

## Research Article

# Moth-Flame Optimization for Early Prediction of Heart Diseases

**S. Haseena,<sup>1</sup> S. Kavi Priya,<sup>2</sup> S. Saroja <sup>1</sup> R. Madavan,<sup>3</sup> M. Muhibbullah <sup>4</sup>,  
and Umashankar Subramaniam<sup>5</sup>**

<sup>1</sup>Department of Information Technology, Mepco Schlenk Engineering College, Sivakasi, India

<sup>2</sup>Department of Computer Science and Engineering, Mepco Schlenk Engineering College, Sivakasi, India

<sup>3</sup>Department of Electrical and Electronics Engineering, P.S.R Engineering College, Sivakasi, India

<sup>4</sup>Department of Electrical and Electronic Engineering, Bangladesh University, Dhaka 1207, Bangladesh

<sup>5</sup>Renewable Energy Laboratory, College of Engineering, Prince Sultan University, Riyadh, Saudi Arabia

Correspondence should be addressed to M. Muhibbullah; [m.muhibbullah@bu.edu.bd](mailto:m.muhibbullah@bu.edu.bd)

Received 15 June 2022; Revised 16 July 2022; Accepted 18 August 2022; Published 12 September 2022

Academic Editor: Xi Lou

Copyright © 2022 S. Haseena et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Heart disease is among the leading causes of mortality globally. Predicting cardiovascular disease is a major difficulty in clinical data analysis. AI has been demonstrated to be powerful in deciding and anticipating an enormous measure of information created by the health domain. We provide a unique method for finding essential traits employing machine learning approaches in this paper, which enhances the effectiveness of identifying heart diseases. Decision tree (DT), support vector machine (SVM), artificial neural network (ANN), and K-nearest neighbor (KNN) are the classification techniques used to create the proposed system. Ensemble stacking integrates the four classification models to create a single best-fit predictive model using logistic regression. Many explorations have been directed at the identification of cardiac infection; however, the exactness of the outcomes is poor. Accordingly, to further enhance the efficiency, Moth-Flame Optimization (MFO) algorithm is proposed. The feature selection strategies are used to improve the classification accuracy while shortening the execution time of the classification system. Medical data are used to assess the probability of heart disease based on BP, age, gender, chest ache, cholesterol, blood sugar, and other variables. Results revealed that the proposed system excelled other existing models, obtaining 99% accuracy in the Cleveland dataset.

## 1. Introduction

Heart disease refers to a group of illnesses that affect people's hearts and veins. The symptoms of cardiac disease differ from person to person [1]. Cardiac disease alludes to a bunch of issues characterized by side effects, for example, hypertension, stroke, respiratory failure, and arrhythmia. The trouble for medical services suppliers is to give top notch care at a reasonable cost. Inadequate outcomes may result from an inaccurate clinical diagnosis and poor therapy. Detecting and diagnosing cardiovascular illness is a never-ending task that can be accomplished by a skilled practitioner with significant experience and understanding [2]. Decision support systems (DSSs) may be used by healthcare organizations to cut costs [3]. Patient records, various disease diagnoses, resource management, and other aspects of healthcare are common.

There has been a lot of work put into establishing remote monitoring gadgets and processes for diagnosing patients. Device exhaustion, then again, has been viewed as a hindrance to adherence [4]. Commercially accessible technologies have been exhibited to conquer this boundary by lessening the necessity for human contact [5], and the precision of action trackers has been shown to be adequate for health workers [6]. Clinical records are continually being produced, processed, and evaluated as a result of the advent of information technology systems. Clinical reports contain data that could be utilized to develop new healthcare services around the world, addressing issues such as social and economic status. Clinical reports, for example, contain a variety of numerical data, medical descriptions, images, etc. All of which can be used to create content-based services to help patients and doctors.

Rapid cardiovascular disease diagnosis of high-risk patients and faster detection using a prediction system has been widely proposed to reduce mortality rates and improve selection for future diagnosis and interventions [7]. A decision support system (DSS) framework may be utilized to assist clinicians in assessing the chance of cardiovascular problems and providing suitable medicines to further prevent the occurrence. Moreover, several studies have indicated that adopting a DSS can enhance preventive services, treatment planning, and better decision-making [8]. An expert DSS based on machine learning (ML) model and metaheuristics approach is used to efficiently identify heart disease.

