Bipolar Fuzzy Hypersoft Set with Heuristic Search Based Customer Retention Prediction Model in Financial Sectors

Alexander Kalinin[1,*], Inomjon Yusubov[2], Tatiana Yakubova[3], Victoria Kruglyakova[4], Tatyana Khorolskaya[5]

[1]Department of Valuation and Corporate Finance, Moscow University for Industry and Finance "Synergy", Moscow, 127015, Russia

[2]Department of Economics, Urgench State University, Urgench, 220100, Uzbekistan

[3]Department of Management, RUDN University, Moscow, 117198, Russia

[4]Department of Construction Organization and Real Estate Management, Moscow State University of Civil Engineering, Moscow, 129337, Russia

[5]Department of Money Circulation and Credit, Kuban State Agrarian University named after I.T. Trubilin, Krasnodar, 350044, Russia

Emails: arkalinin@synergy.ru; inomjon.y@urdu.uz; yakubova-tn@rudn.ru; kruglyakovavm@mgsu.ru; xorolskaya.t@edu.kubsau.ru

Abstract

From a philosophical viewpoint, the theory of neutrosophic set (NS) is a simplification of the concept of Fuzzy Set (FS) and intuitionistic FS (IFS). An NS is illustrated by a truth, an indeterminacy, and a falsity membership functions and every membership degree is an actual standard or a non-standard sub-set of the non-standard unit range of] −0, 1+ [. Customer churn is when clients stop utilizing a company's service or product. Moreover, it is also named customer retention, which is vastly significant metric as it is much less costly to keep the existing customers than to obtain novel customers. The prediction of churn plays an essential part in customer retention because it forecasts clients who are in danger of leaving the organization. In the banking sector, the customer attrition arises when clients quit utilizing the services and goods provided by the bank for some time. So, customer churn is vital in today's economic banking industry. This study proposes a Leveraging Bipolar Fuzzy Hypersoft Set with Heuristic Optimization Algorithms-based Customer Retention Prediction (BFHSS-HOACRP) technique in financial sectors. The BFHSS-HOACRP technique applies optimized techniques to predict the customer retention behavior in the industry of bank. Initially, the mean normalization technique is utilized in the data pre-processing stage to prepare raw data into a suitable format for analysis and modeling. For the selection of feature process, the grasshopper optimization algorithm (GOA) method is employed to identify and select the most relevant features from an input data. In addition, the proposed BFHSS-HOACRP technique implements bipolar fuzzy hypersoft set (BFHSS) method for the classification process. Additionally, the spider monkey optimization (SMO)-based hyperparameter selection process is performed to optimize the classification results of BFHSS model. The efficacy of the BFHSS-HOACRP approach is examined under the bank customer churn prediction dataset. The comparison analysis of the BFHSS-HOACRP approach portrayed a superior accuracy value of 95.41% over existing techniques.

Keywords: Bipolar Fuzzy Hypersoft Set; Customer Retention Prediction; Fuzzy Sets; Neutrosophy; Spider Monkey Optimization; Feature Selection

1. Introduction

For researchers, dealing with inconsistency and uncertainty was a very significant issue in mathematical models. In mathematical models, scholars have presented several estimations for difficulties in inconsistency and uncertainty data [1]. A few of the famous estimations are the fuzzy set (FS) theory introduced and intuitionistic

DOI: https://doi.org/10.54216/IJNS.260305

FS (IFS) theory proposed. FS is classified as membership functionality and an intuitionistic fuzzy set is classified as non-membership and membership functionalities [2]. However, IFS and FS do not manage inconsistent and uncertain pieces of information. Hence, neutrosophic set theory was presented as an overview of IFS and FS depends on Neutrosophy, which is a division of philosophy [3]. Customer retention, otherwise called customer churn, is where customers dismiss their connection with an organization or business. With the background of banking, customer retention takes place when customers close their accounts or stop using the facilities of a specific bank. Efficient handling and understanding of customer churn play a vital role in banks maintaining financial stability and safeguarding bank reputation. The financial effect of customer churn on banks has been important, which results in the possibility of revenue loss within numerous banking services [4].

As a result, nurturing and establishing long-term customer relationships is very crucial for banks. By acquiring vision within abrasion designs, banks will identify customers when they leave the bank and execute plans for retaining them. These methodologies improve the general lifetime value of the customer and boost the profitability of the bank [5]. Additionally, customer churn has consequences on the brand perception and reputation of the bank. Higher churn ratings frequently show fundamental problems namely inefficient processes, lack of competitive products and services, or poor customer experiences. Hence, managing and understanding customer abrasion is very important for banks in addressing those challenges and enhancing their general customer experience [6]. Customer retention analysis outcome has a very high effect on the policy of the bank. The outcomes of churn analysis permit banks to improve their strategies for identifying new customers or improving current customer satisfaction [7]. Moreover, banks play a major role in the financial growth and progression of the country. Therefore, the banking sector is an important aspect of the people's and countries financial stability [8]. Since it is impossible all the time to acquire new customers in this competitive world. The main objective of the bank is to safeguard the current customers and retain them [9]. With the development of technologies customer, churn analysis research states that employing deep learning (DL) and machine learning (ML) become very popular in recent times. With the use of DL and ML, one can measure the performance of the financial sector in several methods [10].

1.1. Research Motivation for Enhancing Customer Retention in the Banking Sector

The banking industry encounters crucial threats in maintaining customer loyalty, particularly with the growing competition and availability of alternative services. High customer churn can have a detrimental impact on a financial stability of the bank, as retaining existing customers is often more cost-effective than acquiring new ones. The significance of effectual customer retention strategies has never been more pronounced, specifically as customers become more discerning and are drawn to competitor's present improved products and services. Understanding the key drivers of customer attrition is crucial for banks looking to stay ahead in a competitive marketplace. This research is motivated by the requirement to develop enhanced models that can predict customer churn, ensuring that banks can take proactive measures to prevent losses.

The motivation for this research is driven by the following key factors:

- **Impact of Customer Churn:** High customer churn rates negatively affect the profitability and market position of a bank, highlighting the requirement for effectual retention strategies.
- **Advancements in ML and DL:** ML and DL models give crucial potential to improve churn prediction accuracy, giving banks with more reliable forecasts.
- **Personalized Services:** There is a growing requirement for banks to give tailored services based on data-driven insights to enhance customer satisfaction and loyalty.
- **Handling Data Inconsistency and Uncertainty:** Addressing the challenges posed by data inconsistency and uncertainty in customer behaviour modelling is substantial for precise churn prediction.
- **Demand for Predictive Analytics:** The growing demand for predictive analytics tools to assist in decision-making underscores the significance of using advanced technologies to improve customer retention efforts.

These factors collectively highlight the requirement for more robust customer retention models, which can assist the banks, mitigate churn and enhance customer loyalty in a competitive market.

1.2. Contribution of the Proposed BFHSS-HOACRP Model

This study proposes a Leveraging Bipolar Fuzzy Hypersoft Set with Heuristic Optimization Algorithms-based Customer Retention Prediction (BFHSS-HOACRP) technique in financial sectors. The BFHSS-HOACRP technique applies optimized techniques to predict the customer retention behavior in the industry of bank. Initially, the mean normalization technique is utilized in the data pre-processing stage to prepare raw data into a suitable format for analysis and modeling. For the selection of feature process, the grasshopper optimization algorithm (GOA) method is employed to identify and select the most relevant features from an input data. In addition, the proposed BFHSS-HOACRP technique implements bipolar fuzzy hypersoft set (BFHSS) method for the classification process. Additionally, the spider monkey optimization (SMO)-based hyperparameter selection

DOI: https://doi.org/10.54216/IJNS.260305

process is performed to optimize the classification results of BFHSS model. The efficacy of the BFHSS-HOACRP approach is examined under the bank customer churn prediction dataset. The major contribution of the BFHSS-HOACRP approach is listed below.

The BFHSS-HOACRP technique utilizes mean normalization to scale all input features uniformly, improving the consistency of data preprocessing. This step assists in ensuring that the model is less sensitive to discrepancies in feature ranges, enhancing its performance. By standardizing the data, the model is better equipped to handle diverse inputs, resulting in predictions that are more accurate.

The BFHSS-HOACRP method employs the GOA approach for feature selection, effectually detecting the most relevant features. This mitigates dimensionality, allowing the model to concentrate on key inputs and improve computational efficiency. By choosing the optimal features, the model improves both accuracy and processing speed during training.

The BFHSS-HOACRP approach implements the BFHSS model for classification, effectually addressing uncertainty and ambiguity in the data. This approach improves the capability of the method to make accurate decisions even in complex, imprecise scenarios. By integrating BFHSS, the model enhances its robustness and reliability in classification tasks.

The BFHSS-HOACRP model uses the SMO approach for model tuning, enabling effectual exploration of the parameter space to find optimal settings. This approach improves the performance of the model by fine-tuning parameters for better accuracy. The utilization of SMO confirms faster convergence and improved overall results during the optimization process.

The BFHSS-HOACRP methodology integrates GOA-based feature selection, BFHSS-based classification, and SMO-based tuning to provide a distinctive solution that handles both data uncertainty and optimization threats in parallel. This incorporates improves the capability of the model to handle complex, uncertain data while optimizing its parameters for superior performance. The novelty is in the simultaneous application of these techniques, giving a comprehensive approach that significantly enhances prediction accuracy in complex tasks.

