Effective Data Classification using Interval Neutrosophic Covering Rough Sets based on Neighborhoods for FinTech Applications

Maksim Kuznetsov[1,2,*], Irina Kosorukova[3,4], Veronika Denisovich[5,6], Elena Klochko[7], Alexey Dengaev[8]

[1]Department of Economics and Management of Elabuga Institute, Kazan Federal University, Kazan, 420008, Russia
[2]Department of Economics and Management, Khorezm University of Economics, Urgench, 220100, Uzbekistan;
[3]Department of Valuation and Corporate Finance, Moscow University for Industry and Finance "Synergy", Moscow, 125315, Russia
[4]Department of Corporate Finance and Corporate Governance, Financial University under the Government of the Russian Federation, Moscow, 125993, Russia
[5]Institute of Digital Technologies and Law, Kazan Innovative University named after V. G. Timiryasov, Kazan, 420111, Russia
[6]Department of International Relations Political Science and Regional Studies, South Ural State University (National Research University), Chelyabinsk, 454080, Russia
[7]Department of Management, Kuban State Agrarian University named after I.T. Trubilin, Krasnodar, 350044, Russia
[8]Faculty of Oil and Gas Fields Development, Gubkin Russian State University of Oil and Gas, Moscow, 119991, Russia

Emails: kuznetsov.m.s@inbox.ru; i.v.kosorukova@yandex.ru; veronika.denisovich@inbox.ru; klochko.e.n@yandex.ru; dengaev.a@gubkin.ru

Abstract

Neutrosophic set (NS) is particularly appropriate in applications where data is incomplete, unclear, or inconsistent, which offers an effectual means for analyzing and exhibiting complex mechanisms. An NS is a mathematical technique to manage uncertainty, indeterminacy, and imprecision. It enlarges classical sets, IF sets, and fuzzy sets by presenting three degrees such as indeterminacy (I), false (F), and truth (T). Financial technology (Fintech) plays an essential part in advancing modern society, technology, economies, and various fields. Financial crisis prediction (FCP) plays a crucial role in shaping economic outcomes. Past research has predominantly focused on using deep learning (DL), machine learning (ML), and statistical methods to forecast the financial stability of business. In this manuscript, we focus on the development of Effective Data Classification using Interval Neutrosophic Covering Rough Sets based on Neighborhoods and Multi-Strategy Improved Butterfly Optimization (EDCINCRS-MSIBO) Algorithm for FinTech Applications. It contains distinct kinds of stages such as data normalization, feature selection, classification, and parameter tuning. In the EDCINCRS-MSIBO technique, a min-max normalization-based data pre-processing model to scale the raw data into a uniform format. For feature subset selection, the whale optimizer algorithm (WOA) is employed to reduce the dimensionality and improve model efficiency by selecting the most relevant features. In addition, the EDCINCRS-MSIBO technique takes place interval neutrosophic covering rough sets (INCRS) classifier is utilized for detection and classification of a financial crisis. Finally, a multi-strategy improved butterfly optimization algorithm (MSIBOA) is exploited for the optimum parameter adjustment of the INCRS model. To confirm the better predictive solution of the EDCINCRS-MSIBO model, a wide range of simulations are executed on the two benchmark databases. The comparative result analysis displays the encouraging outcomes of the EDCINCRS-MSIBO method on the existing techniques

Keywords: Neutrosophic Set; Rough Sets; Financial Crisis Prediction; Neutrosophic Covering Rough Sets; FinTech; Feature Subset Selection

DOI: https://doi.org/10.54216/IJNS.250319
Received: February 26, 2024 Revised: May 28, 2024 Accepted: October 05, 2024

1. Introduction

Neutrosophic Logic is a neonatal investigation region in which every proposal has been approximated to have the percentage (proportion) of truth in subsection T, the percentage of indeterminacy in subsection I, and the percentage of falsity in subsection F [1]. Neutrosophic set (NS) is effectively utilized for unknown information handling and establishes benefits to handle the indeterminacy data and is even a model encouraged for classification and data analysis application [2]. NS offers an accurate and effective method to describe imbalanced information based on the data attributes [3]. Recently, Financial Technology (FinTech) has obtained the highest attention from developers using discerning taxonomy. The Financial Technology areas are widely employed in several domains of firms, and industries where the performance effectiveness was improved with Information Technology (IT) applications [4]. FinTech is measured as an effective method due to its belongings of inspiring forces accompanied by scientific utilization, cost-effective demands, innovative business (market), and user necessities [5]. It is witnessed that FinTech is applied as an important savings in the most high-ranked companies. The wide-ranging application of FinTech has huge challenges in planning and adoption, due to the intercrossed dominions, composite integrated model, and typical necessities. Therefore, the modern and accurate knowledge of FinTech comprises an instant request for experts and academics. Financial crisis prediction (FCP) is one of the most challenging requirements for organizations in making informed financial decisions [6].

Artificial intelligence (AI) and statistical methods were used to recognize the significant features of FCP. In these techniques, AI is applied for performance authorization and predictions of whether the method undergoes a problem or not [7]. Numerous studies were shown on the classification of FCP. Recently, traditional models used arithmetical functions to predict the financial difficulty that distinguishes financial organizations from sturdier and fragile ones [8]. Recently, AI models have been modified to develop standard classification techniques [9]. Nevertheless, the being of different features in the great dimension of financial data outcomes in different difficulties namely overfitting, lower interoperability, and higher computational complexities [10].

This manuscript focuses on the development of Effective Data Classification using Interval Neutrosophic Covering Rough Sets based on Neighbourhoods and Multi-Strategy Improved Butterfly Optimization (EDCINCRS-MSIBO) Algorithm. In the EDCINCRS-MSIBO technique, a min-max normalization-based data pre-processing model to scale the raw data into a uniform format. For feature subset selection, the WOA is utilized. In addition, the EDCINCRS-MSIBO technique takes place interval neutrosophic covering rough sets (INCRS) classifier is utilized for detection and classification of a financial crisis. Finally, a multi-strategy improved butterfly optimization algorithm (MSIBOA) is exploited for the optimal hyperparameter adjustment of the INCRS classifier. The comparative result analysis shows the promising performance of the EDCINCRS-MSIBO method on the recent approaches.

