEEE Access*, 2024.

[25] H. Zhou, G. Sun, S. Fu, L. Wang, J. Hu, and Y. Gao, "Internet financial fraud detection based on a distributed big data approach with node2vec," *IEEE Access*, vol. 9, pp. 43378-43386, 2021.



DOI: https://doi.org/10.54216/IJNS.250418