RESEARCH ARTICLE

INTERNATIONAL RESEARCH JOURNAL OF
MULTIDISCIPLINARY TECHNOVATION

Asian Research Association

DOI: 10.54392/irjmt25413

An Enhanced Extreme Learning Machine Based on a Swarm Intelligence Approach for Heart Disease Detection

P. Vigneshvaran [a, *], A. Vijaya Kathiravan [b]

[a] Department of Data Science, Shanmuga Industries Arts & Science College, Tiruvannamalai, Tamil Nadu, India
[b] Department of Computer Applications, Sri Meenakshi Government College for Women (Autonomous), Madurai, Tamil Nadu, India

* Corresponding Author Email: vigneshvaran.p@gmail.com

DOI: https://doi.org/10.54392/irjmt25413

Received: 09-01-2025; Revised: 27-06-2025; Accepted: 16-07-2025; Published: 23-07-2025

Abstract: Heart disease (HD) is frequently considered the most problematic human disease. Designing a more precise heart disease prediction technique is a challenging task. Recently, several health disease detection problems have been addressed using the extreme learning machine (ELM) approach. The ELM has become increasingly generally employed in various learning domains for prediction and control applications due to its rapid learning rate, straightforward structure, and excessive generalizability. Finding the perfect input weights and hidden bias parameters is the foremost challenge to enhancing the ELM's performance. Randomly selecting these parameters wants more hidden neurons than the traditional gradient learning technique, delaying the prediction response and decreasing the likelihood of finding the ideal output weight. In this study, bacterial colony optimization (BCO) is used to adjust the connection weights and bias of ELM (BCO+ELM) to address the drawbacks above. Additionally, the missing value of the heart dataset is filled with multivariate imputation by chained equation (MICE) and relevant features are selected by using recursive feature elimination (RFE) for obtaining more efficient solution accuracy and enhancing the performance of proposed BCO+ELM. According to the performance measures, BCO+ELM shows excellent prediction performances. The experimental effects indicate that the BCO+ELM creates better outcomes and low computation time with fast convergence time when compared to other approaches. The best-performing BCO+ELM shows consistent gains, although at a somewhat smaller rate (0.4% after imputation, 2.76% after feature selection). This implies that BCO+ELM still gains from the improvements provided by both imputation and feature selection, further optimizing its performance even if it is already highly optimized.

Keywords: Heart Diseases, Extreme Learning Machine, Bacterial Colony Optimization, Convergence Rate, Medical Data Classification

1. Introduction

An essential part of the human body, the heart uses a system of arteries and veins to pump blood enriched with oxygen to other organs. Any condition that harms our hearts is considered heart disease, which is the leading cause of death. According to estimates from the World Health Organization (WHO), cardiovascular diseases were directly responsible for 17.9 million deaths worldwide in 2019, or 32% of all fatalities. Heart attacks were the cause of 85% of these fatalities. The bulk of deaths from CVD occur in nations with modest and poor incomes. In 2019, non-communicable diseases claimed the lives of 17 million people under the age of 70, with cardiovascular diseases accounting for 38% of those deaths. Heart disorders can take many different forms, such as coronary artery, congenital heart disease, and arrhythmia. Breathlessness, excessive perspiration, and chest pain are just a few of the symptoms that people with heart disease may experience. Among other things, smoking, high blood pressure, diabetes, and obesity are the main causes of heart disease. It is expensive and uncomfortable to diagnose the illness by invasive methods. Therefore, a non-invasive, low-cost technique to detect cardiac disease is needed [1].

The main goal of the detection or prediction is to produce a scheme that, in the absence of a physician, can be utilized by any literate person to make an early diagnosis of heart disease. Even with kinds of prediction or detection systems, the doctor may receive assistance in making a choice. The difficulty in finding reputable facilities at reasonable costs is a major issue for healthcare organizations such as clinics hospitals and health centres. For the current models, accuracy, usability, and reliability are the primary issues. People typically seek a heart professional at a fairly advanced stage [2]. Hence, the early finding of heart disease, and the death rate may be decreased with improved monitoring of heart patients. The goal of the study is to identify the effective machine learning (ML) method for making more precise, sensitive, and accurate diagnoses of heart disease. The ML algorithms can support to detection and identify various diseases and create a model using sample data to make decisions [3]. The big challenge of ML is to make a more efficient prediction model with high accuracy. Because some of the conventional ML algorithms have some shortcomings. There are several, inaccurate heart disease decision assistance systems to be found in the literature [4, 5].

ELM for single-hidden-layer feed-forward neural networks (SLFN) was proposed in 2004 to address the shortcomings of gradient-based approaches. ELM uses simple generalized inverse operations on the hidden layer output matrix to choose at random the input weights and hidden biases. The Moore–Penrose (MP) is used to analytically compute the SLFN output weights. ELM learns significantly more quickly, has superior generalization, and also avoids many of the challenges such as terminating criteria, learning rate, learning epochs, and local minimum [6, 7]. However, it has also been observed that ELM tends to want more hidden neurons than conventional gradient-based algorithms, as well as having an issue with suffering due to randomly gathering input weights and hidden biases. The model's prediction accuracy is highly correlated with the number of hidden layer points. If there are too many hidden layer nodes, it may result in over-fitting; alternatively, it may result in under-fitting [8]. Hence, tuning the optimal parameters of ELM is a significant task to obtain more efficient prediction accuracy and low computation time with faster convergence.

Deep learning (DL) models and neural network hyperparameter tuning heavily rely on swarm intelligence (SI) methods. These methods perform global optimization, which is vital for optimizing complicated learning methods, by simulating the collective behavior of biological schemes. SI methods significantly increase method accuracy, generalization, and convergence speed by cleverly navigating the hyperparameter space. Among various SI methods, BCO frequently performs better than traditional methods because of their strong global search and adaptive behavior, which makes them perfect for improving [9]. BCO is a robust and flexible metaheuristic with distinguished benefits in terms of accuracy, convergence, flexibility, and elasticity [10]. Because of its biological inspiration, it can simulate complicated dynamics of problem-solving, which is very useful for DL and ANN-based systems' hyperparameter tweaking and other optimization tasks. BCO can successfully detect global solutions while avoiding local optima by striking a balance between exploration and exploitation [11].

In this study, a brand-new method called BCO-ELM—a combination of ELM and BCO—is proposed. The input weights and hidden biases are optimized using the BCO, and the output weights are calculated analytically using the MP (Moore-Penrose) generalized inverse. The ELM model's best parameters are obtained using the BCO, which also increases the accuracy of heart disease prediction. BCO is a metaheuristic that optimizes complex operations by drawing inspiration from the foraging behavior of bacteria. Examining experimental data confirms if BCO improves ELM's learning capabilities by minimizing training time, avoiding overfitting or underfitting, and optimizing weights and biases. Often, parameter adjustment is required to optimize performance. Analysis of experimental findings makes it possible to modify the ELM and BCO algorithm's parameters (such as the number of neurons or bacteria) to get the best results, which is crucial for sensitive applications. To improve the performance of the BCO+ELM and produce a more accurate solution, the missing value of the heart dataset is also filled with multivariate imputation by chained equation (MICE), and important features are chosen by utilizing recursive feature elimination (RFE). The main objective of the suggested strategy is to create a more accurate detection model for diagnosing heart disease early on. The following are the research's contributions:

- Missing feature values are imputed by using multivariate imputation by chained equations (MICE)
- The REF approach is used to choose relevant features, while the Min-Max is used to normalize the data.
- To improve the performance of ELM, the BCO is employed to determine the ELM's connection parameters.
- The suggested BCO-ELM is used to classify whether a person has an illness or is normal.

The remaining portions of this research are planned as follows; Section 2 defines the related works. Sections 3 and 4 present the standard ELM and BCO algorithms respectively. The details of the BCO-ELM are introduced in Section 5. The result assessments are presented in Section 6. Finally, the conclusion is delivered in Section 7.

2. Related Works

The related works section examines earlier research and methodology in the fields of ML, optimization techniques, and heart disease prediction. It aids readers in comprehending the framework upon which the present study is based. For instance, the authors might provide an overview of earlier research on ELM and how they are used in healthcare, mostly in the diagnosis of heart disease. This proves that the study is necessary and points out weaknesses in the current methods. The present section discusses some recent research work for identifying heart diseases. H. Li et al. (2015) [12] developed a disease detection method for heart illnesses by combining approaches from AI and the Internet of Things (IoT). Data collection, pre-processing, classification, and parameter configuration are the stages. Wearable devices help with data collection, which is then investigated by AI systems to identify illnesses. The proposed techniques use the Cuckoo search optimization (CSO)-based cascaded short-term memory (CuSO-MuLSTM) for disease diagnosis. To expand medical data classification, CuSO is used to fine-tune the "weight" and "bias" parameters of the MuLSTM model. Moreover, the isolation forest technique (iForest) has been used to eliminate outliers from this investigation determination. However, the developed method produced high accuracy. But, it takes a high computation time and is possible to fall into overfitting. N. K. Sharma (2025) [13] developed the ensemble machine learning method, which integrates many methods into a single predictive architecture that makes use of the benefits of several basic methods—often referred to as weak learners—to make up for the shortcomings of each approach. To generate a powerful classifier with high accuracy, the combination of multiple low-performance classifiers using iterative ensemble techniques. The dataset is collected from the IEEE data port which includes around 1190 cases with 11 heart disease characteristics. The conventional ML method has many shortcomings such as low accuracy and may fail to capture composite non-linear dependencies in medical data without being carefully engineered. Y. Khourdifi *et al.* (2019) [14] developed a new hybrid method for processing different forms of data for the categorization of heart disease. To enhance the accuracy of heart disease classification, the fast correlation-based feature selection (FCBF) method to remove duplicate features. Then, selected features are used for classification tasks to detect heart diseases. S. Mohan et al. (2019) propose a unique approach that attempts to identify significant traits, it will be possible to forecast cardiovascular disease more accurately by using ML approaches. Diverse feature groupings and several well-known classification approaches are used to introduce the prediction model [15]. W. He *et al*. (2020) [16] propose a new evolutionary categorization model based on a kernel ELM (KELM) optimized using an improved salp swarm algorithm (ISSA) that serves as the foundation of the prediction model. The space transformation mechanism is added to the optimization core to enhance SSA for finding the best subset of parameters and features to create an ideal KELM model. The best parameters and a select group of features are used to build the SSA-KELM, a KELM for the finding of heart disease.