2. Literature Review

In [11], intends to forecast the migration and attrition of customers who employ internet-banking services utilizing tree-based classifier. The author employed numerous feature optimizer models in tree-based classifiers for forecasting behaviour of customer churn employing transaction data from customers who utilize internet-banking services. Primarily, feature reduction is performed to extract ineffectual aspects and intend the dataset onto a lower-dimensional space. Afterward, the novelist implemented feature significance to give a score to every input feature. In [12], a domain-order-based shallow fusion DL technique, the FCL-CNN-LSTM is intended and implemented as a decision support method for evaluating the risk of customer churn. The presented model is dissimilar to prior analysis resolved the review of unbalanced data and selecting features individually, but resolved either one concern to offer the model of feature selection under unbalanced data. Li and Yan [13] focusing on balancing the data employ sampling methodologies to predict whether a customer will churn, integrate the ML approach to develop a comprehensive CCP methodology, and choose the method with the better performance. Eventually, the causal effects of these variables on CC are investigated by leveraging the R-learner causal inference approach. Panimalar et al. [14] presented an innovative methodology known as MBP-WMLP abbreviated as Multi-path Back Propagation with Weighted MLP, to enhance precision of prediction in telecommunication networks. The relevant method employs a specialized framework of MLP with weights, intended for handling the complexity of telecommunications data. Besides, an enhanced multi-path backpropagation approach was introduced to increase the convergence and optimization of model training.

In [15] primary objective of this method is to examine the varied ML models that are essential to progressing CCP techniques and recognize churns to offer them with retention plans and approaches. During this method, leave subscriptions to gather information about customers by implementing classification models namely ML and RF methods like decision trees (DT) and KNN. It provides an effective business model, which inspects customer churn data and delivers precise predictions of churn customers, thus business administration can take action during the churn period to stop churn along with loss in profit. Pathak et al. [16] projected a strong solution: Personalized recommendation and CCP. Over dual-stage techniques, banks can proactively predict and reduce customer departures by utilizing the ability to advanced analytics. The initial stage requires to assessment of historical data to precisely forecast churn, thus aiding the implementation of targeted initiatives focused on improving customer experiences and rate of curbing attrition. In another stage, the attention shifts to personalized suggestions powered by Explainable AI (XAI), guaranteeing that the decision models are understandable and transparent. The author [17] concentrates on structuring an arithmetic optimizer algorithm (AOA) with stacked bi-LSTM (SB-LSTM) technique for CCP. The projected AOA-SBLSTM technique aims to effectively predict the existence of CC in the

DOI: https://doi.org/10.54216/IJNS.260305

telecommunication industry. Primarily, the presented approach conducts pre-processing by changing the original data to valuable format. Moreover, the SBLSTM approach is utilized for classifying data into non- and churners.

2.1. Limitations of Existing Customer Churn Prediction Models and Their Challenges

The existing studies on CCP have made significant strides in improving prediction accuracy, but they mostly encounter threats in handling large-scale and unbalanced datasets, which can result in model overfitting and inaccurate predictions. Additionally, many techniques face difficulty to maintain robustness in real world, noisy data environments, where diverse external factors can heavily influence the churn prediction outcome. Another limitation is the reliance on conventional ML models, which may not be able to capture complex non-linear relationships and dependencies in data. Furthermore, while feature selection is a crucial component of most models, many approaches fail to adequately account for feature interdependencies, which can restrict model performance. The focus on individual aspects of customer churn, without incorporating holistic strategies, also lessens the comprehensive nature of existing models.

2.2. Research Gaps in Advancing Customer Churn Prediction Techniques and Methodologies

1. There is a requirement for more advanced techniques that can effectually manage imbalanced datasets and enhance prediction accuracy without overfitting.
2. Existing techniques often lack the capacity to capture non-linear relationships in data, which restricts their capability to model real-world complexity effectively.
3. Existing approaches fail to incorporate a holistic strategy for customer churn, often concentrating on isolated features rather than a comprehensive view of all possible contributing factors.
4. There is a gap in utilizing real-time or streaming data in churn prediction, with many models still depending on static datasets that do not reflect dynamic changes in customer behaviour.
5. Many models overlook the interpretability and transparency of decision-making processes, which could affect the trust and acceptance of the model in practical applications.
6. Few studies have explored the integration of DL methods with conventional ML techniques, which could potentially improve predictive performance and generalizability across industries.

3. The Proposed Method

In this study, we presented a BFHSS-HOACRP method in Financial Sectors. The BFHSS-HOACRP technique applies optimized techniques to predict the customer retention behaviour in the bank industry. It includes 4 steps namely pre-processing, feature selection using GOA, BFHSS-based classification process, and hyperparameter tuning. Fig. 1 represents the entire flow of BFHSS-HOACRP model.

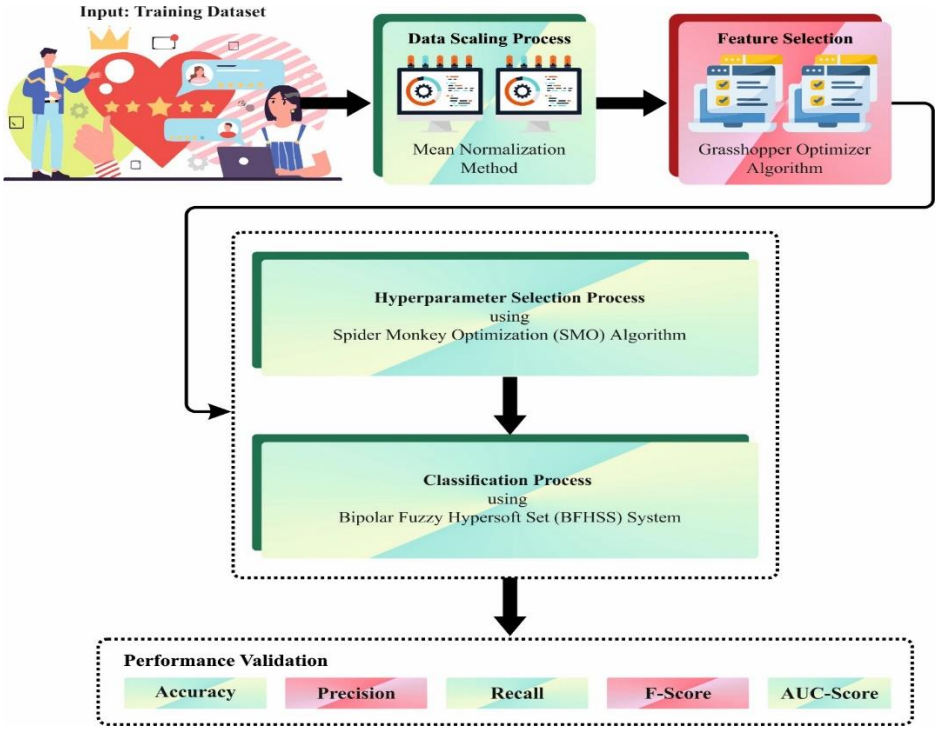

Figure 1. Overall flow of BFHSS-HOACRP model

DOI: https://doi.org/10.54216/IJNS.260305

A. Data Pre-processing

At first, the mean normalization technique has been deployed in the data pre-processing stage to prepare raw data into a suitable format for analysis and modelling [18]. This is a highly effectual technique for scaling features to a common range, typically between -1 and 1. By centering the data on zero, this method assists in eliminating the bias introduced by features with large or small values, ensuring that all features contribute equally to the learning process of the model. Compared to other preprocessing methods such as min-max normalization or standardization, mean normalization is specifically beneficial when the data contains outliers, as it mitigates their impact on the model. This technique is also simple to implement and computationally efficient. It is particularly useful in scenarios where models like neural networks or distance-based algorithms (e.g., KNN) are employed, as it ensures faster convergence during training. Moreover, mean normalization improves model accuracy and stability, making it an ideal choice over more complex normalization methods in many real-world applications.

Mean normalization is a feature-scaling model, which employed in data pre-processing for normalizing the values of data. It converts every feature by subtracting the mean and separating by the range, certifying values are centered on zero. This model aids in removing the bias caused by dissimilar feature measures, making models more effectual. It enhances the convergence speed of ML techniques, particularly those utilizing gradient descent. It is generally employed in regression methods and neural networks for enhancing accuracy. It is chiefly helpful when managing datasets comprising features with fluctuating magnitudes and units.

B. Feature selection using GOA

For the selection of feature process, the GOA technique is employed to identify and select the most relevant features. The GOA is a bio-inspired optimizer model [19]. This model is chosen due to its capability to effectually explore large solution spaces and detect the most relevant features for model training. Unlike conventional methods, GOA is a nature-inspired optimization algorithm that replicates the movement patterns of grasshoppers, giving a dynamic and adaptive approach to feature selection. It presents significant merits in handling complex, high-dimensional datasets by choosing optimal features while minimizing redundancy. The capability of the GOA method to escape local optima and converge towards global solutions makes it superior to simpler methods like genetic algorithms (GA) or particle swarm optimization (PSO), which may face difficulty with more complex feature selection tasks. Moreover, GOA can easily be integrated with other ML models, improving the performance of the model by mitigating overfitting and increasing computational efficiency. Its flexibility and robustness make it a robust candidate for feature selection in a variety of domains.