2. Literature Review

Osadchy et al. [11] presented a jellyfish search algorithm-based feature selection using the optimal DL process (JSAFS-ODL) for the prediction of a financial crisis. The JSAFS-ODL approach uses JSA-based FS (JSA-FS) for selecting the optimal feature set. For the FCP, the RNN-GRU was applied. To assurance the improved functioning of the JSAFS-ODL process, a sequence of examinations was intricate. The author [12] implemented a hybrid hunter-prey optimization using a DL-based FCP (HHPODL-FCP) algorithm. To achieve this, the HHPODL-FCP system exploits the HHPO model for the feature sub-set selection procedure. A SSA-based parameter tuning procedure for enriching the implementation of the GARN method. Wu et al. [13] introduced an original DL technique, a user response guided deep attention network (URGDAN), for predicting economic suffering with recent statements. During the presented model, a DL framework was built to incorporate financial signs, user reactions, and current statement texts. URGDAN utilizes the user reactions to existing statements to direct the semantical features sign of the information, it additionally recognizes result information that has an important connection with corporation economic suffering.

Kalaivani and Saravanan [14] provide a strategy for Enhancing FCP with the COA using the ML (EFCP-COAML) model. To obtain this, the EFCP-COAML model includes dual main procedures like hyperparameter and classification tuning. Initially, the EFCP-COAML model utilizes a KELM-based forecast process. In the succeeding stage, the COA has been used for the hyperparameter choice of the KELM method thereby improving the valuable performance. In [15], a TCN model, which depends on a CNN method is presented to build an initial warning method to anticipate financial difficulty. The recommended TCN in comparison with the logit method and other DL algorithms. The decomposition value of Shapley has been designed for the interpretability of the initial warning method. The author [16] developed an original multi-vs. Optimization (MVO)-based FS using an optimum VAE technique for finiancial prediction. The presented MVOFS-OVAE method mainly pre-processes the economic data with min-max normalization. Moreover, the targets a feature sub-set choice procedure with the MVOFS technique. Accompanied by, the VAE method was applied for the classification process.

DOI: https://doi.org/10.54216/IJNS.250319
Received: February 26, 2024 Revised: May 28, 2024 Accepted: October 05, 2024

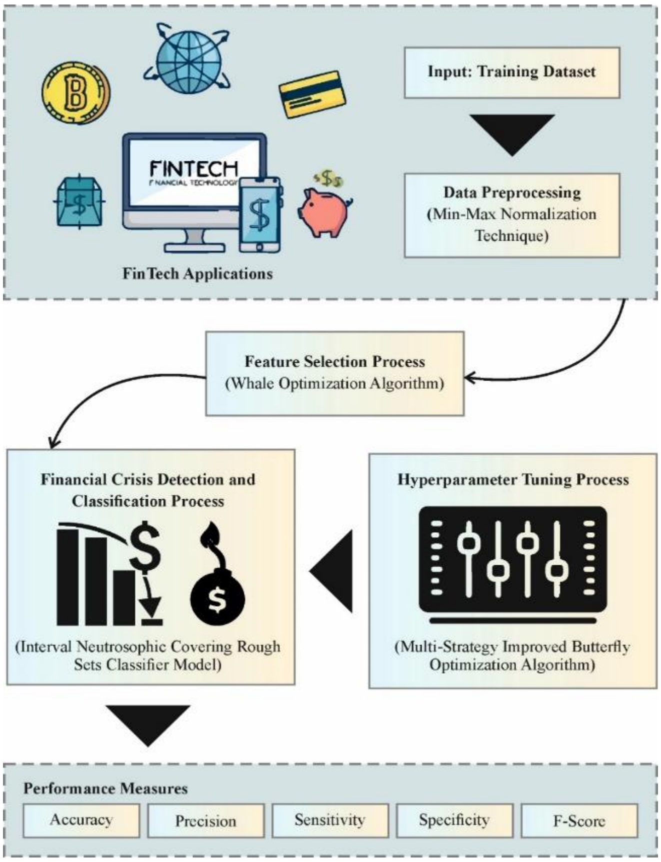

Figure 1. Overall flow of EDCINCRS-MSIBO algorithm

3. Proposed Methodology

In this paper, we concentration on the development of the EDCINCRS-MSIBO approach for FinTech Applications. It contains distinct kinds of stages such as data normalization, feature selection (FS), FCP detection and classification, and parameter optimizer. Fig. 1 illustrates the entire flow of the EDCINCRS-MSIBO algorithm.

A. Stage I: Min-Max Normalization

Initially, the EDCINCRS-MSIBO system undergoes a min-max normalization-based data pre-processing model to scale the raw data into a uniform format. Min-max normalization is vital in the prediction of the financial crisis as it measures the information to an exact range, usually between 0 and 1, which improves the performance of ML models [17]. This normalization system aids in alleviating the effect of outliers and varying scales amongst features, permitting an additional precise analysis of trends and patterns, which leads to financial crises. By giving a reliable structure for equating variables, min-max normalization helps to enhance the reliability and robustness of predictive methods employed in finance.

B. Stage II: WOA-based Feature Subset Selection

For feature subset selection, the WOA is employed to reduce the dimensionality and improve model efficiency by selecting the most relevant features. Mirjalili and Lewis proposed WOA a recent heuristic optimization approach depends on predation behavior of humpback whales [18]. They will shrink to swim or surround the target in a spiral form once humpback whales hunt the prey and it can be mathematically modeled using Eq. (1):

$$\vec{X}(t+1)=\begin{cases}\vec{X}^{*}(t)-\vec{A}\cdot\vec{D} & p<0.5\\ \vec{X}^{*}(t)+\vec{D}'\cdot e^{bl}\cdot\cos(2\pi l) & p\geq 0.5\end{cases} \tag{1}$$

If $p<0.5$ and $|\vec{A}|<1$, then individuals update its location based on the encirclement and contraction model:

DOI: https://doi.org/10.54216/IJNS.250319
Received: February 26, 2024 Revised: May 28, 2024 Accepted: October 05, 2024

$$\begin{cases} \vec{X}(t+1) = \vec{X}^*(t) - \vec{A} \cdot \vec{D} \\ \vec{D} = \left|\vec{C} \cdot \vec{X}(t) - \vec{X}(t)\right| \\ \vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a} \\ \vec{C} = 2 \cdot \vec{r}_2 \\ \vec{a} = 2 - 2t/t_{\max} \end{cases} \tag{2}$$

Where the existing and the global optimum individual locations are $\overrightarrow{X}(t)$ and $\vec{X}^*(t)$, $\vec{A}$ and $\vec{C}$ are coefficient vectors, the distance among the present and the optimum individuals represents $\overrightarrow{D}$, $\vec{a}$ Is the convergence factor, $\vec{r}_1$ and $\vec{r}_2$ are random integers within $[0,1]$, t and $t_{\max}$ are the existing and maximal iterations.