M. S. Nawaz *et al.* (2021) [17] developed a new intelligent method based on the gradient descent method for the analysis of heart disease. The suggested intelligent approach classifies heart disease into positive and negative symptoms. The developed method is generally shallow learners who can't learn complex thoughts. A. F. Subahi *et al.* (2022) [18] proposed a modified self-adaptive Bayesian (MSABA) with an IoT platform to deliver more accurate assessments of HD. The patient's smartwatch and pulse sensor gadget capture and communicate data from the patient's vital signs, such as their electrocardiogram (ECG) and blood pressure, to a computer. To establish if the obtained sensor data is normal or pathological, the MSABA is utilized. The method that was developed produced great accuracy. However, because of the dynamic situation, it generated highly nosy data. P. Rani *et al.* (2021) [19] suggested a method of detection that can aid in the early identification of cardiac disease. The authors used the MICE to address the missing values. A hybridized feature selection method that incorporates the genetic algorithm (GA), RFE, SMOTE (“Synthetic Minority Oversampling Technique”), and common scalar approaches was used to choose appropriate features from the given dataset. Common scalar approaches were also used for pre-processing the data. The patient has been categorized using a variety of classification algorithms. Nevertheless, the traditional ML approach yielded poor accuracy and required a lot of computing time. P. Yang *et al*. (2021) [20] suggest a multi-kernel and ensemble learning-based solution for automatically classifying ECG heartbeats. To create a mixed-kernel-based ELM, the author first employs a linear combination of the RBF and the polynomial kernel (MKELM). Then, an MKELM-RF (MKELM-based RF Binary Classifier) is built. Finally, a collaborative multiclass classifier based on one-vs.-one (OVO) reduction and MKELM-RF are proposed. R.C. Ripan *et al.* (2021) [21] observe anomaly detection in the healthcare trade using unsupervised K-means. To create the clusters needed to discover anomalies, our suggested model first selects an ideal value of K using the Silhouette approach. Then, we remove the data's discovered anomalies and use the five most used ML classification approaches. Y. Kaya *et al.* (2022) [22] considered two phases in the suggested hybrid model. The medical dataset was first preprocessed using FA, and the factors that were derived from this process were then utilized as input features for the ELM model. The suggested model was tested using actual datasets from the UCI library for dermatology, lymphography, and Wisconsin breast cancer. Found that in three distinct medical datasets utilized for the classification challenge, preparing the data improved the average classification success. The suggested hybrid paradigm is thought to be beneficial for medical diagnosis systems' decision-making phase.

R. Siouda *et al.* (2024) [23] suggested combining radial basis function (RBF) with ELM in a hybrid classifier. The excellent performance and complementary nature of these two neural network types serve as the driving forces behind this combination. The fundamental concept is based on using an ELM network with a variety of hidden neurons to supplement a compact RBF network. A GA is used to optimize the types, numbers, and shapes of the ELM hidden neurons. The following is a summary of the proposed classifier's goals. First, it gains from the complementing qualities of ELM and RBF, such as ELM's global reaction and RBFs' local response. Secondly, it utilizes the benefits of ELM, such as quick training and the ability to apply several activation functions. Third, by combining the random initialization of ELM with the systematic initialization of RBF, it mitigates the ill-conditioning issue of ELM. Fourth, because GA is only concerned with the additional neurons, whose function is to supplement the RBF network, the optimization procedure is made simpler. However, GA is fall into local optima and low convergence rate. S. Sucharita *et al.* (2024) [24] suggest a two-step hybrid methodology that will streamline the process of finding the optimal feature selection method without entirely depending on individual techniques. They employ a collection of filter-based heterogeneous feature selection techniques in the first step. The second stage of wrapper-based selection is subsequently applied to these chosen features. Suggested using an ELM as the fitness function in this step of the bio-inspired technique known as Moth-flame optimization (MFO). Leveraging ELM's learning technique with one-pass sample processing is the driving force for its use. A classification model for Cancer Micrao array data was proposed using this hybrid feature selection method, in which ELM is also taken into consideration as a classifier. Model performance is assessed using several measuring indices.

H. A. Al-Alshaikh *et al.* (2024) [25] suggested method increases the robustness of the model by selecting pertinent features by combining the RFE method with the GA. By addressing the problem of data imbalance, approaches such as the under sampling clustering oversampling technique (USCOM) improve the predictive power of the method. A multilayer deep convolutional neural network (MLDCNN) trained with the adaptive elephant herd optimization (AEHO) is used in the classification. According to its thorough evaluation, the suggested ML-based heart disease prediction method (ML-HDPM) performs remarkably well across several significant assessment metrics. We created a risk stratification algorithm based on machine learning to predict CHD. S. Zhang *et al.* (2025) [26] developed an ML-based approach for CHD prediction for both feasible and effective. To guide main prevention and CHD care, the risk scores established potential in recognizing pregnant women at high risk for fetal CHD. K. Jha *et al.* (2025) [27] developed ML-based models which are comprised of ANN, SVM, DT, and RF. The ANN performed commendably in contrast to the other methods, attaining 86% accuracy, 86% precision, 84% recall, and 83% F1 score. On the other hand, ANN was shown to be the most reliable method for forecasting heart disease. According to the study's findings, ANN has the best forecast power and is therefore a capable technique for the early analysis of heart disease. Several SI techniques are employed in the literature to tune ELM hyper parameters. However, the repetitive searching of the SI algorithms frequently results in a large increase in calculation time. The optimization algorithm's hyper parameters have an impact on performance. Particularly in high-dimensional search spaces, it could become stuck in local optima, which would improve training accuracy but not always generalize to new data.

3. Extreme Learning Machine (ELM)

ELM is renowned for its quick training times and effective learning procedures [28]. The ELM is a feedforward neural network (FFNN) with a single hidden layer that offers extremely rapid training by initializing input weights at random and calculating output weights systematically. Because ELM does not rely on iterative learning as neural networks do, it is very active and suitable for applications involving great amounts of data. It was created to help with some of the computational issues with conventional neural networks, especially training complexity and duration. Its emphasis on speed and simplicity gives it a distinctive approach to neural network learning, which makes it ideal for applications demanding quick reaction times and little computational expense. For many real-world applications, ELM is a useful tool and a good choice when rapid, effective learning is crucial, even though it might not be as effective for extremely complicated tasks. *For* N arbitrary different data objects (x_i, t_i). Where $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T \in R^n$, $t_i = [t_{i1}, t_{i2}, \dots, t_{in}]^T \in R^m$. The SLFNs have H hidden nodes with $g(.)$ activation that can approximate these N data samples with zero error, this means that

$$Hwo = T \tag{1}$$

Where,

$$H = \begin{bmatrix} g(wh_1.x_1 + b_1) & \dots\dots & g(wh_1.x_1 + b_H) \\ \cdot & \dots\dots & \cdot \\ \cdot & & \cdot \\ g(wh_1.x_N + b_1) & \dots\dots & g(wh_H.x_N + b_H) \end{bmatrix}_{N\times H} \quad Hw_0 = \begin{bmatrix} {w_o}_1^T \\ \cdot \\ \cdot \\ {w_o}_H^T \end{bmatrix}_{H\times m} \quad T = \begin{bmatrix} t_1^T \\ \cdot \\ \cdot \\ t_H^T \end{bmatrix}_{N\times m} \tag{2}$$

Where, $wh_i = [wh_{i1}, wh_{i2}, \dots., wh_{in}]^T$ - represents the connection weights for i^{th} hidden and input neurons. $wo_i = [wo_{i1}, wo_{i2} \dots., wo_{im}]^T$- represents the connection weights for i^{th} hidden and output neurons. b_i - represents the bias of the hidden neuron. Thus, the purpose of the output weights is to find the least - square (LS) solution to the linear system is

$$wo = H^{+}T \tag{3}$$

Where, H^{+} is the MP inverse of matrix H. Among all LS solutions, the lowest norm LS solution is distinct and has the smallest norm. ELM that uses such an MP inverse technique frequently achieves strong generalization performance. The ELM converges significantly faster than other gradient-based learning algorithms since the solutions are obtained using an analytics method, and all SLFN parameters do not need to be modified.

4. Bacterial Colony Optimization

BCO is a modern evolutionary approach established by Niu *et al.* (2012) [29]. The key distinction between BCO and other bacteria-inspired heuristic algorithms is that, contrasting other bacterial-inspired heuristic algorithms, BCO searches for nutrients by using bacterial communication rather than swimming at random. The three straightforward stages of the BCO are chemotaxis and communication, elimination and reproduction, and migration. The BCO is used to address a variety of real-world applications [10, 30-34]. Throughout the whole BCO lifecycle, chemotaxis is continually accompanied by communication. After a lengthy period of chemotaxis and communication, bacteria have two options. If they can locate food on their own, they may be able to breed or they may starve to death. Some individuals could encounter dangerous situations in a problematic situation by pushing the bounds or searching for space. Chemotaxis and communication are used during the whole BCO process. Conversely, the other two phases are only performed under certain situations, such as after a predetermined number of cycles have been completed, the randomly generated number falls below a predetermined probability, etc. Chemotaxis, like swimming and tumbling, can be modeled after two distinct generations. The swimming process is enhanced by a stochastic direction. Both a perfect searching director and a chaotic director combined have an impact on the search positioning. Each bacterium's position is as follows:

$$Position_i(T) = Position_i(T-1) + C(i) * [f_i * (G_{best} - Position_i(T-1)) + (1-f_i) * (P_{best_i} - Position_i(T-1)) + turb_i] \tag{4}$$

The bacteria will swim towards their optimum location if there is no turbulent. Hence, the bacterium's position will be restructured as follows,

$$Position_i(T) = Position_i(T-1) + C(i) * [f_i * (G_{best} - Position_i(T-1)) + (1-f_i) * (P_{best_i} - Position_i(T-1))] \tag{5}$$

Where, $turb_i$ denotes turbulent direction variance value. $f_i \in \{0,1\}$. P_{best} denotes the personal best. G_{best} denotes the global best. $C(i)$ is a chemotaxis step size which is defined as follows,

$$C(i) = C_{\min} + \left(\frac{Iter_{\max} - Iter_j}{Iter_{\max}}\right)^n (C_{\max} - C_{\min}) \tag{6}$$

Where, $Iter_{max}$ and $Iter_j$ are the maximum and present iterations, respectively. n - Chemotaxis step size. During the elimination and reproduction phase, the high-energy bacteria will self-replicate to produce new individuals, replacing the unhealthy ones. Energy-dense bacteria make great nutrient searchers. If certain conditions are met, the bacteria can move during the final phase, called migration, inside the search range.