This model is derived from the mathematical representation of the grasshopper's swarming behaviour. In this method, searching for solutions takes place in dual stages: exploitation and exploration. Grasshoppers normally show these stages. During this the adult stage, they make fast movements that imitate the exploration stage. In exploration, the search is carried out through the wider region with the largest stages.

$$X_i = S_i + G_i + A_i \tag{1}$$

Here, X_i signifies the location of the ith grasshopper, S_i characterizes the social connections (repulsion and attraction) among grasshoppers, G_i signifies the gravitational force on the ith grasshopper, and A_i embodies the wind action.

$$S_i = \sum_{j=1, j\neq i}^{N} s(d_{ij})\widehat{d_{ij}} \tag{2}$$

Now, d_{ij} characterizes the distance among the ith and the jth grasshopper and is measured based on Eq. (3). s refers to function, which describes the weakness of social powers and is measured based on Eq. (4). $\widehat{d_{ij}}$ stands for unit vector from the ith grasshopper to the jth grasshopper and is computed based on Eq. (5).

$$d_{ij} = |x_j - x_i| \tag{3}$$

$$s(r) = fe^{\frac{-r}{\iota}} - e^{-r} \tag{4}$$

Whereas, f characterizes the attraction intensity, and l signifies the scale of attractive length. During this work, artificial grasshoppers' social behaviors with dissimilar f and l parameters have been noted using an experiment. They further discovered that the distance from 2.079 units to approximately four outcomes in a rise in attraction, and beyond four attraction slowly reduces.

$$\widehat{d_{ij}} = \frac{x_j - x_i}{d_{ij}} \tag{5}$$

The gravitational force on the ith grasshopper is measured as shown:

DOI: https://doi.org/10.54216/IJNS.260305

$$G_i = -g\widehat{e_g} \tag{6}$$

Here, g signifies gravitational constant, and $\widehat{e_g}$ refers to unity vector toward the Earth's center. Wind advection is measured as demonstrated:

$$A_i = u\widehat{e_w} \tag{7}$$

Whereas, u stands for continuous drift, and $\widehat{e_w}$ represents unit vector in the wind direction. Eq. (1) is extended to Eq. (8).

$$X_i = \sum_{j=1, j\neq i}^{N} s\left(\left|x_j - x_i\right|\right)\frac{x_j - x_i}{d_{ij}} - g\widehat{e_g} + u\widehat{e_w} \tag{8}$$

Nevertheless, they observed that the mathematical representation in Eq. (8) should be applied straight to the model.

$$X_i^d = C\left(\sum_{j=1, j\neq i}^{N} C\frac{ub_d - lb_d}{2} s\left(\left|x_j^d - x_i^d\right|\right)\frac{x_j - x_i}{d_{ij}}\right) + \widehat{T_d} \tag{9}$$

Whereas, the gravity module G is not counted. In addition, wind direction A is permanently presumed to be towards the source of food. ub_d specifies the upper limit of the dth dimension, lb_d specifies the lower limit of the dth dimensions, and N designates the grasshopper counts within the population, $\widehat{T_d}$ characterizes the dth size of the global target (best), namely the optimal solution discovered thus far. The distance d_{ij} epitomizes the distance among grasshopper i and j, and is measured based on Eq. (3). The function s is provided in Eq. (4). The outer c in Eq. (9) restricts the region around the global best, therefore improving exploitation, whereas inner c decreases the attraction/repulsion among grasshoppers. For maintaining the balance, parameter c is upgraded in all iterations based on Eq. (10).

$$c = cmax - l\frac{cmoax - cm\dot{m}}{L} \tag{10}$$

In Eq. (10), l indicates the present iteration, and L specifies the maximum iteration counts.

The fitness function (FF) imitates the accuracy of classifier and the quantity of chosen features. It exploits the classifier accuracy and lessens the set dimension of voted features. Therefore, the below-mentioned FF is retained for evaluating an individual solution, as exposed in Eq. (11).

$$Fitness = \alpha * ErrorRate + (1 - \alpha) * \frac{\#SF}{\#All\_F} \tag{11}$$

While $ErrorRate$ is a classification rate of error by using the chosen features. $ErrorRate$ is computed as the ratio of incorrect categorized to several classifications set among 0 and 1. $\#SF$ refers to the number of preferred features and $\#All\_F$ means a total quantity of features in an original dataset. α is employed to control the significance of classifier excellence and sub-set length.

C. BFHSS-based Classification Process

In addition, the proposed BFHSS-HOACRP technique designs BFHSS model for the classification process. In this section, describe some preceding concepts such as SS, HSS, FHSS, and BFS suitable for this study are presented [20]. This classification process is an effective approach for handling uncertainty and imprecision in data, which is usual in real-world classification tasks. BFHSS extends fuzzy logic by presenting a bipolar framework, allowing the model to handle both positive and negative information, which improves its capability to classify ambiguous and conflicting data. Compared to conventional classification methods like DT or SVMs, BFHSS gives a more flexible and robust mechanism for decision-making, specifically when dealing with incomplete or uncertain data. Its hypersoft set structure additionally improves classification by considering multiple levels of membership, giving a finer granularity of decision boundaries. This approach is particularly useful when data contains diverse degrees of uncertainty, making it more reliable in complex and dynamic environments. Furthermore, BFHSS can be easily incorporated with other optimization and ML techniques, improving overall classification performance.

Definition2.1. a BFS through a static non-empty set $\hat{Z}$ was described as the resulting framework:

$$\overline{\overline{N}} = \left\{\left(\hat{z}, F_{\overline{N}}^{+}(\hat{z}), F_{\overline{\overline{N}}}^{-}(\hat{z})\right) | \hat{z} \in \hat{Z}\right\}$$

Whereas $F_{\overline{N}}^{+}(\hat{z}): \hat{Z} \to [0,1]$ and $F_{\overline{\overline{N}}}^{-}(\hat{z}): \hat{Z} \to [0,1]$ represent the negative and positive TMs of $\hat{z} \in \hat{Z}$ using normal state $-1 \leq F_{\overline{N}}^{+}(\hat{z}) + F_{\overline{\overline{N}}}^{-}(\hat{z}) \leq 1$.

DOI: https://doi.org/10.54216/IJNS.260305

Definition2.2. Assume $\overline{\overline{N}}_1 = \left\{\left(\hat{z}, F^+_{\overline{\overline{N}}_1}(\hat{z}), F^-_{\overline{\overline{N}}_1}(\hat{z})\right) | \hat{z} \in \hat{Z}\right\}$ and $\overline{\overline{N}}_2 = \left\{\left(\hat{z}, F^+_{\overline{\overline{N}}_2}(\hat{z}), F^-_{\overline{\overline{N}}_2}(\hat{z})\right) | \hat{z} \in \hat{Z}\right\}$ is a dual BFS. Formerly, having.

$\overline{\overline{N}}_1^c = c(\overline{\overline{N}}_1) = \left\{\left(\hat{z}, 1 - F^+_{\overline{\overline{N}}_1}(\hat{z}), F^-_{\overline{\overline{N}}_1}(\hat{z})\right) | \hat{z} \in \hat{Z}\right\}$.

$\overline{\overline{N}}_1 \subseteq \overline{\overline{N}}_2$ iff $F^+_{\overline{\overline{N}}_1}(\hat{z}) \leq F^+_{\overline{\overline{N}}_2}(\hat{z})$ and $F^-_{\overline{\overline{N}}_1}(\hat{z}) \geq F^-_{\overline{\overline{N}}_2}(\hat{z});\ \forall \hat{z} \in \hat{Z}$.

$\overline{\overline{N}}_1 \cup \overline{\overline{N}}_2 = \left\{\hat{z}, \left(max\left(F^+_{\overline{\overline{N}}_1}(\hat{z}), F^+_{\overline{\overline{N}}_2}(\hat{z})\right), min\left(F^-_{\overline{\overline{N}}_1}(\hat{z}), F^-_{\overline{\overline{N}}_2}(\hat{z})\right)\right) | \hat{z} \in \hat{Z}\right\}$.

$$\overline{\overline{N}}_1 \cap \overline{\overline{N}}_2 = \left\{\hat{z}, \left(min\left(F^+_{\overline{\overline{N}}_1}(\hat{z}), F^+_{\overline{\overline{N}}_2}(\hat{z})\right), max\left(F^-_{\overline{\overline{N}}_1}(\hat{z}), F^-_{\overline{\overline{N}}_2}(\hat{z})\right)\right) | \hat{z} \in \hat{Z}\right\}$$

Definition2.3. An architecture $(\bar{\varphi}, \bar{\bar{v}})$ named SS, whereas $\bar{\varphi}: \bar{\bar{v}} \to \hat{\mathbb{P}}(\hat{Z}), \hat{\mathbb{P}}(\hat{Z})$ refers to power sets of $\hat{Z}$ and now $\hat{Z}, \bar{\bar{v}} \subset \bar{\bar{B}}$ indicate either the features or the universal set correspondingly.

Definition2.4. An BFSS is a framework that comprises SS and BFS and is outlined on a nonempty static set and the features set as demonstrated:

$$\bar{\bar{\psi}}, \bar{\bar{v}} = \bar{\bar{\psi}}(b_j) = \left\{\left(\hat{z}_k, \Im^+_{\bar{\bar{\psi}}}(\hat{z}_k), \Im^-_{\bar{\bar{\psi}}}(\hat{z}_k)\right) | \forall \hat{z}_k \in \hat{\Im}, b_j \in \bar{\bar{v}} \subseteq \bar{\bar{\mathfrak{B}}}, j = 1,2, \dots, n, \qquad k = 1,2, \dots, m\right\}$$

Example2.5. Let $\hat{Z} = \{\hat{z}_1, \hat{z}_2\}$ denote set of dual smartphones and $\bar{\bar{v}} = \{b_1 = Costly, b_2 = Battery, b_3 = lightweight\} \subseteq \bar{\bar{\mathcal{B}}}$ refers to set of features. Formerly the BFSS is inspected as shown:

$$(\bar{\bar{\psi}}, \bar{\bar{v}}) = \begin{Bmatrix} \bar{\bar{\psi}}(b_1) = \{(\hat{z}_1, 0.3, -0.1), (\hat{z}_2, 0.7, -0.5)\} \\ \bar{\bar{\psi}}(b_2) = \{(\hat{z}_1, 0.6, -0.9), (\hat{z}_2, 0.2, -0.7)\} \\ \bar{\bar{\psi}}(b_3) = \{(\hat{z}_1, 0.4, -0.2), (\hat{z}_2, 0.2, -0.8)\} \end{Bmatrix}$$

Definition2.6. A HSS $\left(\bar{\bar{\Omega}}, V_1 \times V_2 \times V_3 \times \dots \times V_n\right)$ on the fixed universe $\hat{\Im}$ is represented as demonstrated.

$$\left\{\left(\alpha, \bar{\bar{\Omega}}(\alpha)\right) : \bar{\bar{\Omega}}(\alpha) \subseteq \hat{\Im}, \forall \alpha \in V_1 \times V_2 \times V_3 \times \dots \times V_n \subseteq \mathfrak{B}_1 \times \mathfrak{B}_2 \times \mathfrak{B}_3 \times \dots \times \mathfrak{B}_n\right\},$$

Here $\Im_{\bar{\bar{\Gamma}}_{(b)}}(\hat{z})$ is the fuzzy TM and $\mathfrak{B}_i: 1,2, \dots, n$ are pair-wise separate sets of feature values.