The ML prescient models need legitimate information for preparing and testing. ML is now being utilized in hospitals to aid in the organization of administrative procedures, the planning and management of infectious diseases, and the customization of medical treatments [9]. Besides, the machine is presently utilized in an assortment of heart-related disciplines, as well as the improvement of new medical operations, the control of patient information and records, and the therapy of persistent illnesses [10, 11].

Supervised method [12] is an ML approach that predicts future data by mapping data flow based on labeled training information. The goal of this approach is aimed at creating a model and then improving the machine's performance as it is exposed to new data. Unsupervised learning technique is an ML method in which the system is given an unlabeled dataset and is required to uncover relationships within the data. Unsupervised learning is aimed at grouping data together and detecting existing patterns. Reinforcement learning algorithm uses trial and error to uncover the pattern or relationship within the data. An agent who makes decisions, the environment in which the agent interacts, and the actions that the agent must complete are the three key components involved. This algorithm's purpose is to discover the optimum policy based on experience to make reliable decisions.

Data mining and machine learning (ML) approaches minimize computer costs and time. One use of ML is the detection of medical disorders and therapies to enhance a patient's quality of life. Heart disease is commonly thought to be a condition that only affects the elderly; nevertheless, it is growing more common in people of all ages.

Information preprocessing is expected for information standardization to further develop ML models' expectation capacities. Preprocessing techniques such as noisy data removal and normalization are carried out. Model performance is also improved by feature extraction and selection strategies. Feature selection is performed using Moth-Flame Optimization technique and feature extraction using principal component analysis. Furthermore, classification is performed using four models, namely, DT, SVM, ANN, and KNN. Ensemble stacking is performed to combine the results from four classification models using logistic regression. The proposed system's result is evaluated using several metrics such as accuracy, precision, recall, and  $F$ -measure.

The main contributions of this research include the following:

- (i) An automated system is designed to classify the dataset using Cleveland dataset
- (ii) Feature selection is performed using Moth-Flame Optimization algorithm that helps in removing unwanted features present in the dataset
- (iii) Feature extraction is carried out using principal component analysis (PCA) that helps in reducing the dimensionality of the dataset further
- (iv) Classification is performed using four different models, and the best results are obtained using the ensemble stacking method
- (v) The system outperforms the existing state-of-art systems concerning accuracy, precision, recall, and  $F$ -measure

The following shows the breakdown of this intended work. Section 2 goes into detail on the work related to the planned task. Section 3 goes into detail about the proposed framework. The fourth section digs into the specifics of experimental outcomes and performance evaluation. Section 5 discusses the conclusion as well as future efforts.

## 2. Literature Works

Researchers have proposed several ML-based diagnostic methodologies for HD. ML algorithms [13, 14] have been extensively utilized in a variety of research such as disease recognition and identification [15]. Gudadhe et al. [16] used multilayer perceptron and SVM to create a detection approach for heart disease (HD) categorization, which achieved an efficiency of 81%. Detrano et al. [17] used ML classification techniques to construct the HD classification system, which had 77 percent accuracy.

Al-Makhadmeh and Tolba [18] introduced a deep belief neural network model-based IoT-based heart disease identification system. The collected data was examined for missing values. The writers looked at how the data was distributed. The authors also employed the studentized approach to use normalized data. Deep belief networks and a high-order Boltzmann machine were used to extract features from noiseless data to be used by the model. The researchers achieved a prediction accuracy of 99.03, which can aid in the reduction of heart disease mortality. Ahmed [19] proposed an HD identification algorithm employing IoT architecture with SVM. To forecast heart disease, the patient data was analyzed using an SVM. The researchers claim that their method predicted cardiac disease with 97.53 percent accuracy. Using an IoT device, they acquired heart data and identified HD. This method detected HD in a short period of time, but the accuracy suffered when a large amount of data was used [20].

Guidi et al. [21] helped to create a DSS for heart failure analysis (HF). They investigated the effectiveness of neural networks (NN), SVM, CART-based fuzzy rules, and random forests, among other ML classifiers (RF). The CART model with random forests produced the best results, with an accuracy of 87.6 percent. Parthiban and Srivatsa [22] explored an SVM approach for detecting HD in diabetic patients with an efficiency of 95%.