5. Methodology

In this part, the suggested BCO-ELM optimized predictive model for disease prediction in cardiac disease data is introduced. The present research work focused on three stages for analyzing the performance of the proposed BCO-ELM such as filling in missing values, selecting relevant features, and detecting using the proposed BCO-ELM. The feature imputations are performed by using MICE, relevant features are selected using the REF method, and disease detection is done by BCO-ELM. The BCO-ELM is used to classify whether the person is diseased or normal. An overview of the disease detection technique is shown in Figure 1. To improve diagnosis accuracy, the suggested system for detecting heart disease uses an intelligent and structured pipeline. The MICE approach is first used for feature imputation on the Heart dataset. The MICE efficiently fill in missing data by making realistic values based on the associations between variables. Before additional processing, this phase promises the dataset's extensiveness and honesty. RFE, which reduces the dimensionality and expands model performance by keeping just the most important features, is used for feature selection after imputation.

This reduces the chance of overfitting while also enhancing computing efficiency. The framework combines BCO-ELM for heart disease identification. The BCO gets around issues with random weight initialization and enhances the model's generalizability by optimizing the ELM model's parameters. This hybrid model produces more precise classification results by utilizing both the broad global search ability of BCO and the rapid learning ability of ELM. After that, the proposed approach splits people into groups based on whether they have a disease or not. To regulate the effectiveness of the suggested approach, a systematic performance analysis is carried out, assessing the accuracy, sensitivity, specificity, precision, and F-measure. A major factor is the measured combination of imputation, feature selection, and swarm-intelligence-based optimization. The following subsection is discussed in detail.

5.1 Dataset

A well-known dataset, the Cleveland heart disease dataset, can be accessed from UCI (University of California, Irvine). A variety of clinically applicable variables and actual patient data make the Cleveland Heart Disease Dataset a vital tool for making and assessing heart disease prediction models. These features record numerous aspects of the patient's health, including age, sex, type of chest discomfort, cholesterol, ECG analyses, and more. The dataset is suitable for visions with medical implications because a number of its properties also fit in nicely with recognized clinical criteria. Bias in illustration is an important additional concern. Because there are more male patients than female patients, prediction models may be skewed and less precise for underrepresented groups. Additionally, certain features, notably NMVCE and thalassemia (THAL), have missing or unclear data. Appropriate data cleaning or imputation techniques are required to address this. Furthermore, unless corrective measures like resampling or cost-sensitive learning are used, class imbalance may result in skewed predictions toward the majority class. On the other hand, the dataset's properties, which are often working in medical diagnostics and have a thorough correlation to cardiovascular health, are clinically maintained and still very applicable. Because of their high extrapolative potential, these features are frequently kept when using feature selection methods such as RFE. In addition to enlightening model correctness, their inclusion encourages clinician hope and interpretability, both of which are serious in healthcare applications.

In this study, the dataset was used to assess the robustness of ML-based prediction systems. The dataset includes two different classes of people—heart disease and healthy—along with 303 data samples (patients), 76 characteristics, and 76 data points. The dataset is summarized in Table 1. In this dataset, information on patients with ages ranging from 29 to 77 is gathered from both male and female patients. Heart disease symptoms include chest discomfort. There are four forms of chest pain: “asymptotic, non-angina discomfort, typical angina, and typical atypical angina”. RBP represents the patient's resting blood pressure. The patient's cholesterol level is shown by the code SCH. In FAB, the level of fasting blood sugar is indicated. This feature stores 1 if the blood sugar level is greater than 120 mg/dl; otherwise, it stores 0. Electrocardiographic results are available via RECR, and MHR stores the patient's Maximum heart rate. If a person experiences exercise-induced angina, their EIGA score is 1, otherwise, it is 0. Exercise-induced ST depression is recorded in STD, which has upsloping, downsloping, and flat values that are denoted by 0, 1, and 2.

Table 1. Explanation of the dataset

S. NO	Code (Attribute)	Particulars (Values)	Values (min-max)
1	AGE (Age)	Age	29 - 77
2	SEX (Sex)	Male (1) Female (0)	0-1
3	CPT (Chest Pain Types)	Atypical angina (1), Typical angina (2), Asymptomatic (3), Non-anginal pain (4)	1-4
4	RBP (Resting Blood Pressure)	Mm Hg	94-200
5	SCH (Serum Cholesterol)	Mg/dl	120-564
6	FBS (Fasting blood sugar) (less than 120 mg / dl)	True (1) False (0)	0-1
7	RECR (Resting Electrocardiographic Results)	Having ST-T=1	1-2
8	MHR (Maximum Heart Rate)	-	71-202
9	EIA (Exercise-Induced Angina)	Yes (1) No (0)	0-1
10	STD (ST (Old peak) depression)	-	0 - 6.2
11	SPE (Slope of the ST segment peak exercise)	Up sloping (1) Flat (2) Down sloping (3)	1-3
12	NMVCE (Number of major vessels colored by fluoroscopy)	-	0-3
13	THALM (Thallium Scan)	Normal (3) Fixed defect (6) Reversible defect (7)	3-7
14	TARG (Target)	Disease (1) Normal (0)	0-1

SPE stands for the slope of peak exertion**.** The number of major vessels that are fluorescently colored is reported in NMVCF. The TARG attribute reveals whether or not a person has heart disease. There are five possible values for this feature: 0 for no heart disease, and 1 to 4 for varying degrees of the condition. To identify the existence of the disease, levels 1-4 are combined. There are 6 instances of missing values in the dataset. The NMVCF feature has four missing values, whereas the THALM feature has two missing values. Eight features are selected by using RFE such as SEX, CPT, RECR, EIA, STD, SPE, NMVCF, and THALM**.**

5.2 Data preprocessing method using MICE

The dataset has six missing values which are two in THALM and four in the NMVCF feature. Hence, these missing values were imputed using the MICE method [35]. MICE is essential for heart disease datasets because it fills in missing values while honoring the complex interrelationships between medical features. By dependably handling missing data, this approach produces predictive models that are more impartial and accurate. Utilizing MICE guarantees that the model for predicting heart disease operates efficiently, offering significant assistance in clinical and diagnostic settings where precise data is crucial for patient results.

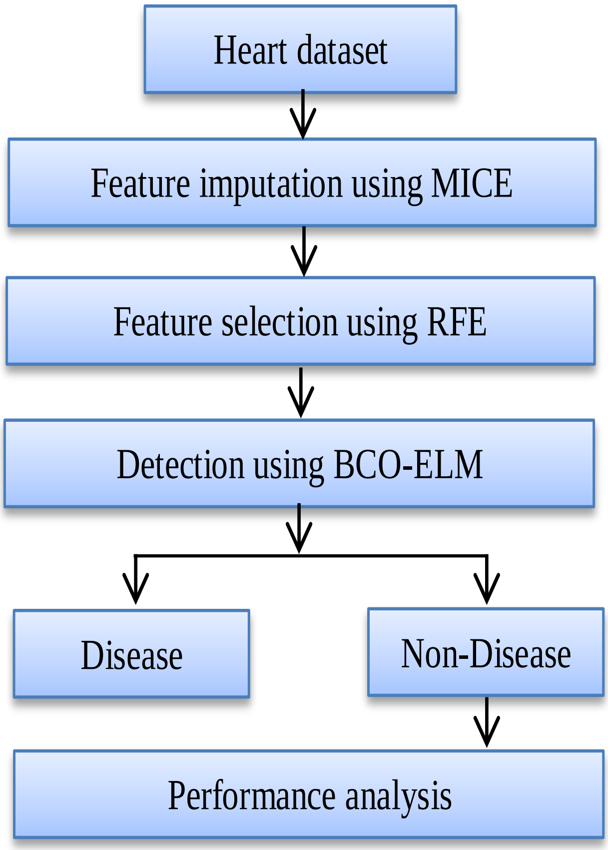

Figure 1. Overview of the proposed prediction system

The MICE performs imputation multiple times and assumes randomly that those missing values. Then, the regression method is used to predict the missing attribute values from the remaining values of the dataset. The steps of the MICE algorithm are shown in Figure 2.

5.3 Feature selection

One well-liked feature selection method for locating the most pertinent features in a dataset is RFE, especially for intricate datasets like those about heart disease. Many characteristics are involved in the prediction of heart disease, although not all of them may be equally significant for prediction accuracy. RFE improves model performance and interpretability by methodically removing the least significant features. RFE models eliminate unnecessary features iteratively [36]. Based on the training data, a classifier is trained, and feature importance is computed. The weakest features are eliminated at each stage of the procedure, and the model is then requalified using the lasting subset of features. Till the necessary number of features is reached; these procedures are repeated recursively.