Now, write the symbols for the benefit of ease as shown: the relation $\mathfrak{B}_1 \times \mathfrak{B}_2 \times \mathfrak{B}_3 \times \dots \times \mathfrak{B}_n = \bar{\bar{\Delta}}$ and $V_1 \times V_2 \times V_3 \times \dots \times V_n = \bar{\bar{\Lambda}}$. Next, describe the FHSS by the architecture $\left(\bar{\bar{\Gamma}}, \bar{\bar{\Lambda}}\right)$.

Here, represents the major description of BFHSS with several numeric samples and basic set operations, in addition to various basic attributes.

Definition3.1. Assume $\hat{Z}$ denote nonempty fixed set, $\hat{P} = \left(\hat{Z}\right)$ signifies to the power of $\hat{Z}$, Let $\bar{\bar{v}}_i: i = 1,2, \dots, n$ are n-well described features whose equivalent feature values are individual, the pair-wise dis-joint set $\bar{\bar{\varepsilon}}_{\bar{\bar{v}}_i}: i = 1,2, \dots, n$. Assume $\bar{\bar{\mathcal{B}}}_{\bar{\bar{v}}_i}$ represents non-empty subset of $\bar{\bar{\varepsilon}}_{\bar{\bar{v}}_i} \forall_i = 1,2, \dots, n$.

A BFHSS $\left(\bar{\bar{\Pi}}, \bar{\bar{\Lambda}}\right)$ is categorized by the mapping $\bar{\bar{\Pi}}: \bar{\bar{\Lambda}} \to \hat{P}\left(\hat{Z}\right)$ of which function value is BFS

$$\bar{\bar{\Pi}}(\alpha) = \left\{\left(\hat{z}, \mathcal{F}^+_{\bar{\bar{\Pi}}(\alpha)}(\hat{z}), \mathcal{F}^-_{\bar{\bar{\Pi}}(\alpha)}(\hat{z})\right) | \forall \hat{z} \in \hat{Z}, \alpha \in \bar{\bar{\Lambda}} \subseteq \bar{\bar{\Delta}}\right\},$$

Whereas $\bar{\bar{\Lambda}} = \bar{\bar{\mathfrak{B}}}_{\bar{\bar{v}}1} \times \bar{\bar{\mathfrak{B}}}_{\bar{\bar{v}}2} \times \dots \times \bar{\bar{\mathfrak{B}}}_{\bar{\bar{v}}n}$, $\bar{\bar{\Delta}} = \bar{\bar{\varepsilon}}_{\bar{\bar{v}}_1} \times \bar{\bar{\varepsilon}}_{\bar{\bar{v}}_2} \times \dots \times \bar{\bar{\varepsilon}}_{\bar{\bar{v}}_n}$ and $\mathcal{F}^+: \bar{\bar{\Lambda}} \to [0,1], \mathcal{F}^-: \bar{\bar{\Lambda}} \to [-1,0]$ signify negative and positive-TMs, correspondingly, like $\mathcal{F}^+_{\bar{\bar{\Pi}}(\alpha)}(\hat{z})$ specifies the acceptability of the feature α^* regarding object $\hat{z}^*$ equivalent to the BFHSS $\left(\bar{\bar{\Pi}}, \bar{\bar{\Lambda}}\right)$ as demonstrated:

$$\left(\bar{\bar{\Pi}}, \bar{\bar{\Lambda}}\right) = \left\{\langle \alpha, \left\{\left(\hat{z}, \mathcal{F}^+_{\bar{\bar{\Pi}}(\alpha)}(\hat{z}), \mathcal{F}^-_{\bar{\bar{\Pi}}(\alpha)}(\hat{z})\right) | \forall \hat{z} \in \hat{Z}\right\}\rangle : \alpha \in \bar{\bar{\Lambda}} \subseteq \bar{\bar{\Delta}}\right\}$$

Example3.2. Assuming that PhD student wants to buy a new computer between 3 computers $\hat{Z} = \{\hat{z}_1, \hat{z}_2, \hat{z}_3\}$.

Assume that the characteristics be $\bar{\bar{v}}_1 = CPU\ class, \bar{\bar{v}}_2 = Size, \bar{\bar{v}}_3 = Hard\ drive\ size$, and their feature's values are $\bar{\bar{\varepsilon}}_{\bar{\bar{v}}_1} = \{b_1 = Dual - core, b_2 = Quad - core\}, \bar{\bar{\varepsilon}}_{\bar{\bar{v}}_2} = \{b_3 = Compact\ Case, b_4 = Full\ Tower, b_5 = Mid\ Tower\}, \bar{\bar{\varepsilon}}_{\bar{\bar{v}}_3} = \{b_6 = 256GB, b_7 = 512GB, b_8 = 1TB\}$.

DOI: https://doi.org/10.54216/IJNS.260305

Assume

$$\overline{\overline{\mathfrak{B}}}_{\bar{\bar{v}}1} = \{b_1 = Dual - core\},$$

$\overline{\overline{\mathfrak{B}}}_{\bar{\bar{v}}2} = \{b_3 = Compact\ case, b_4 = Full\ Tower\}, \overline{\overline{\mathfrak{B}}}_{\bar{\bar{v}}3} = \{b_7 = 512GB, b_8 = 1TB\}$ are sub-sets of $\bar{\bar{\varepsilon}}_{\bar{\bar{v}}_i}, \forall_i =$ 1,2,3. Formerly the BFHSS $(\overline{\overline{\Pi}}, \overline{\overline{\Lambda}})$ is considered as shown

$$(\overline{\overline{\Pi}}, \overline{\overline{\Lambda}}) = \{\{\langle((b_1, b_3, b_7), \{(\hat{z}_1, 0.3, -0.1), (\hat{z}_2, 0.7, -0.5), (\hat{z}_3, 0.2, -0.4)\}),$$
$$((b_1, b_3, b_8), \{(\hat{z}_1, 0.9, -0.2), (\hat{z}_2, 0.8, -0.4), (\hat{z}_3, 0.7, -0.2)\}),$$
$$((b_1, b_4, b_7), \{(\hat{z}_1, 0, -0.5), (\hat{z}_2, 0.2, -1), (\hat{z}_3, 0.3, -0.9)\}),$$
$$((b_1, b_4, b_8), \{(\hat{z}_1, 0.7, -0.6), (\hat{z}_2, 0.4, -0.1), (\hat{z}_3, 1, -0.4)\})\rangle\}\}.$$

Definition3.3. A BFHSS $(\overline{\overline{\Pi}}, \overline{\overline{\Lambda}})$ through a permanent non-empty set $\hat{Z}$, considered to be empty BFHSS, represented by $(\overline{\overline{\Pi}}, \overline{\overline{\Lambda}})_\Phi$ assuming both $\mathcal{F}^+_{\overline{\overline{\Pi}}(\alpha)}(\hat{z}) = \mathcal{F}^-_{\overline{\overline{\Pi}}(\alpha)}(\hat{z}) = 0, \forall \hat{z} \in \hat{Z}, \forall_\alpha \in \overline{\overline{\Lambda}} \subseteq \overline{\overline{\Delta}}$ are expressed as

$$(\overline{\overline{\Pi}}, \overline{\overline{\Lambda}})_\Phi = \{\langle \alpha, \{(\hat{z}, 0,0) | \forall \hat{z} \in \hat{Z}\}\rangle : \alpha \in \overline{\overline{\Lambda}} \subseteq \overline{\overline{\Delta}}\}.$$

Definition3.4. A BFHSS $(\overline{\overline{\Pi}}, \overline{\overline{\Lambda}})$ across a static nonempty set $\hat{Z}$, considered to be absolute BFHSS, represented by $(\overline{\overline{\Pi}}, \overline{\overline{\Lambda}})_{\hat{z}}$ if $\mathcal{F}^+_{\overline{\overline{\Pi}}(\alpha)}(\hat{z}) = 1$ and $\mathcal{F}^-_{\overline{\overline{\Pi}}(\alpha)}(\hat{z}) = -1, \forall \hat{z} \in \hat{Z}, \forall_\alpha \in \overline{\overline{\Lambda}} \subseteq \overline{\overline{\Delta}}$ are outlined as

$(\overline{\overline{\Pi}}, \overline{\overline{\Lambda}})_{\hat{z}} = \{\langle \alpha, \{(\hat{z}, 1, -1) | \forall \hat{z} \in \hat{Z}\}\rangle : \alpha \in \overline{\overline{\Lambda}} \subseteq \overline{\overline{\Delta}}\}.$

At present, clarifying the concept of the complete process of the BFHSS. Before, according to this explanation, further presented a numeric model to get this explanation more obviously into view, also the most essential beliefs in set theory.