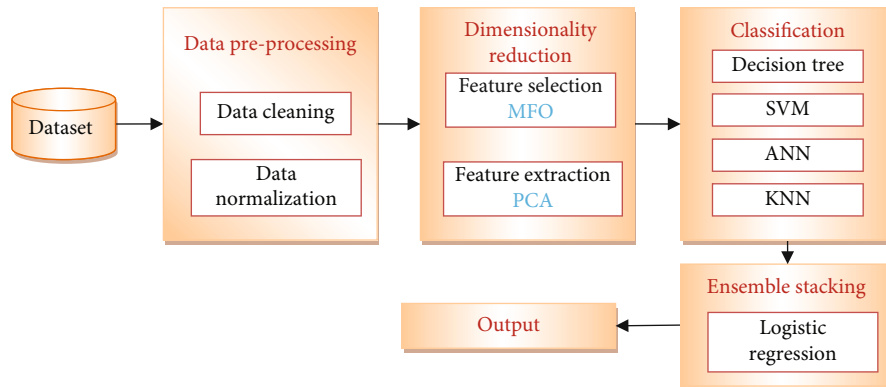


FIGURE 1: Schematic diagram of proposed system.

Melillo et al. [23] led to the formation of an automated classification model that differentiates between those at low and high risk. The sensitivity and specificity of the classification and regression tree (CART) were calculated to be 93.3 percent and 63.5 percent, respectively. To reduce the number of features, Sewak et al. [24] employed a binary classifier that correctly identified heart disease with a 100% improvement in accuracy. Various techniques are available for feature selection [25, 26]. PCA [27, 28] has been employed as a feature retrieval strategy for categorization in healthcare services in recent research.

Classification is performed using Cleveland dataset [29] to identify if a patient had heart disease or not. The ML algorithms combined with feature extraction have three major goals: (i) to discover the unique features, (ii) evaluate the efficiency of PCA, and (iii) investigate the model that produces better results.

### 3. Proposed Methodology

This section describes the proposed system and its components in detail. Figure 1 depicts the proposed system's schematic diagram.

**3.1. Materials.** In this research, the Cleveland dataset is used. There were 303 occurrences and 75 features when this dataset was designed. We preprocessed the dataset in this research, and six samples were excluded due to missing values in this dataset.

**3.2. Data Preprocessing.** Data preparation is a crucial stage in ML that enhances the quality of the data and makes it easier to extract useful knowledge from big data. Data preprocessing is a method of organizing and managing unprocessed data to prepare it for the development and training of ML models. Data cleaning is the process of eliminating erroneous, missing, and inconsistent data from databases as well as restoring missing data. After 6 samples with missing values are discarded, the remaining samples are processed. The absence and the presence of HD are portrayed by a solitary result marked with two classes. After missing values are eliminated, the data should be normalized well within the range from 0 to 1 to make evaluating heart disease trends

easier. The studentized residual methodology is a standard deviation computation-based normalizing method [30]. To standardize the data for HD prediction, several data distributions and regression analysis are employed.

**3.3. Dimensionality Reduction.** The process of lowering the variables analyzed is known as dimensionality reduction [31]. It is employed to retrieve hidden characteristics within the unprocessed without sacrificing the integrity. This work incorporates 2 methods: the Moth-Flame Optimization (MFO) [32] for feature selection and the principal component analysis (PCA) for feature extraction [33].

**3.3.1. Feature Selection by Moth-Flame Optimization (MFO) Algorithm.** MFO is an optimization approach focused on moths' use of transverse orientation to travel at night. Moths may fly greater ranges in a linear fashion by keeping a constant alignment towards the Moon. When moths come into touch with ambient light, they strive to maintain the same angle towards the light source, but the close closeness causes them to become entangled in a spiral route.

The MFO method distributes moths to various alternatives in the optimization problem's solution space, within each fitness function. Every moth carries a flame containing the optimal solution the moth has discovered. The moths fly in a circular route around their flames, changing their locations in each iteration as they explore the solution space. In MFO, moth positions are randomly set inside the solution space. The moths' fitness values are calculated. Each moth's best individual location is identified by the flame. The flame identifies the ideal location for every moth. The moths' locations are improved using a spiral movement algorithm nearing their best individual positions highlighted by a flame in the following iteration, and the flames' locations are replaced with new optimum and most consistent locations. Until the termination requirements are reached, the MFO algorithm keeps updating and generating new positions for the moths and flames.

The following is the computing procedure for MFO:

*Step 1* (creating the preliminary population). Each moth is considered to occupy a  $P$ -dimensional solution space. The

Step 1. Decide on the number of neighbours (K).  
 Step 2. Determine the Euclidean distance between K neighbours.  
 Step 3. Using the obtained Euclidean distance, find the K closest neighbours.  
 Step 4. Count the number of data points in each group among these K neighbours.  
 Step 5. Assign the new data points to the group among with the greatest number neighbours.  
 Step 6. We have completed our model.