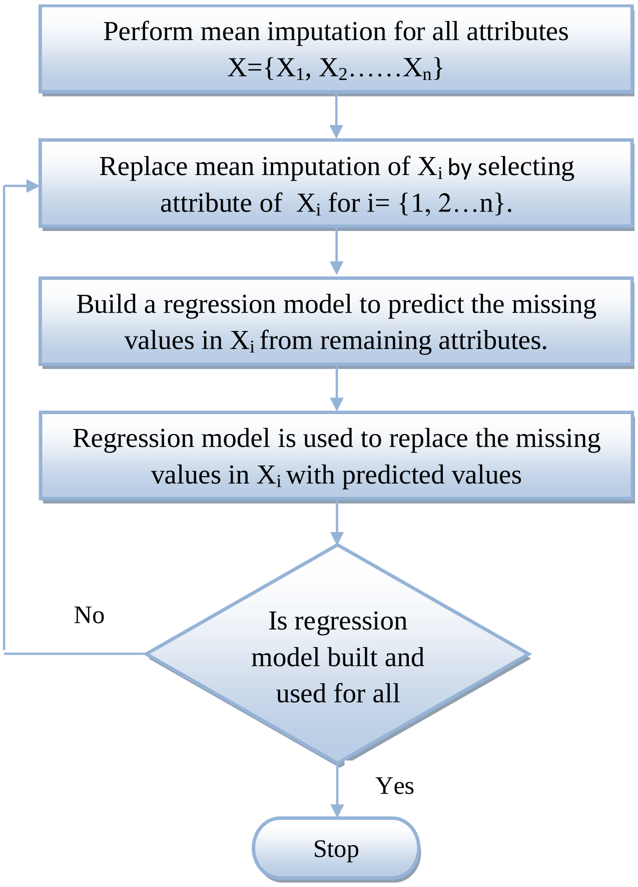

Figure 2. MICE algorithm

The algorithms receive as a parameter the number of characteristics to be kept. The steps of the RFE algorithm are shown in Figure 3. By concentrating the model on the most informative variables, RFE plays a crucial part in feature selection for datasets related to heart disease. In addition to decreasing overfitting and increasing forecast accuracy, this method also lowers processing costs and increases interpretability. To generate reliable cardiac disease prediction models that medical practitioners can rely on and use, RFE helps design a more robust and efficient model by choosing a core set of features.

5.4 Proposed optimized BCO-ELM

ELM has been established as a quick and effective learning technique that attains respectable performance when compared to the conventional gradient descent learning algorithm. In the ELM model, biases and weights are initialized at random. The ELM model's prediction stability suffers as a result of this method's inability to guarantee the correctness of the threshold and weight values. The random thresholds and weights also have an impact on training accuracy. The BCO method chooses the random parameter values in light of the limitations of the ELM model. Therefore, choosing the best ELM parameter settings is a major challenge for attaining improved generalization performance.

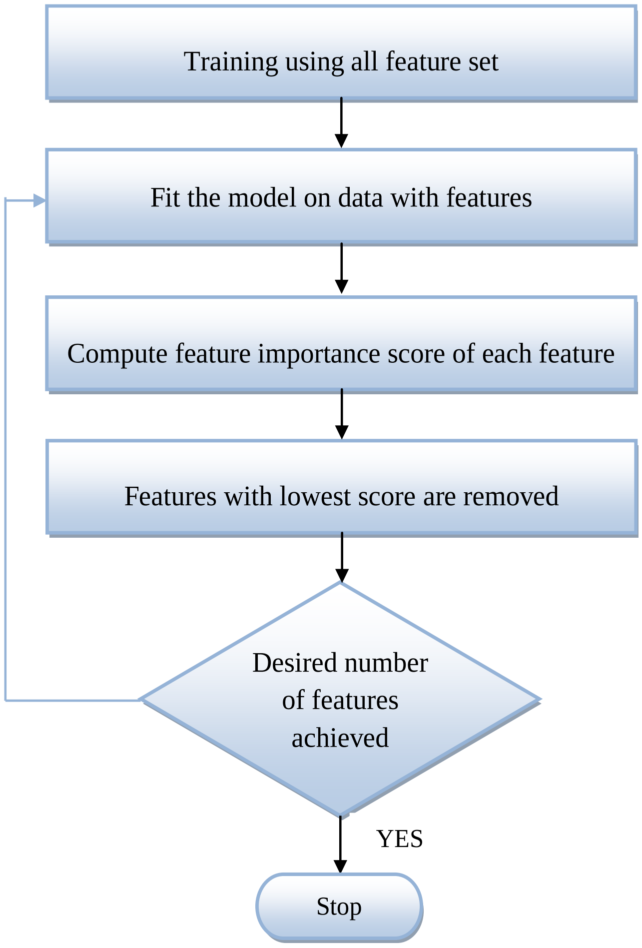

Figure 3. RFE algorithm

Numerous studies have been carried out to optimize the ELM's parameters. For instance, the ELM algorithm has been improved using a genetic algorithm (GA) and it has been used to anticipate short-term traffic flow [37]. The initial weights and biases for the ELM are chosen using the PSO [38]. The proposed approach employs an optimization method based on the BFO to change the weights and biases for ELM [39]. However, the GA, PSO, and BFO have several drawbacks, including a slow convergence rate and premature convergence. To obtain high prediction accuracy, the suggested solution uses the BCO algorithm to improve the ELM algorithm parameter. BCO efficiently discovers complex solution spaces by simulating the behavior of bacterial colonies in locating nutrients through an ordered yet flexible search mechanism. Its method strikes a balance between local and global exploration. BCO can avoid local optima from convergent too soon. When dealing with enormously nonlinear/multimodal landscapes, which are dominant in many real-world applications, BCO's ability to energetically adjust its search behavior is helpful.

The proposed BCO-ELM optimizes important ELM parameters including bias and weights. The BCO approach uses the means square error (MSE) as a fitness function in the BCO-ELM. In the BCO-ELM, finding the optimal ELM hyperparameters, such as the number of hidden neurons, input weights, and biases, is the core goal of reducing error (MSE). Each bacterium denotes a candidate solution and these hyperparameters make up the search space for BCO. A bacterial population is initialized at random to start the optimization process. Every bacterium is determined as a vector with values for input weights, biases, and the number of hidden neurons. The optimal bacteria are selected as the best parameters for ELM which has the highest fitness, and the final ELM model is configured using its related parameters. High accuracy with better generalization is attained by the approach through ELM optimization with BCO for heart disease prediction. By effectively searching high-dimensional spaces, needing fewer control parameters, and adaptively balancing exploration and exploitation, BCO provides advantages over conventional tuning techniques. This makes it a clever and well-organized technique for optimizing ELMs. Simple stages for the suggested BCO+ELM are shown in Algorithm 1 and Flowchart 4. The same research work has been conducted, presented at conferences, and published [40]. However, the conference paper is the only one to take into account parameter optimization. To improve performance and solution correctness, the current research focused on filling in the missing values and choosing the pertinent features of datasets related to heart disease.

In the heart disease prediction pipeline, strong data preprocessing, feature selection, and optimized prediction models are crucial for enhancing predictive performance. To address missing data, MICE is applied, which iteratively methods each imperfect variable based on others, protecting inter-variable relationships and minimizing bias. Following preprocessing, RFE is applied to detect the most appropriate features by recursively training a method and removing the least significant features, thereby reducing dimensionality and refining effectiveness. For classification, an ELM is implemented due to its High-speed learning ability. To overcome its sensitivity to randomly initialized parameters, BCO is applied to adjust the input weights and biases. Inspired by the social behavior of bacteria, BCO enhances the ELM's generalization and accuracy, resulting in an effective and optimized model for heart disease prediction. Overall, the combination of optimization-driven modeling, smart feature selection, and innovative imputation produces a very precise, dependable, and well-organized approach to predicting heart disease, which may help with early detection and expand patient outcomes.

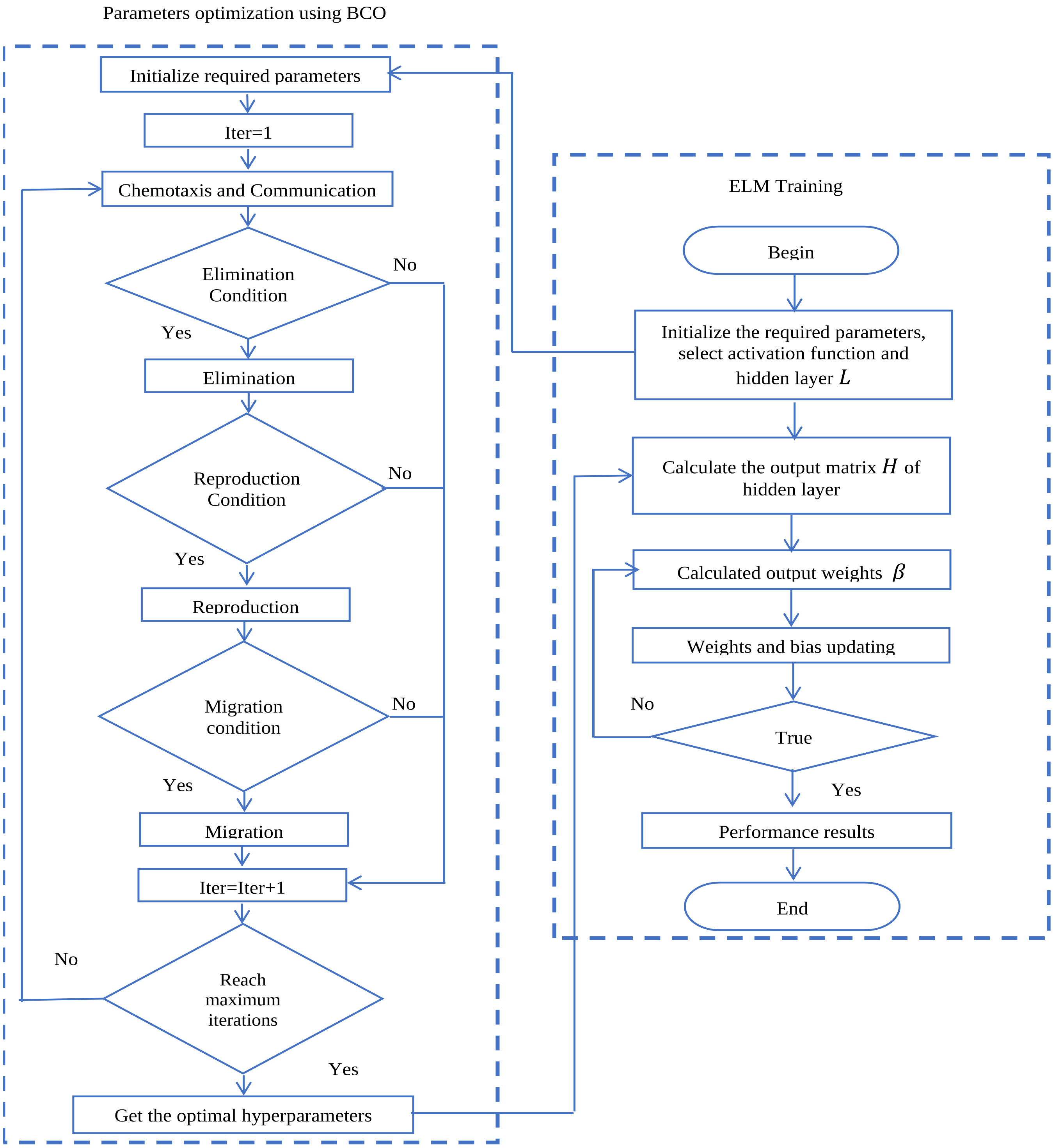

Figure 4. Flowchart of BCO+ELM

.