Definition3.5. Assume $\hat{Z}$ is a universe and $(\overline{\overline{\Pi}}, \overline{\overline{\Lambda}})$ is a BFHSS above $\hat{Z}$, which is well-defined in the following:

$$(\overline{\overline{\Pi}}, \overline{\overline{\Lambda}}) = \left\{\langle \alpha, \left\{\left(\hat{z}, \mathcal{F}^+_{\overline{\overline{\Pi}}(\alpha)}(\hat{z}), \mathcal{F}^-_{\overline{\overline{\Pi}}(\alpha)}(\hat{z})\right) | \forall \hat{z} \in \hat{Z}\right\}\rangle : \alpha \in \overline{\overline{\Lambda}} \subseteq \overline{\overline{\Delta}}\right\}.$$

Next, the pair of $(\overline{\overline{\Pi}}, \overline{\overline{\Lambda}})$ is signified by $(\overline{\overline{\Pi}}, \overline{\overline{\Lambda}})^c = (\overline{\overline{\Pi}}^c, \overline{\overline{\Lambda}})$ and is expressed as

$$(\overline{\overline{\Pi}}, \overline{\overline{\Lambda}})^c = \left\{\langle \alpha, \left\{\left(\hat{z}, \mathcal{F}^+_{\overline{\overline{\Pi}}(\alpha)}(\hat{z}), -1 - \mathcal{F}^-_{\overline{\overline{\Pi}}(\alpha)}(\hat{z})\right) | \forall \hat{z} \in \hat{Z}\right\}\rangle : \alpha \in \overline{\overline{\Lambda}} \subseteq \overline{\overline{\Delta}}\right\}.$$

Example3.6. Illustrate 3.2, next the inverse of the BFHSS $(\overline{\overline{\Pi}}, \overline{\overline{\Lambda}})$ compute according to the definition3.5 as:

$$(\overline{\overline{\Pi}}, \overline{\overline{\Lambda}})^c = \{\{\langle((b_1, b_3, b_7), \{(\hat{z}_1, 0.7, -0.9), (\hat{z}_2, 0.3, -0.5), (\hat{z}_3, 0.8, -0.6)\}),$$
$$((b_1, b_3, b_8), \{(\hat{z}_1, 0.1, -0.8), (\hat{z}_2, 0.2, -0.6), (\hat{z}_3, 0.3, -0.8)\}),$$
$$((b_1, b_4, b_7), \{(\hat{z}_1, 1, -0.5), (\hat{z}_2, 0.8,0), (\hat{z}_3, 0.7, -0.1)\}),$$
$$((b_1, b_4, b_8), \{(\hat{z}_1, 0.3, -0.4), (\hat{z}_2, 0.6, -0.9), (\hat{z}_3, 0, -0.6)\})\rangle\}\}.$$

Proposition3.4. When $(\overline{\overline{\Pi}}, \overline{\overline{\Lambda}})$ refers to BFHSS across the universe $\hat{Z}$.Next $\left((\overline{\overline{\Pi}}, \overline{\overline{\Lambda}})^c\right)^c = (\overline{\overline{\Pi}}, \overline{\overline{\Lambda}})$.

Evidence from Definition3.5, having

$$(\overline{\overline{\Pi}}, \overline{\overline{\Lambda}})^c = \left\{\langle \alpha, \left\{\left(\hat{z}, \mathcal{F}^+_{\overline{\overline{\Pi}}(\alpha)}(\hat{z}), -1 - \mathcal{F}^-_{\overline{\overline{\Pi}}(\alpha)}(\hat{z})\right) | \forall \hat{z} \in \hat{Z}\right\}\rangle : \alpha \in \overline{\overline{\Lambda}} \subseteq \overline{\overline{\Delta}}\right\}.$$

Therefore,

$$\left((\overline{\overline{\Pi}}, \overline{\overline{\Lambda}})^c\right)^c = \left\{\langle \alpha, \left\{\left(\hat{z}, 1 - \left(1 - \mathcal{F}^+_{\overline{\overline{\Pi}}(\alpha)}(\hat{z})\right), -1\left(-1 - \mathcal{F}^-_{\overline{\overline{\Pi}}(\alpha)}(\hat{z})\right)\right) | \forall \hat{z} \in \hat{Z}\right\}\rangle : \alpha \in \overline{\overline{\Lambda}} \subseteq \overline{\overline{\Delta}}\right\}.$$
$$= \left\{\langle \alpha, \left\{\left(\hat{z}, \mathcal{F}^+_{\overline{\overline{\Pi}}(\alpha)}(\hat{z}), \mathcal{F}^-_{\overline{\overline{\Pi}}(\alpha)}(\hat{z})\right) | \forall \hat{z} \in \hat{Z}\right\}\rangle : \alpha \in \overline{\overline{\Lambda}} \subseteq \overline{\overline{\Delta}}\right\}.$$
$$= (\overline{\overline{\Pi}}, \overline{\overline{\Lambda}}).$$

Now, continue to deliberate the concept of the sub-set of dual BFHSSs.

DOI: https://doi.org/10.54216/IJNS.260305

Definition3.8. For dual BFHSS $(\overline{\overline{\Pi}}_1, \overline{\overline{\Lambda}}_1)$ and $(\overline{\overline{\Pi}}_2, \overline{\overline{\Lambda}}_2)$ through $\hat{z}$. $(\overline{\overline{\Pi}}_1, \overline{\overline{\Lambda}}_1)$ refes to sub-set of $(\overline{\overline{\Pi}}_2, \overline{\overline{\Lambda}}_2)$, represented as $(\overline{\overline{\Pi}}_1, \overline{\overline{\Lambda}}_1) \subseteq (\overline{\overline{\Pi}}_2, \overline{\overline{\Lambda}}_2)$ when

$\overline{\overline{\Lambda}}_1 \subseteq \overline{\overline{\Lambda}}_2$.

$\mathcal{F}^{+}_{\overline{\overline{\Pi}}_1(\alpha)}(\hat{z}) \leq \mathcal{F}^{+}_{\overline{\overline{\Pi}}_2(\alpha)}(\hat{z})$ and $\mathcal{F}^{-}_{\overline{\overline{\Pi}}_1(\alpha)}(\hat{z}) \leq \mathcal{F}^{-}_{\overline{\overline{\Pi}}_2(\alpha)}(\hat{z}); \forall \alpha \in \overline{\overline{\Lambda}}_1$ and $\forall \hat{z} \in \hat{Z}$.

At present, explaining the concepts of difference, union, and intersection between binary BFHSSs. Formerly, derived from these definitions, further present's mathematical examples that arrange these definitions more evidently into view, along with the main advances in set theory.

D. Parameter tuning using SMO

Finally, the SMO-based hyperparameter selection process is performed to optimize the classification results of BFHSS model [21]. This is a powerful technique for enhancing the performance of ML methods. SMO, inspired by the behavior of spider monkeys, effectually explores the solution space by balancing exploration and exploitation, making it appropriate for global optimization problems. Unlike conventional gradient-based methods, SMO does not require the model to be differentiable, allowing it to be applied to a wide range of models, comprising complex ones like neural networks. Its capability to avoid local optima and find globally optimal solutions gives it a crucial advantage over other optimization methods, such as grid search or GAs, which may be computationally expensive or prone to overfitting. SMO also illustrates faster convergence and better scalability, particularly when tuning a large number of parameters. This makes it an ideal choice for complex, high-dimensional problems, ensuring effectual and accurate model performance. Fig. 2 represents the flowchart of SMO model.

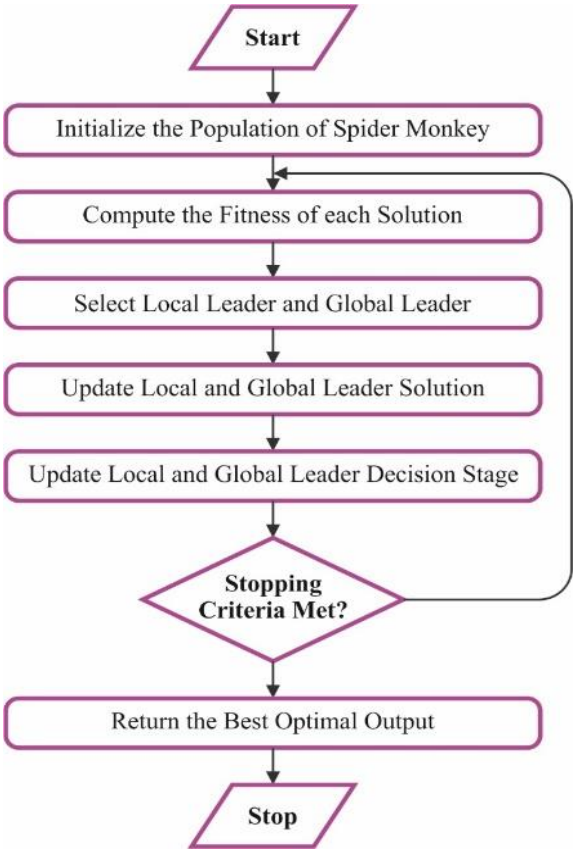

Figure 2. Flowchart of SMO model

A metaheuristic model named SMO was stimulated by how spider monkeys (SM) hunt cleverly. According to their fission-fusion social hierarchy, SMs foraging in flocks. The social system of the group, whereas a female leader makes its decisions on unification and division, becomes major factor in the individualities of the model. Now, the heads of the smaller and the larger groups are called the local and global leaders, correspondingly. The poor growth within the solution, as determined by the SMO model, symbolizes the problem of food shortage. At

DOI: https://doi.org/10.54216/IJNS.260305

least one monkey must be available in each smaller group since SMO is derived from swarm intelligence. Due to this, the time for fusion after another fission leads to minimum of one group having smaller monkeys than necessary, based on the description. A SM is the possible outcome of the SMO method.