If $p \geq 0.5$, then individuals update its location in a spiral mode, and it can be mathematically modeled using Eq. (3):

$$\begin{cases} \vec{X}(t+1) = \vec{X}^*(t) + \vec{D}' \cdot e^{bl} \cdot \cos(2\pi l) \\ \vec{D} = \left|\vec{X}^*(t) - \vec{X}(t)\right| \end{cases} \tag{3}$$

Figure 2. Steps involved in WOA

Where the logarithmic spiral constant is b, l is a randomly generated value $[1,1]$, and the distance between the present optimum and other individuals is $\vec{D}'$. If $p < 0.5$ and $|\vec{A}| \geq 1$, then individuals randomly choose whales from the present whale group to approach.

$$\begin{cases} \vec{X}(t+1) = \vec{X}_{\text{rand}}(t) - \vec{A} \cdot \vec{D} \\ \vec{D} = \left|\vec{C} \cdot \vec{X}_{\text{rand}}(t) - \vec{X}(t)\right| \end{cases} \tag{4}$$

Where the random whale individual location in the present whale population is $\vec{X}_{\text{rand}}(t)$. Fig. 2 demonstrates the steps involved in WOA. In the WOA model, the objectives are combined into a distinct objective formulation such that a current weight classifies every objective significance. In this work, we assume a fitness function (FF) that unites both objectives of FS as exposed in (5).

$$Fitness(X) = \alpha \cdot E(X) + \beta * \left(1 - \frac{|R|}{|N|}\right) \tag{5}$$

DOI: https://doi.org/10.54216/IJNS.250319
Received: February 26, 2024 Revised: May 28, 2024 Accepted: October 05, 2024

Here, $Fitness(X)$ signifies the fitness value of a sub-set X, $E(X)$ denotes the classification rate of error by utilizing the chosen features in the sub-set X, $|R|$ and $|N|$ represents the count of nominated features and the count of original features, correspondingly, α and β means a weight of classifier error and the reduction ratio.

C. Stage III: FCP using INCRS Classifier

In addition, the EDCINCRS-MSIBO technique takes place INCRS classifier is utilized for detection and classification of financial crisis. The summary of concepts and explanations of interval neutrosophic sets (INS), and covering rough sets (RS) [19].

Definition2.1. Assume X as a space of points, with a class of elements signified by x. In X, A $NS\ A$ was shortened by an indeterminacy-membership function (IMF) $I_{A(x)}$, a falsity-membership function (FMF) $F_{A(x)}$, and truth-membership function (TMF) $T_{A(x)}$. The functions $T_{A(x)}, F_{A(x)}$, and $I_{A(x)}$ are or non- and standard sub-sets of] 0^-, 1^+[.

Definition2.2. Assume X as a space of objects, with a type of elements in X meant by x. A single-valued NS N in X is shortened by an IMF $I_{N(x)}$, FMF $F_{N(x)}$, and TMF $T_{N(x)}$. Next, an INS A is represented below:

$$A = \{\langle x, T_A(x), I_A(x), F_A(x)\rangle x \in X\} \quad (6)$$

For ease, we denote to $A = \langle T_A, I_A, F_A\rangle = \{[T_A^L, T_A^U], [I_A^L, I_A^U], [F_A^L, F_A^U]\}$ as an interval neutrosophic number (INN), which is a simple part of INS. Moreover, assume $X = \langle [1,1], [0,0], [0,0]\rangle$ as the largest INN, and $\emptyset = \langle [0,0], [1,1], [1,1]\rangle$ is the lowest INN.

Definition2.3. The complement of an INS $A = \langle T_A, I_A, F_A\rangle = \{[T_A^L, T_A^U], [I_A^L, I_A^U], [F_A^L, F_A^U]\}$ is meant by A^c and which is termed as $A^c = \{[F_A^L, F_A^U], [1 - I_A^U, 1 - I_A^L], [T_A^L, T_A^U$. For any $x, y \in X$, an INS 1_y and its complement $1_{X-\{y\}}$.

Definition2.4. $A = \{\langle x, T_A(x), I_A(x), F_A(x)\rangle\}$ and $B = \{\langle x, T_B(x), I_B(x), F_B(x)\rangle\}$ are dual INS, while $T_A(x) = [T_A^L(x), T_A^U(x)], I_A(x) = [I_A^L(x), I_A^U(x)], F_A(x) = [F_A^L(x), F_A^U(x)]$, and $T_B(x) = [T_B^L(x), T_B^U(x)], I_B(x) = [I_B^L(x), I_B^U(x)], F_B(x) = [F_B^L(x), F_B^U(x)]$.

Definition2.5. A and B are dual INNs, we have below-mentioned basic assets of INNs.

(1) $A \subseteq AUB, B \subseteq AUB$

(2) $A \cap B \subseteq A, A \cap B \subseteq B$

(3) $(A \cup B)^c = A^c \cap B^c$;

(4) $(A^c)^c = A$

Definition2.6. Assume X as a fixed set of objects, and R as a similarity relation on X. Represent by X/R the family of every equivalence class made by R. Obviously, X/R provides a partition of X. (X, R) is named a space of interval neutrosophic approximation. For $x \in X$, the approximations of upper and lower A have been expressed below:

$$R^-(A) = \{x \in X | [x]_R \subseteq A\}, R^+(A) = \{x \in X | [x]_R \cap A \neq \emptyset\},$$

Definition2.7. Assume X as a space of points (objects) and $C = \{C_1, C_2, \cdots, C_m\}$ be a type of sub-sets of X. If not a single element in C is empty and$\bigcup_{i=1}^m C_i = X$, then C is termed a covering of X, and (X, C) is named a space of covering approximation.

Definition2.8. Assume (X, C) be a cover approximation space. For any $x \in X$, the region of x is described as $\bigcap_{i=1}^m \{C_i \in C | x \in C_i\}$ that is represented by N_x.

Definition2.9. Assume (X, C) be a cover approximation space. For $x \in X$, the lower and upper approximations of A are expressed below:

$$C^-(A) = \{x \in X | N_x \subseteq A\}, C^+(A) = \{x \in X | N_x \cap A \neq \emptyset\}$$

Depending upon the description of the neighbourhood, the novel cover rough models can be attained.

DOI: https://doi.org/10.54216/IJNS.250319
Received: February 26, 2024 Revised: May 28, 2024 Accepted: October 05, 2024

The description of INCRS is defined, and we will also utilize a few instances for the sake of perception. Furthermore, we will give a few properties and their evidence of INCRS.