Algorithm 1. BCO+ELM

```
BEGIN
Step 1: Initialize the necessary parameters
Step 2: Chemotaxis and communication
Step 3: Elimination and reproduction
Step 4: Migration
Step 5: Find the optimum best
Step 6: perform Step 2 if the termination condition is not satisfied; else, step 7.
Step 7: Perform testing using optimum values
Step 8: De-normalize the output
```

6. Experimental Results Analysis

The experimental findings guarantee that a BCO-optimized ELM for heart disease prediction satisfies both technical and clinical requirements by offering a thorough assessment of its efficacy and viability. In the end, this analysis helps show that the model is feasible for possible clinical uses. MATLAB 2015b is utilized to create compared prediction algorithms. The dataset's specifics, parameter settings, and discussion of the results were covered in this section. The suggested BCO+ELM algorithm's stability is compared to that of MLP [41] and a few ELM [39] algorithm versions, including GA+ELM [42], PSO+ELM [43], IPSO+ELM [44], and BFO+ELM [45]. The averages of 50 trials were used to generate the results of the findings in this study.

6.1 Parameter settings

All algorithm's best parameters show a significant role in positively a solution's performance. Both BCO and ELM have several parameters which decide the performance of given solutions. Hence, the present section discusses its parameter values and Tables 2 and 3 show the parameters and values for BCO and ELM respectively. The convergence rate of BCO is determined according to its two main features such as chemotaxis step (N_c) values and swim step (N_s). A high chemotaxis value will require more computational time. Henceforward, the current paper chooses a small number of chemotaxis steps as $N_C = 100$. Likewise, the small value of the swim step is selected as $N_s = 4$. The reproduction value is selected as $N_{re} = 4$, and the dispersal step value is selected as $N_{ed} = 2$. The step size and probability of elimination and dispersal values are all significant features in defining the BCO algorithm's performance. Two different step size values are measured such as the lowest (C_{min}) and highest (C_{max}) which are selected as 0.015 and 0.2 respectively.

An elimination and dispersal probability P_{ed} value is another factor of BCO which is helps to escape from the local optima problem. Hence, the best performance is produced when selecting a probability value is 0.25.

The ELM algorithm has many parameters including several input, hidden, and output neurons, learning rate, error rate, and epochs which decide the performance. The present research work uses single input, hidden, and output layers in the ELM algorithm. The number of input neurons is selected based on application. For heart disease prediction, 13 and 8 are used before and after feature selections are considered. The number of hidden neurons is selected as 20. The weights and bias are selected between the ranges of 0 and 1. The TanH is selected as an activation function.

Table 2. BCO Parameter settings

Factors	Values	Factors	Values
S	100	N_{ed}	2
N_C	100	P_{ed}	0.26
N_{re}	4	C_{min}	0.015
N_s	4	C_{max}	0.2

Table 3. Parameter setting of ELM

Parameters	Values
Number of Input neurons before feature selection	13
Number of Input neurons after feature selection	8
Number of Hidden neurons (L)	20
Number of output neurons	1
Learning rate	0.4
Error rate	0.0005
Epochs	1000
Weights and bias range	0 to 1
Activation functions	TanH

6.2 Performance measures

The proposed prediction method such as BCO-ELM computed its performance-based performance measures.

- *Accuracy* is a performance measure of the prediction method and how well a system can predict the future which is defined as follows,

$$Accuracy = \frac{TP+TN+FP+FN}{TP+FP} \quad (7)$$

- *Sensitivity* is a performance measure that assesses a system's ability to correctly predict positive outcomes which is defined as follows,

$$Sensitivity = \frac{TP}{TP+FN} \quad (8)$$

- *Specificity* is a performance metric that assesses the system's ability to correctly forecast the negative which is defined as follows,

$$Specificity = \frac{TN}{TN+FP} \quad (9)$$

- *Precision* is used to analyze the ability of a system to deliver only suitable results, which is defined as follows,

$$Precision = \frac{TP}{TP+FP} \quad (10)$$

- *F-Measure* combines the results of precision and sensitivity using the harmonic mean which is defined as follows,

$$F - Measure = 2 \times \frac{Sensitivity \times Precision}{Sensitivity + Precision} \quad (11)$$

Where, *True Positive (TP)* represents that the patient's heart disease is accurately predicted by the model. *False Positive (FP)* is a false alarm that occurs when the approach incorrectly guesses heart disease in a healthy patient. *True Negative (TN)* represents that the patient's health is precisely predicted by the approach. *False Negative (FN)* represents that the patient has heart disease yet the model indicates when they are healthy.

6.3 Results analysis

A study on Optimized ELM utilizing BCO for heart disease prediction must include a results analysis and discussion section to fully comprehend the effectiveness of the suggested model in comparison to more conventional approaches and the importance of the optimization strategies used [46]. The performance of the suggested BCO-ELM prediction system is discussed in the following section. It is possible to quantify the improvement in accuracy, sensitivity, specificity, and precision by comparing the BCO-optimized ELM with conventional models or ELMs without optimization. These measures show the model's ability to distinguish between patients with and without heart disease, which is critical for clinical decision-making in heart disease prediction. Since BCO is intended for rapid optimization, the investigation should determine whether it accelerates ELM training without sacrificing quality. A quicker convergence makes the model more useful for predicting cardiac illness in real-time or almost real-time.

Examine whether BCO determines the best weights and biases for the ELM, producing predictions that are more accurate and trustworthy.

Table 4. Performance of prediction algorithm on whole feature sets

Prediction model	Accuracy	Sensitivity	Specificity	Precisions	F-Measures
MLP	78.15	84.28	76.09	79.65	82.39
ELM	81.28	87.09	83.16	82.49	85.16
GA-ELM	83.82	89.26	86.85	85.45	86.79
PSO+ELM	86.79	90.54	89.74	88.71	88.62
IPSO+ELM	89.52	93.27	90.92	90.86	90.61
BFO+ELM	90.69	95.42	92.27	93.48	92.93
BCO+ELM	**92.27**	**97.48**	**95.34**	**95.61**	**94.27**

Table 5. Performance of prediction algorithm after imputation

Prediction model	Accuracy	Sensitivity	Specificity	Precisions	F-Measures
MLP	79.13	83.24	78.84	80.59	83.47
ELM	81.84	88.20	83.97	82.95	86.11
GA-ELM	84.06	89.92	87.15	86.07	87.82
PSO+ELM	86.94	91.58	89.84	88.85	89.62
IPSO+ELM	89.51	93.92	90.92	90.70	92.01
BFO+ELM	91.28	95.62	93.28	94.46	92.93
BCO+ELM	**92.67**	**98.43**	**96.32**	**96.47**	**94.89**

Table 6. Performance of prediction algorithm after feature selection

Prediction model	Accuracy	Sensitivity	Specificity	Precisions	F-Measures
MLP	81.67	85.96	81.46	82.18	84.59
ELM	83.28	88.85	84.71	83.73	87.94
GA-ELM	84.61	90.37	87.65	86.64	88.64
PSO+ELM	87.08	93.08	90.72	89.28	90.86
IPSO+ELM	89.75	94.19	91.36	91.09	92.84
BFO+ELM	93.47	97.48	94.84	94.07	94.17
BCO+ELM	**95.43**	**99.10**	**97.73**	**97.63**	**95.54**