(1) Initialization

In the beginning stage, SMO generates the original swarm of N SMs, which are equally distributed, with the ith SM in the swarm. This initialization is used for every SM_{ij}:

$$SM_{ij} = SM_{minj} + U(0,1) * \left(SM_{maxj} - SM_{minj}\right) \tag{12}$$

whereas $U(0,1)$ denotes random integer uniformly distributed in the interval of (0, 1) and SM_{minj} and SM_{maxj} represents upper and lower boundaries of the searching region in jth dimensions.

(2) Local Leader Stage

This phase characterizes the method by which each SM changes its geographical ranges; this model depends on SM determination and social inspiration. According to the members' personal experiences and the group's local leaders, it contains a positive social influence. The greedy model by which noticeable SM is selected defines regional changes. Eq. (13) displays the position modification method:

$$PO_{Newxy} = PO_{xy} + a_1 \times \left(le_{gy} - PO_{xy}\right) + a_2 \times \left(PO_{ay} - PO_{xy}\right) \tag{13}$$

Whereas PO_{xy} specifies the endurance of the xth PO in the yth dimensions, le_{gy} denote local leader of the gth group, and PO_{ay} refers to ath randomly selected SM. a_1 refers to randomly generated values among (0, 1), and a_2 varies between (− 1, 1).

(i) Global Leader Phase (GLP)

SM upgrades its perspective after being inspired by the global head and taking inspiration from the neighbor. In this step, the position is upgraded depending on fitness, and using SM with greater fitness has a better probability of upgrading itself than SM with low fitness.

$$Fitness = fit_i \begin{cases} \dfrac{1}{1+f_i}, & if \quad f_i \geq 0 \\ 1 + abs(f_i), & if \quad f_i < 0 \end{cases} \tag{14}$$

According to the roulette wheel selection, the selection possibility is measured. The likelihood that SM should be selected for the GLP is calculated utilizing the dual equations below if f_i denote fitness of fit_i SM:

$$Prob_i = \frac{fitness}{\sum_{i=1}^{N} fitness_i} \; or \tag{15}$$

$$Prob_i = 0.9 * \frac{fit_i}{\max(fit_i)} + 0.1 \tag{16}$$

The working knowledge of the global head, the mastery of neighboring SM, and $SM's$ own endurance is all applied to upgrade the location. Eq. (26) defines the position modification process:

$$PO_{newxy} = PO_{xy} + a_1 \times \left(CB_y - PO_{xy}\right) + a_2 \times \left(PO_{ay} - PO_{xy}\right) \tag{17}$$

CB_y refers to global head of the bevy's position. Afterward the change, individuals are vulnerable to selection of greed, by which a monkey whose fitness is higher is preferable to its novel position through this previous one. When the $fit_{New}(PO) > fit_{old}(PO)$ denote, implement the new upgraded location in Eq. (26), or else perform the older location in Eq. (18).

(ii) Global Leader Learning Phase (GLLP)

During this stage, the group's GL should be influenced, and the proper monkey will be selected as the GL. Only the GL transfers, the global limit scores are reorganized to 0 and or else increased by 1.

(iii) Local Leader Learning Phase (LLLP)

The heads of all sub-groups are chosen in this phase. A tally named the local limit scores is improved by 1 when a local leader doesn't inverse sequence.

(iv) Local Leader Decision Phase (LLDP)

DOI: https://doi.org/10.54216/IJNS.260305

Each of the monkeys in the group changed their settings. When the local head is not encouraged to a particular edge names as the Limit of the Local Leader LeI, both by the randomization method or by utilizing global leader data over the rate of perturbation presented in Eq. (19):

$$PO_{Newxy} = PO_{xy} + a_1 \times \left(CB_y - PO_{xy}\right) + a_1 \times \left(PO_{xy} - Lel\right) \tag{19}$$

(v) Global Leader Decision Phase (GLDP)

When the global head is not stimulated to a particular edge named the Limit of the Global Leader (LeI), the bevy is divided into small sub-groups or combined into only greater groups.

The SMO technique originates an FF to completely augment performance of classifier. It describes an optimistic number to epitomize the superior performance of candidate solution. The classification error rate reduction is dignified as FF and formulated in Eq. (20).

$$fitness(x_i) = ClassifierErrorRate(x_i)$$
$$= \frac{no.of\ misclassified\ samples}{Total\ no.of\ samples} \times 100 \tag{20}$$

4. Experimental Validation

The experimental analysis of BFHSS-HOACRP model is examined under bank customer churn prediction dataset [22]. This dataset includes 4000 instances based on 2 classes namely Exited-yes and Exited-No as depicted in Table 1. There are 13 features in total but only 10 features are chosen.

Table 1: Details of dataset

Class	No. of Instances
Exited-yes	2000
Exited-No	2000
Total Instances	4000

4.1. Evaluation Metrics for Classification Models:

$Accu_y$, $Prec_n$, $Reca_l$, $F1_{Score}$, and AUC_{score}

The performance of the classification models is evaluated using diverse metrics. Eq. (21) represents $accu_y$, which measures the overall proportion of correct predictions. Eq. (22) computes $prec_n$, the ratio of correct positive predictions. Eq. (23) defines $reca_l$, which evaluates the model's ability to detect true positives. Eq. (24) exhibits the $F1_{score}$,, a metric that incorporates $prec_n$ and $reca_l$ into a single value to balance their trade-offs, particularly in cases with imbalanced classes. Finally, Eq. (25) represents AUC_{score}, which computes the balance between classification accuracy for both classes, giving a more balanced performance evaluation in imbalanced datasets. These metrics provide a comprehensive evaluation of model performance as demonstrated by the following equations:

$$Accu_y = \frac{TP + TN}{TP + TN + FP + FN} \tag{21}$$

$$Prec_n = \frac{TP}{TP + FP} \tag{22}$$

$$Reca_l = \frac{TP}{TP + FN} \tag{23}$$

$$F1_{Score} = \frac{2 \cdot Prec_n \cdot Reca_l}{Prec_n + Reca_l} \tag{24}$$

$$AUC = \int_0^1 TPR(FPR)d(FPR) \tag{25}$$

DOI: https://doi.org/10.54216/IJNS.260305

Here, TP represents True Positives, TN indicates True Negatives, FP denotes False Positives, and FN stands for False Negatives. These metrics provide a comprehensive evaluation of the model's performance, including its ability to correctly identify both positive and negative instances, as well as its handling of class imbalances. The inclusion of $prec_n$, $reca_l$, $F1_{score}$, and AUC_{score} presents insights into the efficiency of the model in detecting relevant patterns while minimizing errors across diverse classification scenarios.

4.2. Results and Discussion: Comprehensive Evaluation and Comparative Analysis of Model Performance

Fig. 3 displays the classifier performances of the BFHSS-HOACRP model based on 70%TRAPHA and 30%TESPHA. Figs. 3a-3b represents the confusion matrices using specific classification and identification of all classes. Fig. 3c exemplifies the PR outcome, notifying superior performance across all class labels. At last, Fig. 3d illustrates the ROC inspection, representing skillful performances through high ROC values for different class labels.

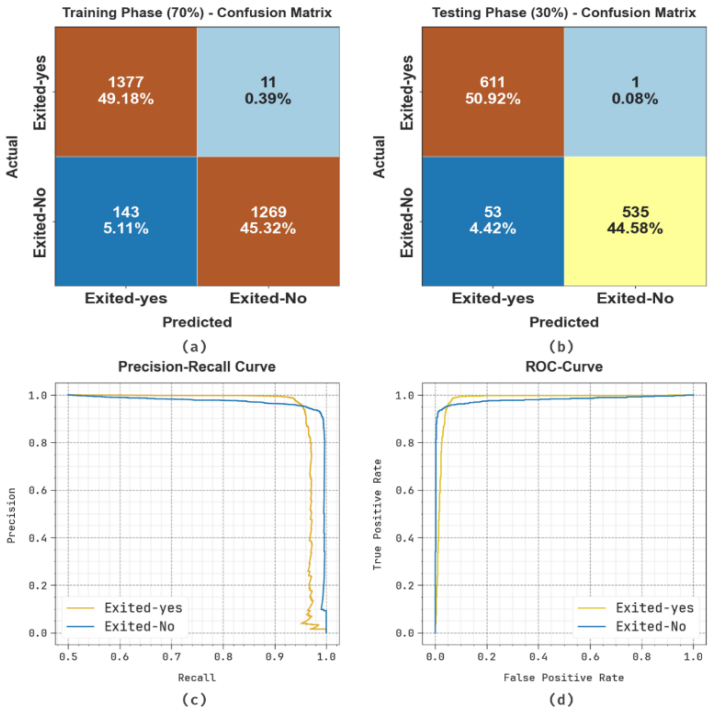

Figure 3. Confusion matrices of BFHSS-HOACRP model (a-b) 70:30 of TRAPHA/TESPHA and (c-d) curves of PR and ROC

In Table 2 and Fig. 4, the customer churn detection of the BFHSS-HOACRP model based on 70%TRAPHA and 30%TESPHA is showcased. The performances depicted that the BFHSS-HOACRP model has gained efficacious performance. On 70%TRAPHA, the BFHSS-HOACRP method reaches average $accu_y$, $prec_n$, $reca_l$, $F1_{score}$, and AUC_{score} of 94.54%, 94.87%, 94.54%, 94.49%, and 94.54%. Also, on 30%TESPHA, the BFHSS-HOACRP method attains an average $accu_y$, $prec_n$, $reca_l$, $F1_{score}$, and AUC_{score} of 95.41%, 95.92%, 95.41%, 95.48%, and 95.41%.