Definition3.1. Assume X as a space of points (objects). For any $[s,t] \in [0,1]$ and $C = \{C_1, C_2, \cdots, C_m\}$, while $C_i = \{T_{c,i} I_{c_i}, F_{c_i}\}$ and $C_i \in INS(i = 1, 2_{\leq} \cdots, m)$. For $\forall x \in X, \exists C_k \in C$, then $C_k(x) \geq [s,t]$, whereas $T_{C_k}(x) \geq [s,t], I_{C_k}(x) \leq [1-t, 1-s], F_{C_k}(x) \leq [1-t, 1-s]$.

Definition3.2. Let $C = \{C_1, C_2, \cdots, C_m\}$ be an interval neutrosophic $[s,t]$ cover of X. If $0 \leq [s', t'] \leq [s,t]$, C denotes an interval neutrosophic $[s', t']$ cover of X.

Definition3.3. Assume $C = \{C_1, C_2, \cdots, C_m\}$ denotes an interval neutrosophic $[s,t]$ cover of X. If $s = t = \beta$, then C is termed an interval neutrosophic β cover of X.

Definition3.4. Let $C = \{C_1, C_2, \cdots, C_m\}$ be an interval neutrosophic $[s,t]$ cover of X, while $C_i = \left\{T_{c,i} I_{c_i}, F_{c_i}\right\}$ and $C_i \in INS(i = 1,2, \cdots, m)$. For $\forall x \in X$, the interval neutrosophic $[s,t]$ region of x is expressed below:

$$N_x^{[s,t]}(y) = \cap \left\{C_i \in C \middle| T_{C_i}(x) \geq [s,t], I_{C_i}(x) \leq [1-t, 1-s], F_{C_i}(x) \leq [1-t, 1-s]\right\}$$

Definition3.5. Suppose $C = \{C_1, C_2, \cdots, C_m\}$ be an interval neutrosophic $[s,t]$ cover of X, whereas $C_i = \left\{T_{c,i} I_{c_i}, F_{c_i}\right\}$ and $C_i \in INS(i = 1,2, \cdots, m)$. When $s = t = \beta$, then the interval neutrosophic $[s,t]$ region of x is embarrassed as an interval neutrosophic β district of x.

Theorem 3.6. Assume $C = \{C_1, C_2, \cdots, C_m\}$ as an interval neutrosophic $[s,t]$ cover of X, whereas $C_i = \{T_{c,i} I_{c_i}, F_{c_i}\}$ and $C_i \in INS(i = 1,2, \cdots, m). \forall x, y, z \in X$.

Example1. Let X be the objects, with a class of elements in X signified by $x, C = \{C_1, C_2, C_3, C_4\}$ refers to an interval neutrosophic cover of X as in table 1. Set $[s,t] = [0.4,0.5]$, and it can be acquired that C refers to interval neutrosophic [0.4,0.5] cover of X.

$$N_{x_1}^{[0.4,0.5]} = C_1 \cap C_2 \cap C_3, N_{x_2}^{[0.4,0.5]} = C_1 \cap C_2 \cap C_4, N_{x_3}^{[0.4,0.5]} = C_2 \cap C_4, N_{x_4}^{[0.4,0.5]} = C_1 \cap C_2.$$

The neighborhood of interval neutrosophic [0.4, 0.5] $x_i (i = 1,2,3,4)$. The neighborhood of interval neutrosophic [0.4,0.5] $x_i (i = 1,2,3,4)$ is a cover of X.

The interval neutrosophic $[s,t]$ cover was shown in the preceding section. Depending upon this, the coverage approximation space can be gained.

Definition 3.7. Assume $C = \{C_1, C_2, \cdots, C_m\}$ Is an interval neutrosophic $[s,t]$ cover of X, whereas $C_i = \left\{T_{c_i} I_{c_i}, F_{c_i}\right\}$ and $C_i \in INS(i = 1,2, \cdots, m)$. Then (X, C) is named an interval neutrosophic $[s,t]$ cover approximation space.

Definition3.8. Suppose (X, C) is an interval neutrosophic $[s,t]$ cover approximation space, for any $A \in INS$, the approximation operator of upper $C^{[s,t]}(A)$ and the lower $\underline{C}^{[s,t]}(A)$ of interval, neutrosophic A are definite below: $\underline{C}^{[s,t]}(A) = \left\{T_{\underline{C}^{[s,t]}(A)}, I_{\underline{C}^{[s,t][s,t]}(A)}, F_{\underline{C}^{[s,t]}(A)}\right\}, C^{[s,t]}(A) = \left\{T_{\underline{C}^{[s,t]}(A)}, I_{\underline{C}^{[s,t]}(A)}, F_{\underline{C}^{[s,t][s,t]}(A)}\right\}$, whereas

$$T_{\underline{C}^{[s,t]}(A)} = \wedge \{T_A(y) \vee F_{N_x^{[s,t]}}(y) | y \in X\}, I_{\underline{C}^{[s,t]}(A)} = \vee \left\{I_A(y) \wedge \left([1,1] - I_{N_x^{[s,t]}}(y)\right) \middle| y \in X\right\},$$

$$F_{\underline{C}^{[s,t]}(A)} = \vee \{F_A(y) \wedge T_{N_x^{[s,t]}}(y) | y \in X\}, T_{\bar{C}^{[s,t]}(A)} = \vee \left\{T_A(y) \wedge T_{N_x^{[s,t]}}(y) \middle| y \in X\right\},$$

$$I_{\bar{C}^{[s,t]}(A)} = \wedge \{I_A(y) \vee I_{N_x^{[s,t]}}(y)) T | y \in X\}, F_{\bar{C}^{[s,t]}(A)} = \wedge \{F_A(y) \vee F_{N_x^{[s,t]}}(y) | y \in X\}.$$

D. Stage IV: MSIBOA-based Parameter Tuning

Finally, the MSIBOA is exploited for the optimum hyperparameter adjustment of the INCRS classifier. The effectual optimization ability of MSIBOA facilitates accurate and quick determination of super-parameters [20]. BOA is a new meta-heuristic approach depends on the foraging behavior of butterflies. Using Eq. (7), the algorithm randomly generates a butterfly position and enables the transition of search manners during the iteration.