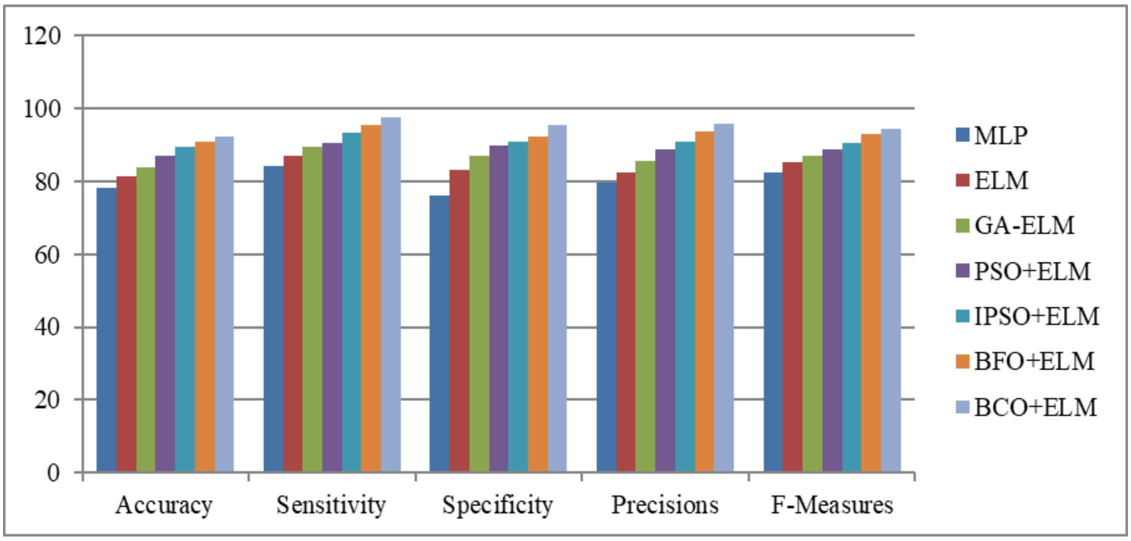

Figure **5.** Performance of prediction algorithm on whole feature sets

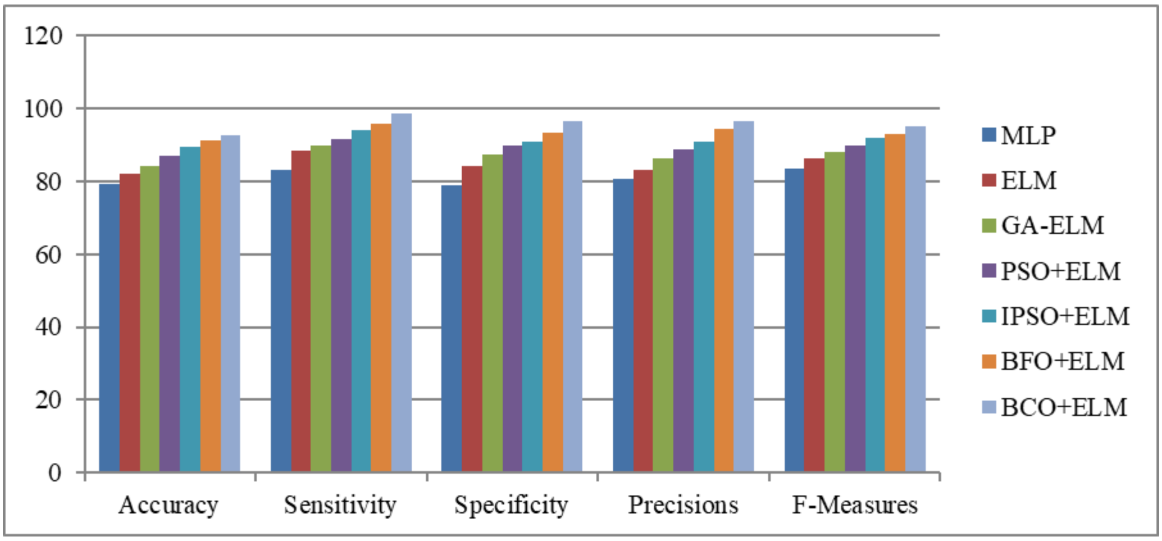

Figure 6. Performance of prediction algorithm after imputation

The quality of the solution has a direct impact on how well the model generalizes to new inputs. The three stages of the process have been carried out in the suggested procedure including filling in missing values, choosing appropriate features, and obtaining optimal parameters of ELM. The effectiveness of the detection system is predicated on various phases, such as raw data, after filling in the missing values, and after choosing appropriate features. The computation time is taken into consideration when analyzing the performance. The suggested BCO+ELM algorithm's stability is compared to that of MLP [41] and a few ELM [39] algorithm versions, including GA+ELM [42], PSO+ELM [43], IPSO+ELM [44], and BFO+ELM [45].

Performance comparisons of the proposed algorithm are shown in Tables 4, 5, and 6, as well as in Figures 5, 6 and 7. The performance of the whole feature set for various prediction algorithms, including the ELM [39] and MLP [41] variants, is displayed in Table 4. The BCO+ELM model performs significantly better than the others in the first scenario, where the models are trained using the entire feature set without any preprocessing. Its accuracy is 92.27%, sensitivity is 97.48%, specificity is 95.34%, precision is 95.61%, and F-measure is 94.27%. According to this, BCO+ELM is perfect for balanced classification jobs since it accurately detects both positive and negative cases. The BFO+ELM [45] model performs admirably, albeit with somewhat lower results (accuracy: 90.69%, sensitivity: 95.42%, F-measure: 92.93%). Additionally, IPSO+ELM [44] performs well (accuracy: 89.52%), suggesting that classification reliability is increased by combining PSO with enhanced exploration capabilities. The limitations of conventional learning methods without optimization are highlighted by the fact that ELM [39] and MLP [41] display the lowest metrics, whereas PSO+ELM [43] and GA-ELM [42] get intermediate outcomes.

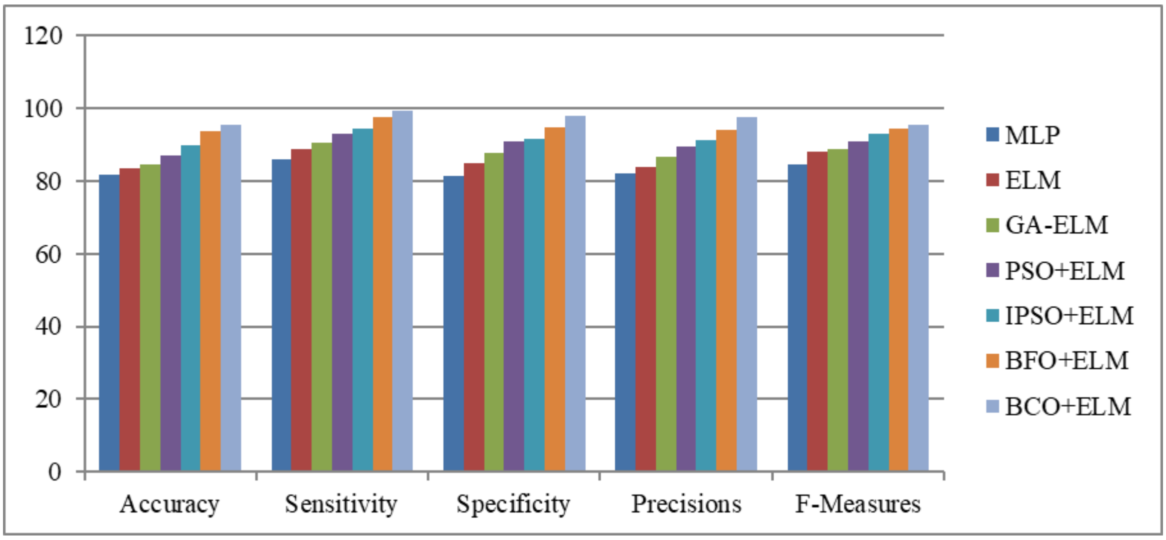

Figure 7. Performance of prediction algorithm after feature selection

Table 7. Performance of prediction algorithm based on computation time (Seconds)

Prediction model	Whole Features	After imputation	After feature selection
GA-ELM	9.5674	9.5187	8.5674
PSO+ELM	10.2874	10.1594	9.4967
IPSO+ELM	9.3687	9.2682	8.8517
BFO+ELM	8.2675	8.2196	7.6585
BCO+ELM	**7.4975**	**7.4837**	**7.0197**

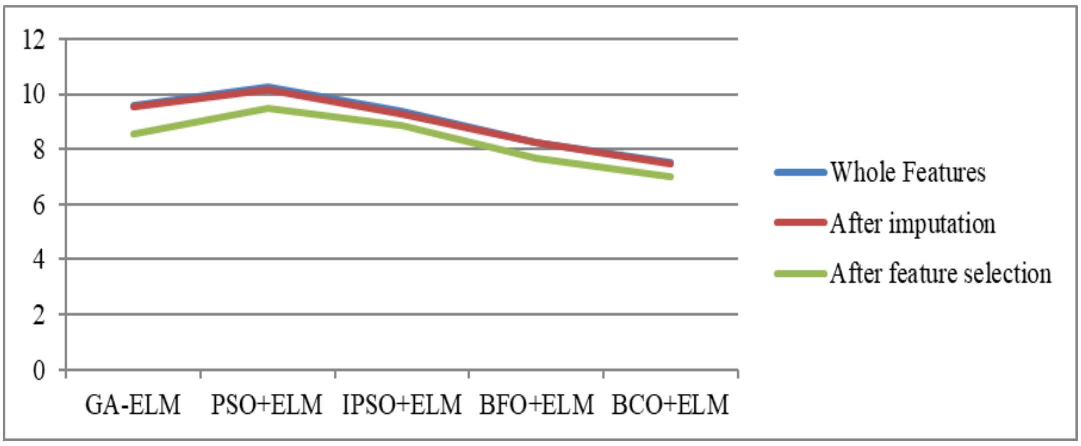

Figure 8. Computational time analysis

The performance analysis following imputation is displayed in Table 5. All models show better or stabilized performance after using imputation to manage missing variables, particularly in sensitivity and F-measure. With a minor performance improvement (accuracy: 92.67%, sensitivity: 98.43%, F-measure: 94.89%), BCO+ELM holds its lead once more, demonstrating its resilience while handling defective or incomplete datasets. The efficacy of BFO+ELM [45] (accuracy: 91.28%) and IPSO+ELM [44] (accuracy: 89.51%) in managing noisy real-world data is confirmed by their continued great predictive power, especially in sensitivity and precision. The significance of hybrid techniques in improving data-driven decision-making post-imputation is further supported by the fact that MLP [41] and ELM [39] continue to lag while PSO+ELM [43] and GA-ELM [42] demonstrate modest gains.

The examination of performance results following feature selection is displayed in Table 6. All model's predicted accuracy is greatly increased by the use of feature selection. The BCO+ELM model's remarkable accuracy of 95.43%, sensitivity of 99.10%, specificity of 97.73%, precision of 97.63%, and F-measure of 95.54%—the highest numbers ever recorded in the experiment—best illustrates this. This demonstrates both the model's capacity for categorization and the efficiency of eliminating superfluous or irrelevant features. Feature selection also helps BFO+ELM [45] and IPSO+ELM [44] a lot. BFO+ELM [45] obtains 93.47% accuracy and a remarkable 94.17% F-measure, while IPSO+ELM [44] reaches 89.75% accuracy and 92.84% F-measure. These findings demonstrate how neural learning and feature reduction, in conjunction with evolutionary and swarm-based optimization techniques, can greatly improve classification performance. MLP [41] and ELM [39], on the other hand, make slight progress but still lag well behind the optimized hybrid models. Table 8 shows the average accuracy of developed methods. The optimized method achieved the highest accuracy in all situations, showing its strong ability in learning from features, handling imputed data, and benefiting from feature selection.

Average accuracy growth from MLP [41] to BCO+ELM across all circumstances is almost 13.8% to 14.1%, which is important in predictive modeling. BCO+ELM consistently yields the best scores in all important performance parameters, making it the most reliable and successful model across all experimental setups. The model's versatility and exceptional learning ability are demonstrated by its outstanding performance following both imputation and feature selection. The competitiveness of BFO+ELM [45] and IPSO+ELM [44], particularly after pre-processing, demonstrates how hybrid optimization techniques greatly improve ELM's predictive performance. The three tables clearly show that, even in cases of high dimensionality or inadequate data, combining metaheuristic optimization—particularly BCO and BFO—with ELM offers a strong foundation for accurate and trustworthy classification.