Table 2: Customer churn detection of BFHSS-HOACRP model under 70%TRAPHA and 30%TESPHA

Class Labels	$Accu_y$	$Prec_n$	$Reca_l$	$F1_{score}$	AUC_{score}
TRAPHA (70%)					
Exited-yes	99.21	90.59	99.21	94.70	94.54
Exited-No	89.87	99.14	89.87	94.28	94.54

DOI: https://doi.org/10.54216/IJNS.260305

Average	94.54	94.87	94.54	94.49	94.54
TESPHA (30%)					
Exited-yes	99.84	92.02	99.84	95.77	95.41
Exited-No	90.99	99.81	90.99	95.20	95.41
Average	95.41	95.92	95.41	95.48	95.41

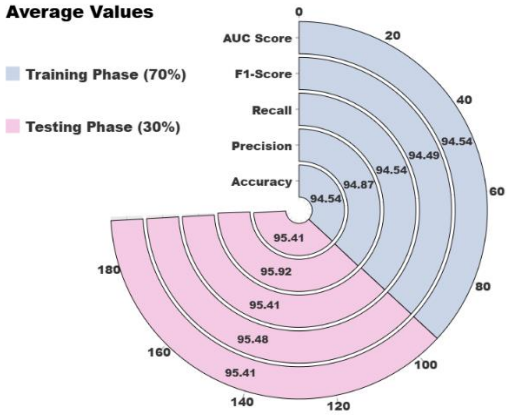

Figure 4. Average of BFHSS-HOACRP technique based on 70%TRAPHA and 30%TESPHA

In Fig. 5, the training $accu_y$ (TRAAC) and validation $accu_y$ (VLAAC) performances of BFHSS-HOACRP approach are exemplified. The $accu_y$ values are calculated across an interlude of 0-25 epochs. The figure underscored that both values express a cumulative propensity, indicating the competency of the BFHSS-HOACRP approach using maximum reach through multiple repetitions. In addition, both values ruin nearer over the epochs, indicating diminished overfitting and presenting maximum success of the BFHSS-HOACRP technique, ensuring dependable prediction on hidden samples.

Figure 5. $Accu_y$ Curve of BFHSS-HOACRP technique

In Fig. 6, the TRA loss (TRALO) and VLA loss (VLALO) graph of the BFHSS-HOACRP approach is exemplified. The loss values are computed through a period of 0-25 epochs. It is signified that both values represent a declining propensity, indicating the proficiency of the BFHSS-HOACRP approach in harmonizing an equilibrium between generalization and data fitting. The progressive dilution in values of loss also guarantees the better result of the BFHSS-HOACRP technique and tunes the prediction solutions after a while.

DOI: https://doi.org/10.54216/IJNS.260305

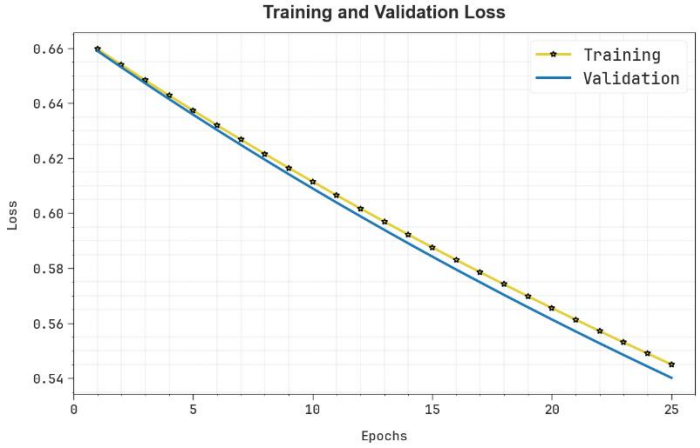

Figure 6. Loss curve of BFHSS-HOACRP technique

4.3. Comprehensive Comparison Study of the BFHSS-HOACRP Method with Existing Techniques Across Multiple Performance Metrics and Scenarios

Table 3 and Fig. 7 illustrate the comparative study of BFHSS-HOACRP model with existing methodologies is depicted [23-25]. The values of table specified that the proposed BFHSS-HOACRP model has attained effective performance. The results underscored that the SVM, NN, RF, KNN, DTC, RTC, Stochastic GB methods have stated poorer performance. In addition, the proposed BFHSS-HOACRP method indicated maximum performance with increased $accu_y$ of 95.41%, $prec_n$ of 95.92%, $reca_l$ of 95.41%, and $F1_{score}$ of 95.48%.

Table 3: Comparative study of BFHSS-HOACRP technique with existing models

Technique	$Accu_y$	$Prec_n$	$Reca_l$	$F1_{score}$
SVM Model	83.68	87.96	89.89	87.86
Neural Network	89.36	89.68	88.25	89.56
Random Forest	81.99	89.89	83.78	87.86
KNN Algorithm	90.98	83.20	90.68	82.06
DTC Model	92.23	81.34	82.95	84.96
RTC Method	80.50	87.53	91.73	85.78
Stochastic GB	90.91	80.54	84.20	87.13
BFHSS-HOACRP	95.41	95.92	95.41	95.48

DOI: https://doi.org/10.54216/IJNS.260305

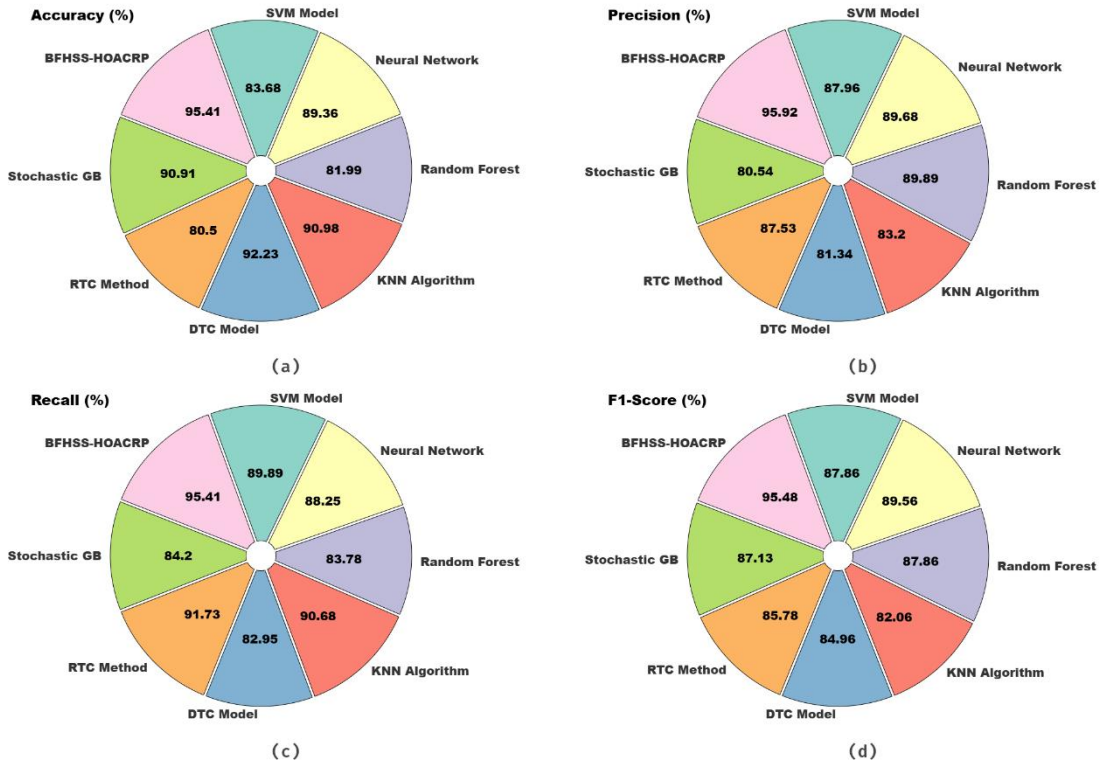

Figure 7. Comparative analysis of BFHSS-HOACRP technique with existing models

4.4. Detailed CT Analysis of the BFHSS-HOACRP Method with Existing Techniques in Various Performance Aspects and Use Cases

The computational time (CT) outcome of BFHSS-HOACRP approach with existing techniques are exemplified in Table 4 and Fig. 8. Based upon time, the BFHSS-HOACRP approach delivers minimal time of 7.81min while the SVM, NN, RF, KNN, DTC, RTC, Stochastic GB methodologies gain better time of 23.03min, 18.29min, 12.45min, 13.28min, 21.90min, 22.31min, and 14.14min, respectively.