DOI: https://doi.org/10.54216/IJNS.250319
Received: February 26, 2024 Revised: May 28, 2024 Accepted: October 05, 2024

$$\begin{cases} X = lb + r \times (ub - lb) \\ x_i^{t+1} = x_i^t + \left(r^2 \times x_j^t - x_k^t\right) \times f_i \,, r \geq P \\ x_i^{t+1} = x_i^t + (r^2 \times g^* - x_i^t) \times f_i \,, r < P \end{cases} \quad (7)$$

BOA outperforms other similar techniques, still, it suffer from inadequate local optimization, early convergence, and limited global search capability. To overcome this issue, the reference proposes three strategies that were introduced in BOA:

1) Integrating the Hénon mapping initialization method. The Hénon map is a 2D discrete chaotic map that generates a uniform distribution sequence. Initializing the population with the Hénon chaotic map assists in preventing local optima, reducing search stagnation, and improving population diversity. The mathematical modelling of the Hénon map is given below:

$$\begin{cases} x(k) = 1 - \gamma(x(k-1))^2 + y(k-1) \\ y(k) = \varphi x(k-1) \\ Z_i^k = \dfrac{x_i^k - \varphi_i}{\gamma_i - \varphi_i}, j = 1,2, \cdots, n \end{cases} \quad (8)$$

$$x_i^t = lb_i + z_i^k(ub_i - lb_i) \quad (9)$$

Where $r; \varphi; x(0) = 0; y(0) = 0; k$ refers to the chaotic iteration number. n chaotic variable z_i^k is produced by Eq. (8), and the individual search space x_i^t variable is attained using inverse mapping.

2) Introduce a feedback-sharing model into the search formula, integrating the feedback factor and sharing coefficient. Negative or positive feedbacks are generated by comparing and calculating the fitness value of the butterfly. Then, the optimal position or a random individual position is updated.

$$\alpha = (F_{\max} - F_{\min})r + F_{\min} \quad (10)$$

$$\theta = \frac{1}{1 + \exp(-\lambda)} \quad (11)$$

$$x_i^t = \begin{cases} x_i^t + (g^* - x_i^t)\alpha + (x_h^t - x_i^t)\alpha + \theta, F(x_h^t) < F(x_i^t) \\ x_i^t + (g^* - x_i^t)\alpha + \theta, F(x_h^t) \geq F(x_i^t) \end{cases} \quad (12)$$

Where α represents the sharing coefficient; $Fmax$ and $Fmin$ are the fitness ranges of the population. θ refers to the feedback factor, λ denotes the existing iteration count; x_h^t shows the location of random butterflies; F denotes the computation fitness.

3) The group synergistic effect location updating method is introduced. The golden section number was utilized to enhance the distance formula using the golden sine algorithm as a medium, and population coordination factor $\psi_2 \sin\psi_1$and cooperative individual correction factor $|\sin\psi_1|$ are introduced to guide population levels, population search direction, and individual flight adjustment correspondingly.

$$\begin{cases} l_1 = -\pi + 2\pi(1 - \tau) \\ l_2 = -\pi + 2\pi\tau \\ \tau = \dfrac{\sqrt{5} - 1}{2} \end{cases} \quad (13)$$

$$x_i^{t+1} = x_i^t |\sin \psi_1| + \left|l_1 x_j^t - l_2 x_i^t\right| f_i \psi_2 \sin\psi_1 \quad (14)$$

$$x_i^{t+1} = x_i^t |\sin\psi_1| + |l_1 g^* - l_2 x_i^t| f_i \psi_2 \sin\psi_1 \quad (15)$$

Where τ denotes the golden ratio; l1 and l2 indicate the distance renewal coefficients. The random number of $\psi_1 \in [0_7\tau], \psi_2 \in [0,2\pi]$. The fitness selection is the substantial factor manipulating the efficiency of MSIBOA. The hyperparameter range method includes the solution-encoded process to evaluate the efficiency of the candidate solution.

$$Fitness \;=\; \max{(P)} \quad (16)$$

DOI: https://doi.org/10.54216/IJNS.250319
Received: February 26, 2024 Revised: May 28, 2024 Accepted: October 05, 2024

$$P = \frac{TP}{TP + FP} \quad (17)$$

Where TP signifies the true positive and FP signifies the false positive value.

4. Experimental Result and Analysis

The performance evaluation of the EDCINCRS-MSIBO model is studied under two databases namely German credit dataset (GCD)[21] and Australian credit dataset (ACD) [22]. Firstly, the GCD has 1000 instances with 2 classes. Next, the ACD includes 690 instances with 2 classes as represented in Table 1. Table 2 provides details of selected attributes.

Table 1: Details of Database

Datasets	Total Instances	Total Features	Class labels	Financial Crisis/Non-Financial Crisis
GCD	1000	24	2	300/700
ACD	690	14	2	383/307

Table 2: Selected features of GCD and ACD

Dataset	Selected Attributes
GCD	3,5,6,7,9,12,13,15,17,18,19,22,23
ACD	2,4,5,8,9,10,11,12

Table 3 signifies the classifier results of the EDCINCRS-MSIBO system on GCD. The outcomes suggest that the EDCINCRS-MSIBO approach correctly recognized the samples. With 70%TRAPH, the EDCINCRS-MSIBO system offers average $accu_y$, $prec_n$, $sens_y$, $spec_y$, and $F1_{score}$ of 94.43%, 95.19%, 91.91%, 91.91%, and 93.32%, similarly. Also, with 30%TESPH, the EDCINCRS-MSIBO methodology offers average $accu_y$, $prec_n$, $sens_y$, $spec_y$, and $F1_{score}$ of 96.33%, 97.63%, 93.04%, 93.04%, and 95.04%, correspondingly.