Table 8. Performance comparison based on average accuracy

Models	Average Accuracy (%)
MLP	79.65
ELM	82.13
GA-ELM	84.16
PSO+ELM	86.94
IPSO+ELM	89.59
BFO+ELM	91.81
BCO+ELM	93.46

The computation complexity of ELM is determined as $\mathrm{O}(\mathrm{NdL}+\mathrm{NL}^2\mathrm{L}^3)$. The computational complexity of BCO is determined as $\mathrm{O}(\mathrm{GPd})$. Hence, the computation complexity of BCO-based ELM is determined as $\mathrm{O}(\mathrm{GP}(\mathrm{NdL}+\mathrm{NL}^2\mathrm{L}^3))$. Here, P represents the number of bacteria, G denotes the number of generations, d is the dimensionality of the solution vector (weights and biases of ELM). N is the number of training samples. n is the number of input features. L is the number of hidden neurons.

The computational time is taken into account when evaluating the effectiveness of the suggested BCO-ELM, which is depicted in Table 7 and Figure 8. According to Table 7, the suggested BCO-ELM outperforms various ELM variants in terms of using complete features, filling in missing values, and selecting the most appropriate features all the while requiring less computation time. For example, the value of using complete features, filling in missing values, and selecting the most appropriate features are produced at 7.4975, 7.4837, and 7.0197 respectively. In terms of using all features, after imputation, and after choosing the best features, the suggested BCO-ELM outperformed other variants of the ELM and MLP algorithms, according to the evaluation of the overall findings.

Figure 9 shows the AUC analysis and the most consistent performance is shown by BCO-ELM, which reliably preserves a high TPR over a broad range of FPR among the models measured. This proposes that BCO-ELM is better at distinguishing across classes, lessening classification mistakes while increasing detection accurateness. Additionally, BFO-ELM [45] performs competitively, mainly in the mid-to-high FPR range, demonstrating that it is a good substitute when a small increase in FPR is acceptable in exchange for better sensitivity.

When compared to BCO-ELM, PSO+ELM [43], IPSO+ELM [44], and BFO+ELM [45] models perform discreetly, remaining near each other on the ROC curve but somewhat underperforming.

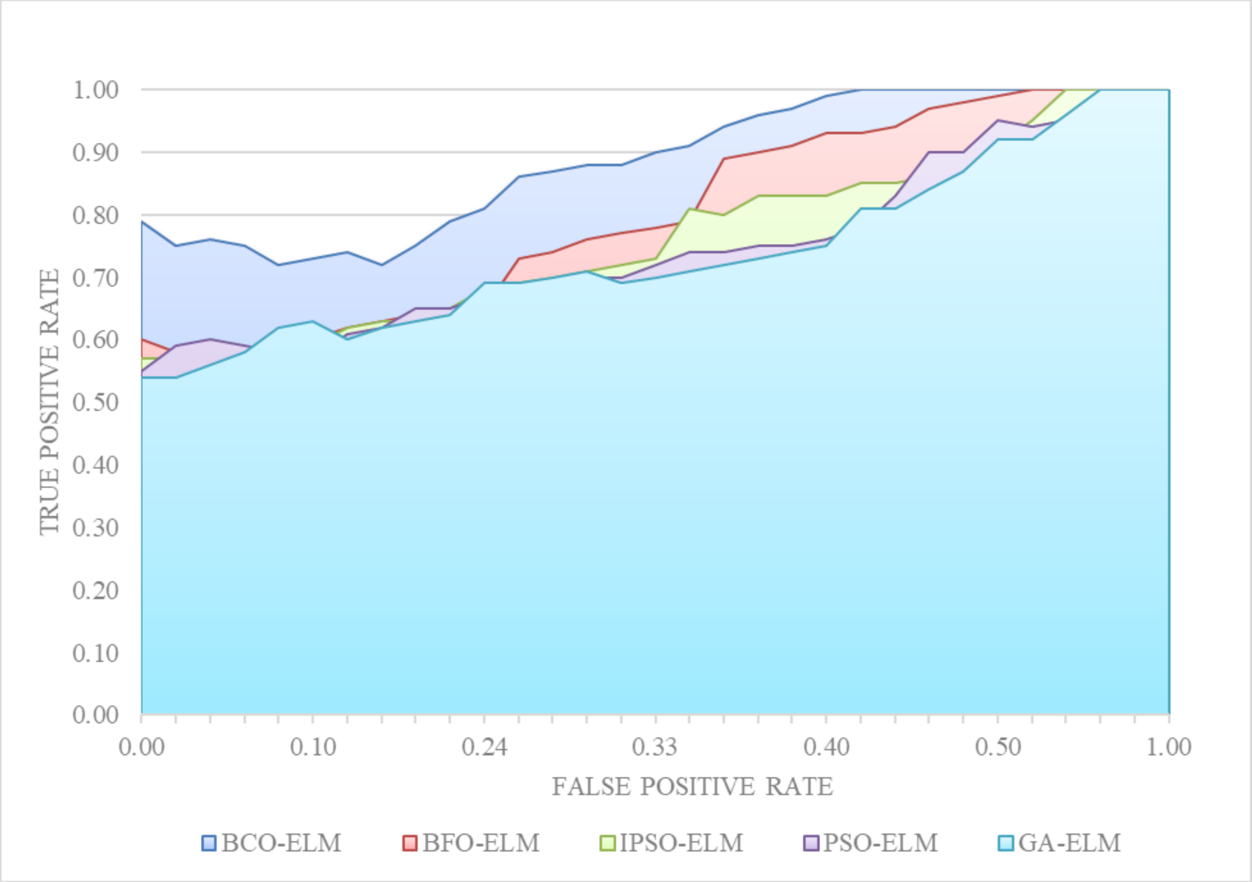

Figure 9. AUC analysis for compared methods

Although their efficacy decreases at key decision thresholds with reduced FPR, these techniques offer a balance between precision and recall. However, among the assessed models, GA+ELM [42] seems to be the least effective, particularly in the lower TPR, signifying poorer classification capability. With better trade-offs between TPR and FPR, BCO-ELM is the most talented technique for refining ELM performance in classification tasks, according to the analysis.

7. Conclusions

The current study work has suggested a new disease prediction technique based on optimized ELM to predict disease early. The optimized ELM uses the BCO algorithm for optimizing the parameters of ELM. The major goal of the suggested BCO-ELM prediction method is to diagnose heart disease more accurately than the current approach. Using heart disease datasets, the performance of the BCO-ELM is compared with some well-known prediction methods. Predictive accuracy and robustness can be greatly increased by efficiently optimizing the ELM model's parameters using BCO. By lowering overfitting and improving model generalization, this optimized model performs better than conventional ELMs and other popular machine-learning models. The experimental findings support the assertion that the suggested prediction method provided high accuracy compared to the current algorithms. However, the proposed performance is heavily dependent on the superiority and size of the dataset. Incomplete/imbalanced data can lead to biased or imprecise forecasts. Though BCO improves accuracy by optimizing ELM parameters, the optimization process presents additional computational complexity and runtime. Hence, handling class imbalance and new optimization method for hyper parameters optimization will be considered as future works.

References

[1] C.B.C. Latha, S.C. Jeeva, Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques. Informatics in Medicine Unlocked, 16, (2019) 100203. https://doi.org/10.1016/j.imu.2019.100203

[2] S.P. Patro, G.S. Nayak, N. Padhy, Heart disease prediction by using novel optimization algorithm: a supervised learning prospective. Informatics in Medicine Unlocked, 26, (2021) 100696. https://doi.org/10.1016/j.imu.2021.100696

[3] R. Kavitha, K. Saravanan, S.A. Jebakumari, K. Velusamy, Machine learning algorithms for IoT applications," Artificial Intelligence for Internet of Things: CRC Press, (2022) 185-214.

[4] Y. Zhao, E.P. Wood, N. Mirin, S.H. Cook, R. Chunara, Social determinants in machine learning cardiovascular disease prediction models: a systematic review. American journal of preventive medicine, 61(4), (2021) 596-605. https://doi.org/10.1016/j.amepre.2021.04.016

[5] J. Azmi, M. Arif, M.T. Nafis, M.A. Alam, S. Tanweer, G. Wang, A systematic review on machine learning approaches for cardiovascular disease prediction using medical big data. Medical Engineering & Physics, 105, (2022) 103825. https://doi.org/10.1016/j.medengphy.2022.103825

[6] G.S. Kushwah, V. Ranga, Voting extreme learning machine based distributed denial of service attack detection in cloud computing. Journal of Information Security and Applications, 53, (2020) 102532. https://doi.org/10.1016/j.jisa.2020.102532

[7] G.S. Kushwah, V. Ranga, Optimized extreme learning machine for detecting DDoS attacks in cloud computing. Computers & Security, 105, (2021) 102260. https://doi.org/10.1016/j.cose.2021.102260B

[8] F. Weng, Y. Chen, Z. Wang, M. Hou, J. Luo, Z. Tian, Gold price forecasting research based on an improved online extreme learning machine algorithm. Journal of Ambient Intelligence and Humanized Computing, 11, (2020) 4101-4111. https://doi.org/10.1007/s12652-020-01682-z

[9] M.T. Hussan, G.V. Reddy, P. Anitha, A. Kanagaraj, P. Naresh, DDoS attack detection in IoT environment using optimized Elman recurrent neural networks based on chaotic bacterial colony optimization. Cluster Computing, 27(4), (2024) 4469-4490. https://doi.org/10.1007/s10586-023-04187-4

[10] K. Tamilarisi, M. Gogulkumar, K. Velusamy, (2021) Data clustering using bacterial colony optimization with particle swarm optimization. fourth international conference on electrical, computer and communication technologies (ICECCT), (2021), IEEE, India. https://doi.org/10.1109/ICECCT52121.2021.9616695

[11] V. Prakash, V. Vinothina, K. Kalaiselvi, K. Velusamy, An improved bacterial colony optimization using opposition-based learning for data clustering. Cluster Computing, 25(6), (2022) 4009-4025. https://doi.org/10.1007/s10586-022-03633-z

[12] H. Li, Y. Xia Dou, Resource optimization in smart electronic health systems using IoT for heart disease prediction via feedforward neural networks," Cluster Computing, vol. 28, (2025), no. 1, p. 21, https://doi.org/10.1007/s10586-024-04726-7

[13] N.K. Sharma, A.S. Chauhan, S. Fatima, S. Saxena, Enhancing heart disease diagnosis: Leveraging classification and ensemble machine learning techniques in healthcare decision-making. Journal of Integrated Science and Technology, 13(1), (2025) 1016-1016. https://doi.org/10.62110/sciencein.jist.2025.v13.1016

[14] Y. Khourdifi, M. Bahaj, M. Bahaj, Heart disease prediction and classification using machine learning algorithms optimized by particle swarm optimization and ant colony optimization. International Journal of Intelligent Engineering and Systems, 12(1), (2019) 242-252.