ALGORITHM 1: Algorithm of KNN.

TABLE 1: Cleveland dataset used for the proposed system.

Name	Cleveland
Total # of instances	303
Number of attributes	75
Omitted values	Yes
Dataset type	Multivariate
Attribute type	Categorical, integer, real
Tasks performed	Classification

collection of moths might be written as

$$N = \begin{bmatrix} N_{1,1} & \cdots & N_{1,p} \\ N_{2,1} & \cdots & N_{2,p} \\ \vdots & \ddots & \vdots \\ N_{m,1} & \cdots & N_{m,p} \end{bmatrix}, \quad (1)$$

where  $m$  denotes the moth and  $p$  denotes the dimensions. The fitness function is given by

$$F = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{bmatrix}. \quad (2)$$

Two components of the MFO include flames and their respective fitness functions as shown below:

$$FM = \begin{bmatrix} FM_{1,1} & \cdots & FM_{1,p} \\ FM_{2,1} & \cdots & FM_{2,p} \\ \vdots & \ddots & \vdots \\ FM_{m,1} & \cdots & FM_{m,p} \end{bmatrix}, \quad (3)$$

$$OF = \begin{bmatrix} of_1 \\ of_2 \\ \vdots \\ of_m \end{bmatrix}.$$

In MFO, moths seek the optimum solution in each iter-

ation, with flames depicting the optimum solution discovered. Location of the flame is then restructured.

*Step 2* (positions of the moths are being updated). MFO employs three functions to setup the moths' random placements ( $R$ ), optimal solution ( $S$ ), and stop process ( $T$ ):

$$MFO = (R, S, T). \quad (4)$$

To setup the moths' location in the optimal position, any random distribution can be utilized. The  $R$  function's implementation can be written as

$$N(A, B) = (\text{up}(A) - \text{lo}(B)) * \text{rand}() + \text{lo}(A), \quad (5)$$

where  $\text{up}$  and  $\text{lo}$  are arrays that determine the upper limit and lower limit of the function, respectively. The direction of moths in the optimal position is described using a logarithmic function that is susceptible to the conditions listed as follows:

- (i) The initial point of the spiral should be the moth
- (ii) The location of the flame should represent the spiral's ultimate point
- (iii) The spiral's reach should not vary more than the problem space

As a result, the  $S$  represents the movement as shown as follows:

$$S(N_A, FM_B) = D_A \cdot e^{ct} \cdot \cos(2\pi t) + FM_B, \quad (6)$$

where  $c$  is a constant used to establish the logarithmic shape,  $t$  is a random value between 1 and -1, and  $D_A$  is the length.

$$D_A = |FM_B - N_A|. \quad (7)$$

The spiral path of the moth encircling the flame assures that the optimal solution is explored and utilized. Each cycle sorts the ideal solution (flames) to keep the moths from being stuck in local optima, and it hovers surrounding its associated flame using OF and OM matrix.

*Step 3* (flame update). Equation (8) would be used to reduce the number of flames in the MFO algorithm, resulting in the moths solely circling around the optimal solution in the



FIGURE 2: Heat map of Cleveland dataset.

TABLE 2: Instances in the dataset after preprocessing.

Name	Instances present	With HD	Without HD
Cleveland	283	157 (55%)	126 (45%)

method’s final phase.

$$\text{Flame Number} = \text{round} \left( \text{Max} - \text{Cur} * \frac{\text{Max} - \text{Cur}}{\text{Iter}} \right), \quad (8)$$

where Cur stands for the existing number of epochs, Max stands for the highest number of flames, and Iter stands for the number of iterations. The solution space developments are balanced by reducing the number of flames.

*Step 4 (termination condition).* The termination criterion decides when the algorithm should be stopped. The choice of a good termination criterion is critical for the algorithm’s correct convergence. The MFO is frequently terminated based on the number of repetitions, the amount of improved performance, and the length of time it has been running.

*3.3.2. Feature Extraction by Principal Component Analysis (PCA).* We use the eigenvalue-one criterion to identify the number of meaningful components to keep in the analysis. We were able to keep all components with an eigenvalue larger than 1 as a result of this. As an independent variable, each element accounts for one unit of variation. As a result, components having an eigenvalue larger than one accounted

for higher variance compared to individual contributions. Components having eigenvalues less than 1, on the other hand, delivered lesser than the individual and were eliminated.