Table 4: Time outcome of BFHSS-HOACRP method with existing models

Method	**CT (min)**
SVM Model	23.03
Neural Network	18.29
Random Forest	12.45
KNN Algorithm	13.28
DTC Model	21.90
RTC Method	22.31
Stochastic GB	14.14
BFHSS-HOACRP	7.81

DOI: https://doi.org/10.54216/IJNS.260305

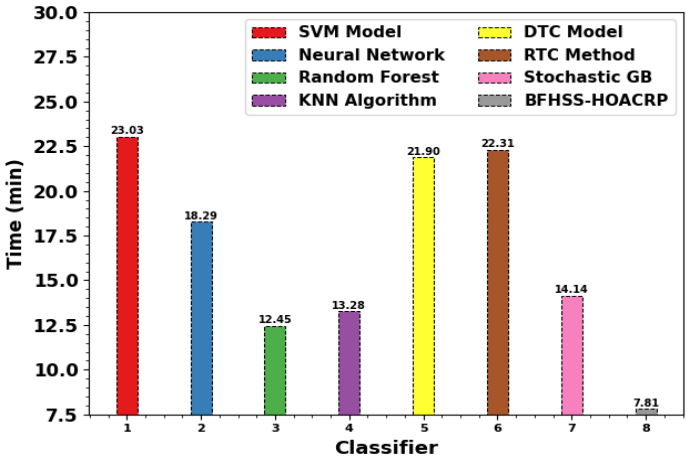

Figure 8. Time outcome of BFHSS-HOACRP method with existing models

4.5. Comprehensive Ablation Study of the BFHSS-HOACRP Method to Assess its Impact on Model Performance

Table 5 and Fig. 9 demonstrates the ablation study of the BFHSS-HOACRP methodology. The ablation study results demonstrate the superior performance of the BFHSS-HOACRP model compared to other methods. The BFHSS-HOACRP model attained an $accu_y$ of 95.41%, a $prec_n$ of 95.92%, $reca_l$ of 95.41%, and an $F1_{score}$ of 95.48%, outperforming GOA, SMO, and BFHSS across all metrics. The improvement in all evaluation metrics illustrates that the integration of the proposed technique significantly enhances the capability of the model to accurately classify and predict outcomes. While the BFHSS and SMO methods also perform well, the BFHSS-HOACRP model exhibits a notable advantage in balancing both $prec_n$ and $reca_l$, highlighting its robust capability in handling complex prediction tasks. These results underscore the efficiency of the proposed BFHSS-HOACRP model in giving a more accurate and reliable solution for the task.

Table 5: Result analysis of the ablation study of BFHSS-HOACRP methodology

Methodology	$Accu_y$	$Prec_n$	$Reca_l$	$F1_{score}$
GOA	93.36	93.87	93.57	93.53
SMO Algorithm	93.90	94.51	94.18	94.32
BFHSS	94.66	95.20	94.71	94.92
BFHSS-HOACRP	95.41	95.92	95.41	95.48

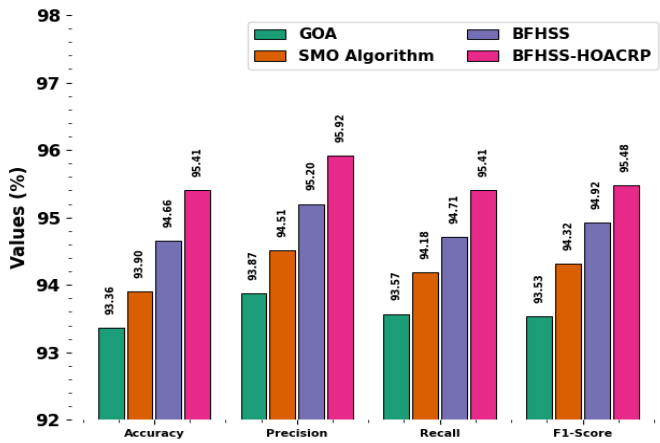

Figure 9. Result analysis of the ablation study of BFHSS-HOACRP methodology

DOI: https://doi.org/10.54216/IJNS.260305

5. Conclusion

This study proposes a BFHSS-HOACRP technique in financial sectors. The BFHSS-HOACRP technique applies optimized techniques to predict the customer retention behaviour in the industry of bank. Initially, the mean normalization technique is utilized in the data pre-processing stage to prepare raw data into a suitable format for analysis and modelling. For the selection of feature process, the GOA is employed to identify and select the most relevant features from an input data. In addition, the proposed BFHSS-HOACRP technique implements the BFHSS method for the classification process. Additionally, the SMO-based hyperparameter selection process is performed to optimize the classification results of BFHSS model. The efficacy of the BFHSS-HOACRP approach is examined under the bank customer churn prediction dataset. The comparison analysis of the BFHSS-HOACRP approach portrayed a superior accuracy value of 95.41% over existing techniques. The limitations of the BFHSS-HOACRP approach comprise its reliance on a specific dataset, which may not generalize well to other industries or domains. The performance of the model could be affected by the quality and completeness of the data, particularly in the presence of noise or missing values. Furthermore, the study does not account for real-time data fluctuations, which could affect the accuracy of predictions in dynamic environments. The complexity of the model may also result in an enhanced computational need, limiting its scalability for large datasets. Moreover, the interpretability of the model could be an issue, as more complex techniques may lack transparency in decision-making. Future work may explore incorporating real-time data sources, enhancing model interpretability through XAI, and improving the scalability of the model to handle larger, more diverse datasets. Additionally, the performance of the model could be further validated utilizing cross-domain applications and diverse business scenarios.

Funding: “This research received no external funding”

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

References

[1] R. Farokhzad Rostami, "Fixed point theorems in ordered non-Archimedean fuzzy metric spaces," *Int. J. Nonlinear Anal. Appl.*, vol. 14, no. 9, pp. 345–356, Sep. 2023.

[2] A. M. Abd-Elal, "Fixed point results for fuzzy mappings in F-metric spaces," *Int. J. Nonlinear Anal. Appl.*, vol. 14, no. 1, pp. 575–584, Jan. 2023.

[3] T. Bhandari, K. B. Manandhar, and K. Jha, "Fixed point theorem in fuzzy b-metric space using compatible mapping of type (A)," *Pure Appl. Math. J.*, vol. 13, no. 6, pp. 100–108, 2024.

[4] M. Mecheraoui, "On some fixed point results in E−fuzzy metric spaces," *J. Math.*, vol. 2021, 2021.

[5] M. K. Tiwari, "Fixed point theorem in fuzzy metric space," *Int. J. Eng. Res. Technol.*, vol. 9, no. 1, Jan. 2020.

[6] F. Shirazi and M. Mohammadi, "A big data analytics model for customer churn prediction in the retiree segment," *Int. J. Inf. Manage.*, vol. 48, pp. 238–253, Oct. 2019.

[7] R. A. De Lima Lemos, T. C. Silva, and B. M. Tabak, "Propension to customer churn in a financial institution: a machine learning approach," *Neural Comput. Appl.*, vol. 34, no. 14, pp. 11751–11768, Jul. 2022.

[8] M. K. Abdela et al., "The effect of customer relationship management on customer loyalty on banking sector," *Int. J. Manage.*, vol. 14, no. 5, pp. 20–37, 2023.

[9] M. R. Machado, S. Karray, and I. T. De Sousa, "LightGBM: An effective decision tree gradient boosting method to predict customer loyalty in the finance industry," in *Proc. 14th Int. Conf. Comput. Sci. Educ. (ICCSE)*, Aug. 2019, pp. 1111–1116.

[10] R. S. Subramanian et al., "Ensemble-based deep learning techniques for customer churn prediction model," *Kybernetes*, 2024.

[11] F. Ehsani and M. Hosseini, "Customer churn analysis using feature optimization methods and tree-based classifiers," *J. Serv. Mark.*, vol. 39, no. 1, pp. 20–35, 2025.

[12] C. Wang et al., "Risk assessment of customer churn in telco using fclcnn-lstm model," *Expert Syst. Appl.*, vol. 248, p. 123352, 2024.

[13] Y. Li and K. Yan, "Prediction of bank credit customers churn based on machine learning and interpretability analysis," *Data Sci. Finance Econ.*, vol. 5, no. 1, pp. 19–35, 2025.

DOI: https://doi.org/10.54216/IJNS.260305

[14] S. A. Panimalar, A. Krishnakumar, and S. S. Kumar, "Intensified customer churn prediction: connectivity with weighted multi-layer perceptron and enhanced multipath back propagation," *Expert Syst. Appl.*, vol. 265, p. 125993, 2025.

[15] S. K. Wagh et al., "Customer churn prediction in telecom sector using machine learning techniques," *Results Control Optim.*, vol. 14, p. 100342, 2024.

[16] P. Pathak et al., "Customer churn prediction and personalised recommendations in banking," in *Int. Conf. Artif. Intell. Smart Energy*, Cham: Springer Nature Switzerland, Mar. 2024, pp. 409–421.

[17] V. Haridasan, K. Muthukumaran, and K. Hariharanath, "Arithmetic optimization with deep learning enabled churn prediction model for telecommunication industries," *Intell. Autom. Soft Comput*, vol. 35, no. 3, 2023.

[18] J. Zhang et al., "Deformation stage division and early warning of landslides based on the statistical characteristics of landslide kinematic features," *Landslides*, vol. 21, no. 4, pp. 717–735, 2024.

[19] A. Babalik and A. Babadag, "A binary grasshopper optimization algorithm for solving uncapacitated facility location problem," *Eng. Sci. Technol. Int. J.*, vol. 65, p. 102031, 2025.

[20] A. Al-Quran et al., "Bipolar fuzzy hypersoft set and its application in decision making," *Infinite Study*, 2023.

[21] G. L. Anoop, C. Nandini, and E. Naresh, "3TFL-XLnet-CP: A novel transformer-based crop yield prediction framework with weighted loss based 3-tier feature learning model," *SN Comput. Sci.*, vol. 6, no. 3, p. 275, 2025.

[22] S. Y. Al-Sultan and I. A. Al-Baltah, "An improved random forest algorithm (ERFA) utilizing an unbalanced and balanced dataset to predict customer churn in the banking sector," *IEEE Access*, 2024.

[23] M. Maduna et al., "Smart customer churn management system using machine learning," *Procedia Comput. Sci.*, vol. 237, pp. 552–558, 2024.

[24] B. Prabadevi, R. Shalini, and B. R. Kavitha, "Customer churning analysis using machine learning algorithms," *Int. J. Intell. Netw.*, vol. 4, pp. 145–154, 2023.

[25] H. Y. Daher and J. R. Kider, "Some properties of convex fuzzy normed space," *Int. J. Math. Comput. Sci.*, vol. 19, no. 2, pp. 345–355, 2024.

DOI: https://doi.org/10.54216/IJNS.260305