Table 3: Classifier outcome of EDCINCRS-MSIBO model on GCD

GCD					
Class Labels	$\boldsymbol{Accu_y}$	$\boldsymbol{Prec_n}$	$\boldsymbol{Sens_y}$	$\boldsymbol{Spec_y}$	$\boldsymbol{F1_{score}}$
70% TRAPH					
Financial Crisis	94.43	96.91	85.07	98.75	90.60
Non-Financial Crisis	94.43	93.48	98.75	85.07	96.04
Average	94.43	95.19	91.91	91.91	93.32
30% TESPH					
Financial Crisis	96.33	100.00	86.08	100.00	92.52
Non-Financial Crisis	96.33	95.26	100.00	86.08	97.57
Average	96.33	97.63	93.04	93.04	95.04

DOI: https://doi.org/10.54216/IJNS.250319
Received: February 26, 2024 Revised: May 28, 2024 Accepted: October 05, 2024

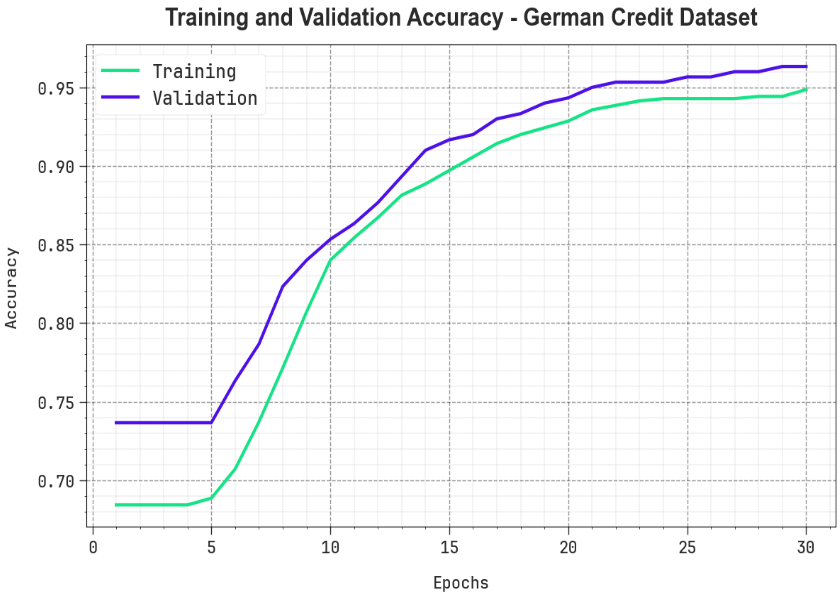

Figure 3. $Accu_y$ Curve of EDCINCRS-MSIBO model on GCD

In Fig. 3, the training $accu_y$ (TRAAY) and validation $accu_y$ (VLAAY) analysis of the EDCINCRS-MSIBO technique on GCD is displayed. The $accu_y$ values are computed over the range of 0-30 epochs. The outcome highlighted that the TRAAY and VLAAY curves show a rising tendency which informed the capacity of the EDCINCRS-MSIBO method with higher performance over multiple iterations.

In Fig. 4, the TRA loss (TRALO) and VLA loss (VLALO) values of the EDCINCRS-MSIBO methodology on GCD are shown. The loss graph is computed over an interval of 0-30 epochs. It is denoted that the TRALO and VLALO analysis exemplifies a reduced tendency, notifying the capacity of the EDCINCRS-MSIBO system in balancing a trade-off between generalization and data fitting.

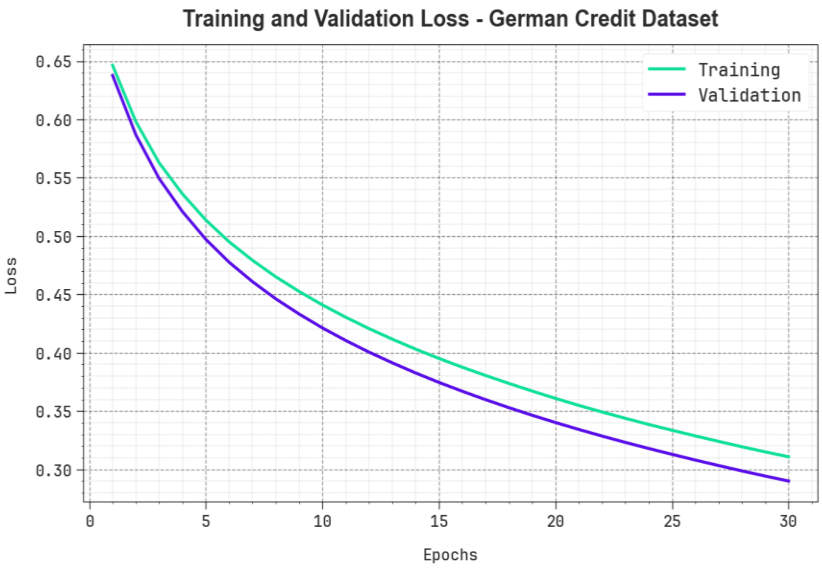

Figure 4. Loss curve of EDCINCRS-MSIBO system on GCD

Table 4 denotes the classifier results of the EDCINCRS-MSIBO method on ACD. The results suggest that the EDCINCRS-MSIBO system correctly identified the samples. With 70%TRAPH, the EDCINCRS-MSIBO approach offers average $accu_y$, $prec_n$, $sens_y$, $spec_y$, and $F1_{score}$ of 96.48%, 96.46%, 96.41%, 96.41%, and 96.43%, consistently. Also, with 30%TESPH, the EDCINCRS-MSIBO methodology offers average $accu_y$, $prec_n$, $sens_y$, $spec_y$, and $F1_{score}$ of 95.17%, 95.35%, 94.92%, 94.92%, and 95.10%, respectively.

DOI: https://doi.org/10.54216/IJNS.250319
Received: February 26, 2024 Revised: May 28, 2024 Accepted: October 05, 2024

Table 4: Classifier result of EDCINCRS-MSIBO model on ACD

Australian Credit Dataset					
Class Labels	$Accu_y$	$Prec_n$	$Sens_y$	$Spec_y$	$F1_{score}$
TRAPH (70%)					
Financial Crisis	96.48	96.67	97.03	95.79	96.85
Non-Financial Crisis	96.48	96.24	95.79	97.03	96.02
Average	96.48	96.46	96.41	96.41	96.43
TESPH (30%)					
Financial Crisis	95.17	94.07	97.37	92.47	95.69
Non-Financial Crisis	95.17	96.63	92.47	97.37	94.51
Average	95.17	95.35	94.92	94.92	95.10

In Fig. 5, the TRAAY and VLAAY results of the EDCINCRS-MSIBO method on ACD are exhibited. The $accu_y$ values are computed over the range of 0-30 epochs. The outcome highlights that the TRAAY and VLAAY graphs illustrate a growing tendency which informed the capacity of the EDCINCRS-MSIBO methodology with higher performance over multiple iterations.

In Fig. 6, the TRALO and VLALO values of the EDCINCRS-MSIBO methodology on ACD are depicted. The loss analysis is calculated over an interval of 0-30 epochs. It is denoted that the TRALO and VLALO graphs exemplify a diminishing tendency, reporting the capacity of the EDCINCRS-MSIBO system to balance a trade-off between data fitting and generalized.