[15] S. Mohan, C. Thirumalai, G. Srivastava, Effective heart disease prediction using hybrid machine learning techniques. IEEE access, 7, (2019) 81542-81554. https://doi.org/10.1109/ACCESS.2019.2923707

[16] W. He, Y. Xie, H. Lu, M. Wang, H. Chen, Predicting coronary atherosclerotic heart disease: an extreme learning machine with improved salp swarm algorithm. Symmetry, 12(10), (2020) 1651. https://doi.org/10.3390/sym12101651

[17] M.S. Nawaz, B. Shoaib, M.A. Ashraf, Intelligent Cardiovascular Disease Prediction Empowered with Gradient Descent Optimization. Heliyon, 7(5), (2021) e06948. https://doi.org/10.1016/j.heliyon.2021.e06948

[18] A.F. Subahi, O.I. Khalaf, Y. Alotaibi, R. Natarajan, N. Mahadev, T. Ramesh, Modified Self-Adaptive Bayesian Algorithm for Smart Heart Disease Prediction in IoT System. Sustainability, 14(21), (2022) 14208. https://doi.org/10.3390/su142114208

[19] P. Rani, R. Kumar, N.M.S. Ahmed, A. Jain, A decision support system for heart disease prediction based upon machine learning. Journal of Reliable Intelligent Environments, 7(3), (2021) 263-275. https://doi.org/10.1007/s40860-021-00133-6

[20] P. Yang, D. Wang, W.-B. Zhao, L.H. Fu, J.L. Du, H. Su, Ensemble of kernel extreme learning machine based random forest classifiers for automatic heartbeat classification. Biomedical Signal Processing and Control, 63, (2021) 102138. https://doi.org/10.1016/j.bspc.2020.102138

[21] R.C. Ripan, I.H. Sarker, S.M.M. Hossain, M.M. Anwar, R. Nowrozy, M.M. Hoque, M.H. Furhad, A data-driven heart disease prediction model through K-means clustering-based anomaly detection. SN Computer Science, 2(2), (2021) 112. https://doi.org/10.1007/s42979-021-00518-7

[22] Y. Kaya, F. Kuncan, A hybrid model for classification of medical data set based on factor analysis and extreme learning machine: FA+ ELM. Biomedical Signal Processing and Control,

78 (2022) 104023. https://doi.org/10.1016/j.bspc.2022.104023

[23] R. Siouda, M. Nemissi, H. Seridi, Diverse activation functions based-hybrid RBF-ELM neural network for medical classification. Evolutionary Intelligence, 17, (2024) 829-845. https://doi.org/10.1007/s12065-022-00758-3

[24] S. Sucharita, B. Sahu, T. Swarnkar, S.K. Meher, Classification of cancer microarray data using a two-step feature selection framework with moth-flame optimization and extreme learning machine. Multimedia Tools and Applications, 83(7), (2024) 21319-21346. https://doi.org/10.1007/s11042-023-16353-2

[25] H.A. Al-Alshaikh, P. Prabu, R.C. Poonia, A.K.J. Saudagar, M. Yadav, H.S. AlSagri, A.A. AlSanad, Comprehensive evaluation and performance analysis of machine learning in heart disease prediction. Scientific Reports, 14(1), (2024) 7819. https://doi.org/10.1038/s41598-024-58489-7

[26] S. Zhang, C. Kang, J. Cui, H. Xue, S. Zhao, Y. Chen, H. Lu, L. Ye, D. Wang, F. Chen, Y. Zhao, L. Pei, Development of machine learning-based models to predict congenital heart disease: A matched case-control study. International Journal of Medical Informatics, 195, (2025) 105741. https://doi.org/10.1016/j.ijmedinf.2024.105741

[27] K.M. Jha, V. Velaga, K.K. Routhu, G. Sadaram, S.B. Boppana, Evaluating the Effectiveness of Machine Learning for Heart Disease Prediction in Healthcare Sector. Journal of Cardiobiology, 9(1), (2025) 1.

[28] G.B. Huang, Q.Y. Zhu, C.K. Siew, Extreme learning machine: theory and applications. Neurocomputing, 70(1-3), (2006) 489-501. https://doi.org/10.1016/j.neucom.2005.12.126

[29] B. Niu, H. Wang, Bacterial colony optimization. Discrete Dynamics in Nature and Society, 2012, (2012). https://doi.org/10.1155/2012/698057

[30] K. Vijayakumari, V. Baby Deepa, Fuzzy C-Means Hybrid with Fuzzy Bacterial Colony Optimization. Advances in Electrical and Computer Technologies: Select Proceedings of ICAECT 2020, (2021) 75-87. https://doi.org/10.1007/978-981-15-9019-1_7

[31] H. Wang, L. Tan, B. Niu, Feature selection for classification of microarray gene expression cancers using Bacterial Colony Optimization with multi-dimensional population. Swarm and Evolutionary Computation, 48, (2019) 172-181. https://doi.org/10.1016/j.swevo.2019.04.004

[32] B. Niu, T. Xie, Y. Bi, J. Liu, Bacterial colony optimization for integrated yard truck scheduling and storage allocation problem. International Conference on Intelligent Computing, (2014) 431-437. https://doi.org/10.1007/978-3-319-09330-7_50

[33] B. Niu, Q. Liu, Z. Wang, L. Tan, L. Li, Multi-objective bacterial colony optimization algorithm for integrated container terminal scheduling problem. Natural computing, 20(1), (2021) 89-104. https://doi.org/10.1007/s11047-019-09781-3

[34] S. İlkin, T.H. Gençtürk, F.K. Gülağız, H. Ozcan, M.A. Altuncu, S. Şahin, hybSVM: Bacterial colony optimization algorithm based SVM for malignant melanoma detection. Engineering Science and Technology, an International Journal, 24(5), (2021) 1059-1071. https://doi.org/10.1016/j.jestch.2021.02.002

[35] I.R. White, P. Royston, A.M. Wood, Multiple imputation using chained equations: issues and guidance for practice. Statistics in medicine, 30(4), (2011) 377-399. https://doi.org/10.1002/sim.4067

[36] K. Yan, D. Zhang, Feature selection and analysis on correlated gas sensor data with recursive feature elimination. Sensors and Actuators B: Chemical, 212, (2015) 353-363. https://doi.org/10.1016/j.snb.2015.02.025

[37] A. Sadeghi-Niaraki, P. Mirshafiei, M. Shakeri, S.-M. Choi, Short-term traffic flow prediction using the modified elman recurrent neural network optimized through a genetic algorithm. IEEE Access, 8, (2020) 217526-217540. https://doi.org/10.1109/ACCESS.2020.3039410

[38] F. Han, H.F. Yao, Q.H. Ling, An improved evolutionary extreme learning machine based on particle swarm optimization. Neurocomputing, 116, (2013) 87-93. https://doi.org/10.1016/j.neucom.2011.12.062

[39] J.H. Cho, D.J. Lee, M.G. Chun, Parameter optimization of extreme learning machine using bacterial foraging algorithm. Journal of the Korean Institute of Intelligent Systems, 17(6), (2007) 807-812. https://doi.org/10.5391/JKIIS.2007.17.6.807

[40] P. Vigneshvaran, A.V. Kathiravan, (2022) Heart Disease Prediction using an optimized Extreme Learning Machine with Bacterial Colony optimization. International Conference on Smart Electronics and Communication (ICOSEC), IEEE, India. https://doi.org/10.1109/ICOSEC54921.2022.9952051

[41] J. Kaur, B.S. Khehra, A. Singh, Back propagation artificial neural network for diagnose of the heart disease. Journal of Reliable Intelligent Environments, 9(1), (2024) 57-85. https://doi.org/10.1007/s40860-022-00192-3

[42] A.S. Alencar, A.R.R. Neto, J.P.P. Gomes, A new pruning method for extreme learning machines via genetic algorithms. Applied Soft Computing, 44, (2016) 101-107. https://doi.org/10.1016/j.asoc.2016.03.019

[43] Eshtay M, Faris H, Obeid N. Improving extreme learning machine by competitive swarm optimization and its application for medical diagnosis problems. Expert Systems with Applications. 104 (2018) 134-52. https://doi.org/10.1016/j.joim.2021.08.001

[44] C.Y. Zhang, S.L. Wang, C.M. Yu, Y.X. Xie, C. Fernandez, Improved particle swarm optimization-extreme learning machine modeling strategies for the accurate lithium-ion battery state of health estimation and high-adaptability remaining useful life prediction. Journal of the Electrochemical Society, 169(8), (2022) 080520. https://doi.org/10.1149/1945-7111/ac8a1a

[45] X. Lv, H. Chen, Q. Zhang, X. Li, H. Huang, G. Wang, An improved bacterial-foraging optimization-based machine learning framework for predicting the severity of somatization disorder. Algorithms, 11(2) (2018) 17. https://doi.org/10.3390/a11020017

[46] H. Yang, J. Tian, B. Meng, K. Wang, C. Zheng, Y. Liu, J. Yan, Q. Han, Y. Zhang, Application of extreme learning machine in the survival analysis of chronic heart failure patients with high percentage of censored survival time. Frontiers in cardiovascular medicine, 8, (2021) 726516. https://doi.org/10.3389/fcvm.2021.726516

Authors Contribution Statement

Vigneshvaran and Vijaya Kathiravan was done the data collection, conduct the experimental and write the paper. Finally, both authors were verified and finalized the paper.

Funding

The authors declare that no funds, grants or any other support were received during the preparation of this manuscript.

Competing Interests

The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

Data Availability

The dataset can download from the following website: https://archive.ics.uci.edu/dataset/45/heart+disease

Has this article screened for similarity?

Yes

About the License