*3.4. Classification.* The proposed research’s subsequent phase is classification. A class label is forecasted for a particular example of input data in classification, which is a predictive modeling (PM) job in ML. PM design is the method of estimating the feature representation from discrete input parameters to discrete independent variables. The primary purpose is to determine which group the additional knowledge corresponds to. There are four classification algorithms incorporated in the proposed research that includes decision tree, SVM, ANN, and KNN.

*3.4.1. Decision Tree (DT).* DT is a visual representation of decision taking using an algorithm. A DT is created by asking a yes/no question and then breaking the answer into two parts to lead to another decision. The question is at the node, and the decisions that follow are at the leaves. It may be used to tackle classification and regression problems.

*3.4.2. Support Vector Machine (SVM).* SVM is a regulated ML methodology that can tackle classification and regression issues. It may also be utilized to address categorization challenges. Each piece of data is represented as a locus in geometry, without the value of each attribute representing the algorithm’s value of a single locus. The categorization is then accomplished by determining the hyperplane that

TABLE 3: Experimental results of preprocessing technique.

S. no	Method	Precision (%)	Recall (%)	F-measure (%)	Accuracy (%)
1	No preprocessing	80	80	81	79
2	Noise removal	82	80	86	85
3	Normalization using min-max	86	84	80	87
4	Normalization using Z-score	87	85	86	86
5	Combining 2, 3, and 4	90	89	89	91

TABLE 4: Experimental results of feature selection approach.

S. no	Methodology	Precision (%)	Recall (%)	F-measure (%)	Accuracy (%)
1	Genetic algorithm	82	81	76	81
2	Particle swarm optimization	87	86	84	89
3	Moth-Flame Optimization	90	87	87	91

TABLE 5: Experimental results of feature extraction approach.

S. no	Methodology	Precision (%)	Recall (%)	F-measure (%)	Accuracy (%)
1	Linear discriminant analysis (LDA)	81	79	73	80
2	Nonnegative matrix factorization (NMF)	80	83	84	85
3	Principal component analysis (PCA)	90	86	87	90

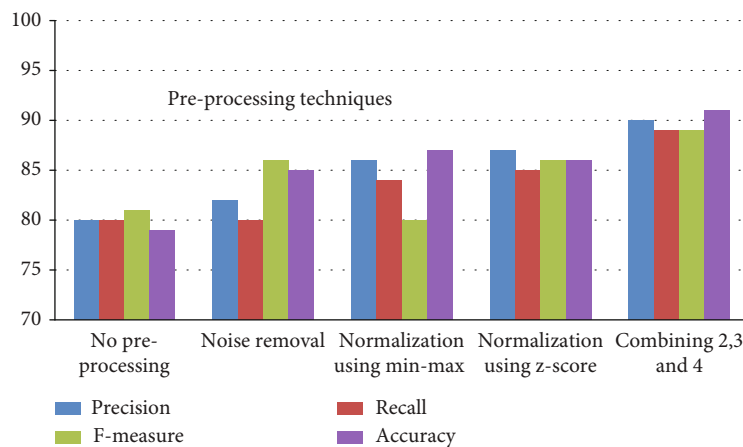


FIGURE 3: Comparison of preprocessing techniques.

TABLE 6: Experimental results of classification approach.

S. no	Methodology	Precision (%)	Recall (%)	F-measure (%)	Accuracy (%)
1	DT	94	95	98	98
2	SVM	96	98	98	99
3	ANN	98	94	97	98
4	KNN	97	97	98	96
5	Ensemble stacking	98	98	99	99

clearly differentiates the 2 groups. An SVM model in high-dimensional space is simply a representation of different groups in a hyperplane. To decrease the error, SVM will cre-

ate the hyperplane progressively. The goal of SVM is to divide the datasets into groups such that a maximal marginal hyperplane may be identified (MMH). Using these support vectors, we optimize the classifier's margin. If the support vectors are removed, the location of the hyperplane will change.

**3.4.3. Artificial Neural Network (ANN).** ANNs are a crude model of how the human brain learns. Neurons, which are in charge of layer generation, make up an ANN. These neurons are also known as tuned factors. The result of each layer is forwarded on to the subsequent layer. Every layer has its own nonlinear activation function, which helps with the process of learning and producing the result. Terminal neurons are another name for the output layer. Each epoch, the

TABLE 7: Evaluation with other existing studies.