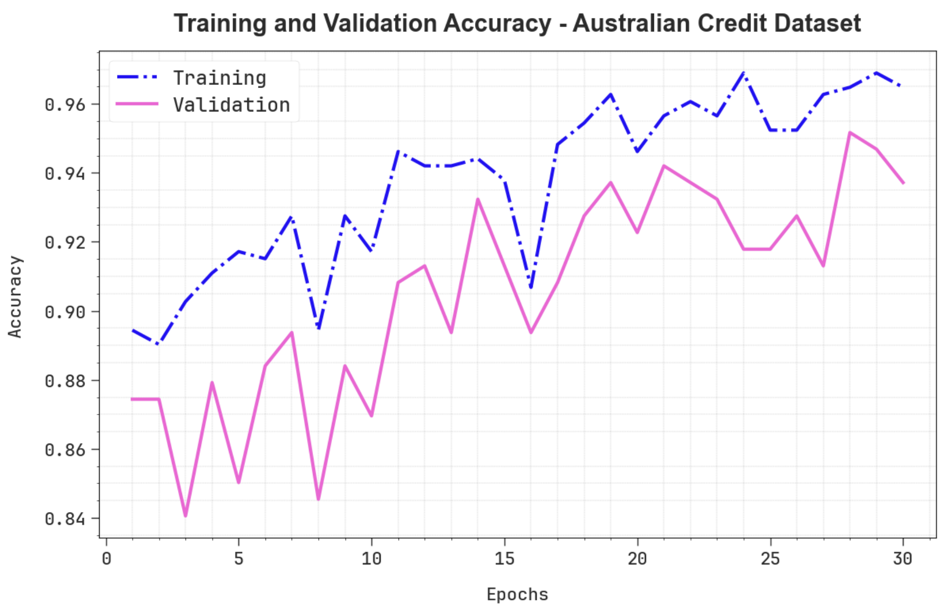

Figure 5. $Accu_y$ curve of EDCINCRS-MSIBO model on ACD

DOI: https://doi.org/10.54216/IJNS.250319
Received: February 26, 2024 Revised: May 28, 2024 Accepted: October 05, 2024

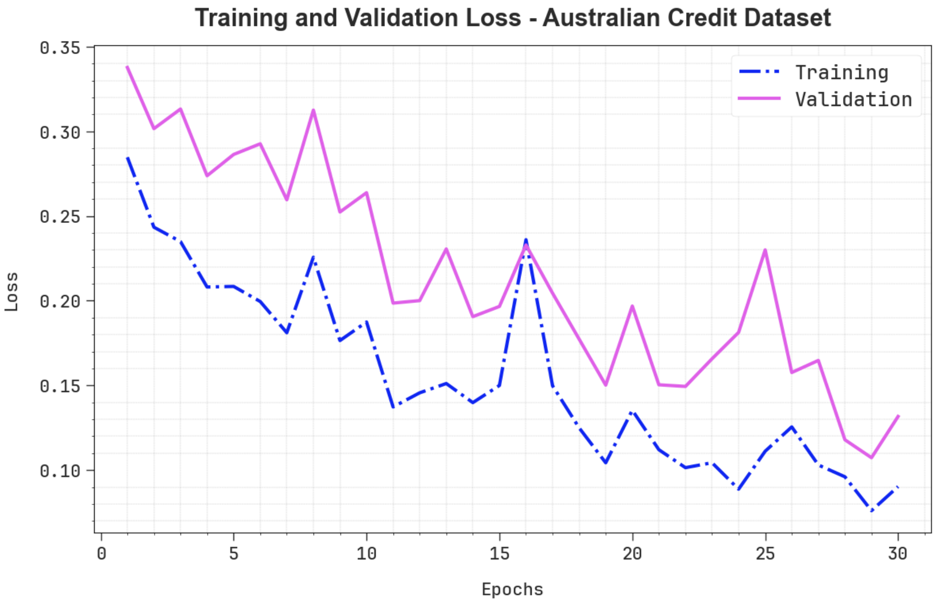

Figure 6. Loss curve of EDCINCRS-MSIBO model on ACD

Table 5 inspects the comparison results of the EDCINCRS-MSIBO methodology with existing models on GCD. The results highlighted that the QABO-LSTM-RNN model, LSTM-RNN approach, ACO classifier, MLP algorithm, SVM technique, and AdaBoost algorithm have reported inferior performance. Meanwhile, HHPODL-FCP methodology has gained closer outcomes. Additionally, the EDCINCRS-MSIBO algorithm reported higher performance with maximum $sens_y$, $spec_y$, $accu_y$, and $F1_{score}$ of 93.04%, 93.04%, 96.33%, and 95.04%, respectively.

Table 5: Comparative values of EDCINCRS-MSIBO method with existing methods on GCD

German Credit				
Technique	$Sens_y$	$Spec_y$	$Accu_y$	$F1_{Score}$
EDCINCRS-MSIBO	93.04	93.04	96.33	95.04
HHPODL-FCP	92.16	92.54	94.93	93.75
QABO-LSTM-RNN	87.23	92.60	91.99	90.11
LSTM-RNN	82.22	88.58	84.58	88.74
ACO Methodology	78.32	69.32	75.79	85.42
Multilayer Perceptron	73.95	66.96	71.04	75.22
Support Vector Machine	72.80	66.49	71.26	71.84
AdaBoost Model	71.47	61.41	67.60	71.35

Table 6 studies the comparison results of the EDCINCRS-MSIBO algorithm with existing systems on ACD. The results emphasized that the QABO-LSTM-RNN model, LSTM-RNN approach, ACO classifier, MLP algorithm, SVM technique, and AdaBoost algorithm have reported diminished performance. While HHPODL-FCP technique has accomplished closer outcomes. Besides, the EDCINCRS-MSIBO algorithm reported superior performance with highest $sens_y$, $spec_y$, $accu_y$, and $F1_{score}$ of 96.41%, 96.41%, 96.48%, and 96.43%, similarly.