S. no	Approaches	Precision (%)	Recall (%)	F-measure (%)	Accuracy (%)
1	Miao et al. [40]	81	71	80	80
2	Naidu and Rajendra [41]	80	85	82	85
3	Shamosollahi et al. [42]	92	90	90	92
4	Hungarian dataset	93	91	95	94
5	Proposed method (Cleveland dataset)	98	98	99	99

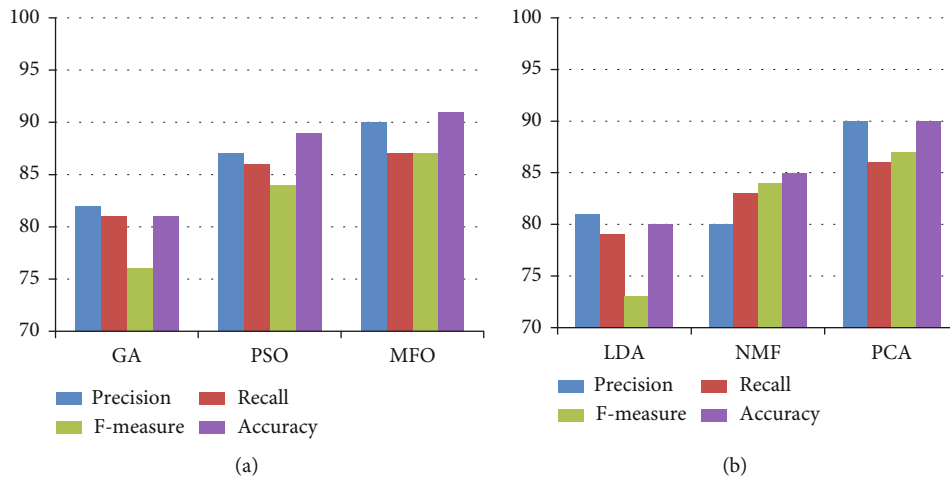


FIGURE 4: Comparison of (a) feature selection techniques and (b) feature extraction technique.

weights connected to the neurons and responsible for overall prediction are modified. Various optimizers are incorporated to increase the learning process. Every ANN has objective function that decreases as learning progresses. The best weights for which the cost function gives the best outcomes are then used.

**3.4.4. K-Nearest Neighbor (KNN).** KNN is a simplistic, convenient solution for dealing with classification and regression tasks. The KNN technique ensures that the newest scenario and existing scenarios are identical and places the new case in the group that is most close to the original group. KNN is discussed below.

### 3.5. Ensemble Stacking

**3.5.1. Stacking.** Ensemble learning is a machine learning terminology in which several models are trained to handle the same problem and then integrated to improve results. The main idea is that by correctly combining weak models, we can get more accurate and/or resilient results. A single method may not be able to deliver the best estimate for a given dataset. ML algorithms have limitations, and it is difficult to create a high-accuracy model. The total precision might be enhanced by developing and combining many models. Integrating the results of each system achieves reduced error while keeping generality. This sort of aggregate may be implemented in a variety of ways. Meta-algorithms are a term used in some textbooks to describe such designs.

Stacking is an ML approach used in ensembles. Using a metalearning strategy, it learns how to merge the predictions from two or more ML algorithms. Stacking has the benefit of integrating the characteristics of many high-performing systems to provide predictions that surpass any single model in composition of a classification or regression task.

**3.5.2. Logistic Regression.** A method for evaluating the probability of a finite result given independent variables is known as logistic regression. A binary outcome is incorporated in the majority common logistic regression methods. To model scenarios with more than 2 distinct outcomes, multinomial logistic regression can be employed. Logistic regression is a powerful statistical method for determining if a fresh data matches well inside a group. Because cyber security components, like threat detection, are categorization issues, logistic regression is a useful analysis technique.

## 4. Results and Discussion

The Cleveland dataset is used to run the proposed architecture. It performs classification using the following steps, namely, preprocessing, feature selection using MFO, feature extraction using PCA, and classification using decision tree, SVM, ANN, and KNN, followed by ensemble staking using logistic regression. The proposed system achieves higher accuracy with reduced computation.

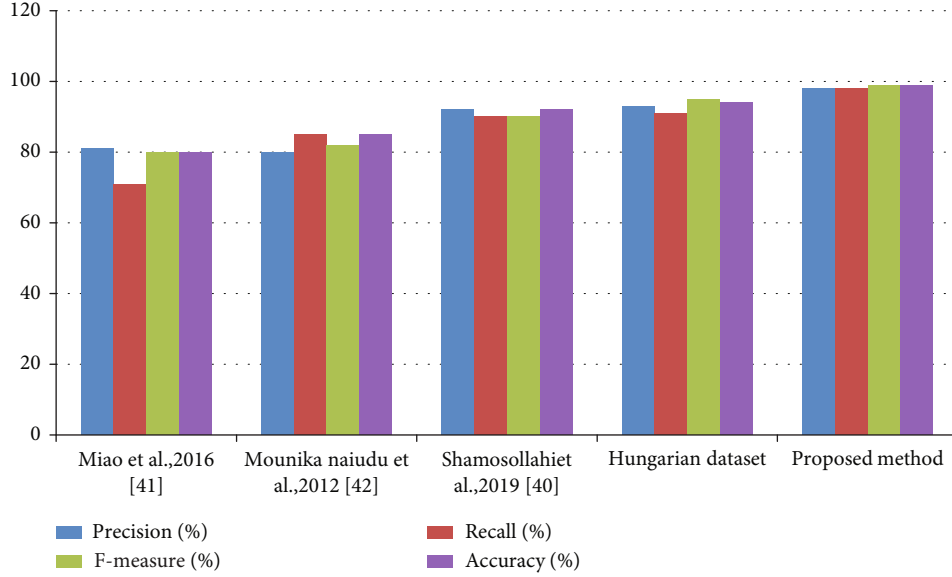


FIGURE 5: Comparison of various state-of-art methods.

Table 1 shows the description of Cleveland dataset. A heat map helps in visualizing the data present in the dataset. Figure 2 represents the heat map for Cleveland dataset. Table 2 represents the instances in the dataset after preprocessing.

For the proposed research, the datasets are divided into two halves: (1) a training part with 70% of the information and (2) a testing part with 30% of the information. Four performance metrics, namely, accuracy, precision, recall, and  $F$ -measure, are considered in the proposed research. The accuracy rate is obtained through dividing the number of accurate classifications by the total number of classes in the dataset. It is represented in

$$\text{Accuracy} = \frac{\text{Num of correct classes identified}}{\text{Total Num of classes}}. \quad (9)$$

Precision, recall, and  $F$ -measure are estimated using the following equations:

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \times 100, \quad (10)$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \times 100, \quad (11)$$

$$\text{F-Measure} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \times 100. \quad (12)$$

Accuracy, precision, recall, and  $F$ -measure should improve over period. Table 3 shows the research findings without preprocessing and with preprocessing technique like noise removal by removing missing values, normalization using min-max, normalization using  $Z$ -score, and combined results. Various metaheuristic algorithms are applied for various applications [34–39]. Table 4 displays the results of experimentation using several feature selection strategies such as the genetic algorithm, particle swarm optimization,

and Moth-Flame Optimization. The proposed method using MFO achieves the best result when compared with other techniques. Table 5 displays the results of the experiment of several feature extraction strategies, with the PCA employed in the proposed method producing the best results in comparison with other approaches. Figure 3 depicts a comparison of preprocessing methods.

Table 6 shows the experimental classification approaches like DT, SVM, ANN, and KNN. The result after applying ensemble stacking using logistic regression produces the best results than applying individual models. When compared to previous strategies, the current technique employing MFO delivers superior results.

Table 7 shows the result of comparison of the proposed method with various existing techniques. Figure 4 shows the comparison with various feature selection and feature extraction techniques. Figure 5 shows the comparison with several state-of-art methodologies. The recommended strategies produce superior performance in the aspects of accuracy, precision, recall, and  $F$ -measure.

## 5. Conclusions and Future Work

Many lives are saved to healthcare monitoring and prediction technologies, especially when patients are placed distant. The classifier's objective was to determine whether a patient had cardiovascular disease or not. When system resources must be considered, using all functionalities is not possible. In this work, we utilized dimensionality reduction methods to enhance the original collected data. Classification was performed using decision trees, SVM, ANN, and KNN, and ensemble stacking called logistic regression helped in achieving the best results from the various classification models. Our technology may be used to evaluate large volumes of data and discover the risk variables associated with various diseases in a variety of real-world applications or in other medical diagnoses. We intend to test our strategy



on a larger dataset and analyze a different disease using alternative feature selection strategies. Our key limitation is that the limited sample size makes it difficult to generalize these findings to heart disease. In further work, we intend to expand the dataset to which our method will be used and analyze more diseases using various feature selection methods.