DOI: https://doi.org/10.54216/IJNS.250319
Received: February 26, 2024 Revised: May 28, 2024 Accepted: October 05, 2024

Table 6: Comparative results of EDCINCRS-MSIBO technique with existing algorithms on ACD

Australian Credit				
Classifiers	$Sens_y$	$Spec_y$	$Accu_y$	$F1_{Score}$
EDCINCRS-MSIBO	96.41	96.41	96.48	96.43
HHPODL-FCP	94.37	94.62	95.12	94.69
QABO-LSTM-RNN	90.96	93.22	93.39	94.70
LSTM-RNN	86.04	93.08	93.08	91.66
ACO Methodology	79.93	89.48	89.83	82.12
Multilayer Perceptron	76.67	84.59	84.59	78.89
Support Vector Machine	71.30	75.87	76.31	77.57
AdaBoost Model	69.27	67.47	67.85	68.63

5. Conclusion

In this manuscript, we focus on the development of the EDCINCRS-MSIBO Algorithm for FinTech Applications. It contains distinct kinds of stages such as data normalization, feature selection, classification, and parameter tuning. In the EDCINCRS-MSIBO technique, a min-max normalization-based data pre-processing model to scale the raw data into a uniform format. For feature subset selection, the WOA is utilized to reduce the dimensionality and improve model efficiency by selecting the most relevant features. In addition, the EDCINCRS-MSIBO technique takes place INCRS classifier is utilized for detection and classification of financial crisis. Finally, the MSIBOA is exploited for the optimal hyperparameter adjustment of the INCRS model. To ensure the enhanced predictive performance of the EDCINCRS-MSIBO model, extensive simulations are conducted on two benchmark databases. The comparative result analysis displays the promising performance of the EDCINCRS-MSIBO method on the recent techniques.

Funding: “This research received no external funding”

Conflicts of Interest: “The authors declare no conflict of interest.”

References

[1] Almuhur, E., Miqdad, H., Al-labadi, M. and Idrisi, M.I., 2024. µ-L-Closed Subsets of Noetherian Generalized Topological Spaces. International Journal of Neutrosophic Science, 23(3), pp.148-48.

[2] Tashtemirovich, A.O., Balba, M.E., Ibrohimjon, F. and Batirova, N., Investigating the Impact of Artificial Intelligence on Digital Marketing Tactics Strategies Using Neutrosophic Set.

[3] Sivakumar, C., Al-Qadri, M.O., Alsaraireh, A.A., Al-Husban, A., Meenakshi, P.M., Rajesh, N. and Palanikumar, M., 2024. q-rung square root interval-valued neutrosophic sets with respect to aggregated operators using multiple attribute decision making. International Journal of Neutrosophic Science, 23(3), pp.154-54.

[4] Gharib, M., Fakhry, A.E., Ali, A.M., Abdelhafeez, A. and Elbehiery, H., 2024. Single Valued Neutrosophic Sets for Analysis Opinions of Customer in Waste Management. International Journal of Neutrosophic Science, 23(3), pp.184-84.

[5] Ali Alqazzaz. "Integrated Neutrosophic methodology and Machine Learning Models for Cybersecurity Risk Assessment: An exploratory study." International Journal of Neutrosophic Science, Vol. 23, No. 3, 2024 ,PP. 195-207.

[6] Balmaseda, V.; Coronado, M.; de Cadenas-Santiagoc, G. Predicting Systemic Risk in Financial Systems Using Deep Graph Learning. Intell. Syst. Appl. 2023, 19, 200240.

DOI: https://doi.org/10.54216/IJNS.250319
Received: February 26, 2024 Revised: May 28, 2024 Accepted: October 05, 2024

[7] Uthayakumar, J.; Metawa, N.; Shankar, K.; Lakshmanaprabu, S.K. An intelligent hybrid model for financial crisis prediction using machine learning techniques. Inf. Syst. e-Bus. Manag. 2020, 18, 617–645.

[8] Khuwaja, P.; Khowaja, S.A.; Dev, K. Adversarial learning networks for fintech applications using heterogeneous data sources. IEEE Internet Things J. 2021, 10, 2194–2201.

[9] S. K. S. Tyagi and Q. Boyang, "An intelligent internet of things aided financial crisis prediction model in FinTech," IEEE Internet of Things Journal, vol. 10, pp. 1, 2021.

[10] J. Uthayakumar, N. Metawa, K. Shankar and S. K. Lakshmanaprabu, "Financial crisis prediction model using ant colony optimization," International Journal of Information Management, vol. 50, no. 5, pp. 538–556, 2020.

[11] Osadchy, E., Abdullayev, I., Bakhvalov, S., Klochko, E. and Tagibova, A., 2024. Jellyfish Search Algorithm Based Feature Selection with Optimal Deep Learning for Predicting Financial Crises in the Economy and Society. Full Length Article, 14(2), pp.186-86.

[12] Katib, I., Assiri, F.Y., Althaqafi, T., AlKubaisy, Z.M., Hamed, D. and Ragab, M., 2023. Hybrid Hunter–Prey Optimization with Deep Learning-Based Fintech for Predicting Financial Crises in the Economy and Society. Electronics, 12(16), p.3429.

[13] Wu, C., Jiang, C., Wang, Z. and Ding, Y., 2024. Predicting financial distress using current reports: A novel deep learning method based on user-response-guided attention. Decision Support Systems, 179, p.114176.

[14] Kalaivani, R. and Saravanan, A., 2023, November. Exploiting pattern recognition using chimp optimization algorithm with machine learning for financial crisis prediction. In 2023 International Conference on Sustainable Communication Networks and Application (ICSCNA) (pp. 944-950). IEEE.

[15] Chen, S., Huang, Y. and Ge, L., 2024. An early warning system for financial crises: a temporal convolutional network approach. Technological and Economic Development of Economy, 30(3), pp.688-711.

[16] Muthukumaran, K., Hariharanath, K. and Haridasan, V., 2023. Feature Selection with Optimal Variational Auto Encoder for Financial Crisis Prediction. Computer Systems Science & Engineering, 45(1).

[17] Shantal, M., Othman, Z. and Bakar, A.A., 2023. A novel approach for data feature weighting using correlation coefficients and min–max normalization. Symmetry, 15(12), p.2185.

[18] Song, W., Han, X. and Qi, J., 2023. Prediction of Gas Emission in the Working Face Based on LASSO-WOA-XGBoost. Atmosphere, 14(11), p.1628.

[19] Xu, D., Xian, H. and Lu, X., 2021. Interval neutrosophic covering rough sets based on neighborhoods. Infinite Study.

[20] Wang, R., Dong, Y. and Lu, J., 2024. Short-term Wind Power Forecasts based on VMD-KPCA and CFSBOA-BiLSTM. International Journal of Electric Power and Energy Studies, 2(1), pp.39-55.

[21] https://archive.ics.uci.edu/ml/datasets/statlog+(german+credit+data)

[22] http://archive.ics.uci.edu/ml/datasets/statlog+(australian+credit+approval)

DOI: https://doi.org/10.54216/IJNS.250319
Received: February 26, 2024 Revised: May 28, 2024 Accepted: October 05, 2024