### Data Availability

The authors confirm that the data supporting the proposed work are taken from publically available datasets.

### Conflicts of Interest

The authors ensure that there is no conflict of interest in this proposed work.

### References

- [1] M. Kavousi, S. Elias-Smale, J. H. W. Rutten et al., "Evaluation of newer risk markers for coronary heart disease risk classification," *Annals of Internal Medicine*, vol. 156, no. 6, pp. 438–444, 2012.
- [2] H. Tada, O. Melander, J. Z. Louie et al., "Risk prediction by genetic risk scores for coronary heart disease is independent of self-reported family history," *European Heart Journal*, vol. 37, no. 6, pp. 561–567, 2016.
- [3] G. J. Njie, K. K. Proia, A. B. Thota et al., "Clinical decision support systems and prevention," *American Journal of Preventive Medicine*, vol. 49, no. 5, pp. 784–795, 2015.
- [4] M. K. Ong, P. S. Romano, S. Edgington et al., "Effectiveness of remote patient monitoring after discharge of hospitalized patients with heart failure," *JAMA Internal Medicine*, vol. 176, no. 3, pp. 310–318, 2016.
- [5] W. Speier, E. Dzubur, M. Zide et al., "Evaluating utility and compliance in a patient-based eHealth study using continuous-time heart rate and activity trackers," *Journal of the American Medical Informatics Association*, vol. 25, no. 10, pp. 1386–1391, 2018.
- [6] T. Ferguson, A. V. Rowlands, T. Olds, and C. Maher, "The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: a cross-sectional study," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 12, no. 1, pp. 1–9, 2015.
- [7] P. Greenland, J. S. Alpert, G. A. Beller et al., "2010 ACCF/AHA guideline for assessment of cardiovascular risk in asymptomatic adults: a report of the American College of Cardiology Foundation/American Heart Association task force on practice guidelines," *Circulation*, vol. 122, no. 25, pp. e584–e636, 2010.
- [8] J. Hunt, J. Siemieniczuk, W. Gillanders et al., "The impact of a physician-directed health information technology system on diabetes outcomes in primary care: a pre and post-implementation study," *Journal of Innovation in Health Informatics*, vol. 17, no. 3, pp. 165–174, 2009.
- [9] S. I. Ansarullah and P. Kumar, "A systematic literature review on cardiovascular disorder identification using knowledge mining and machine learning method," *International Journal of Recent Technology and Engineering*, vol. 7, no. 6S, pp. 1009–1015, 2019.
- [10] K. Arunachalam and K. S. Chandran, "Effects of heartfulness meditation practices on control of Alzheimer's disease: a comprehensive review," *Iranian Journal of Public Health*, vol. 50, no. 12, pp. 2427–2438, 2021.
- [11] R. Divya and R. Shantha Selva Kumari, "Genetic algorithm with logistic regression feature selection for Alzheimer's disease classification," *Neural Computing and Applications*, vol. 33, no. 14, pp. 8435–8444, 2021.
- [12] S. Haseena, M. B. B. Pepsi, and S. Saroja, "Multi criteria decision making technique for machine learning algorithms: iterative and noniterative algorithms," *International Journal of Scientific & Technology Research*, vol. 9, no. 1, 2020.
- [13] M. B. B. Pepsi and S. N. Kumar, "Supervised learning techniques for classification of students' tweets," *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 12, no. 12, pp. 1714–1722, 2021.
- [14] M. S. Bhuvanewari and K. Muneeswaran, "A parallel approach for web session identification to make recommendations efficient," *International Journal of Business Intelligence and Data Mining*, vol. 19, no. 2, pp. 189–213, 2021.
- [15] R. C. Deo, "Machine learning in medicine," *Circulation*, vol. 132, no. 20, pp. 1920–1930, 2015.
- [16] M. Gudadhe, K. Wankhade, and S. Dongre, "Decision support system for heart disease based on support vector machine and artificial neural network," in *International Conference on Computer and Communication Technology (ICCCCT)*, pp. 741–745, Allahabad, India, September 2010.
- [17] R. Detrano, A. Janosi, W. Steinbrunn et al., "International application of a new probability algorithm for the diagnosis of coronary artery disease," *The American Journal of Cardiology*, vol. 64, no. 5, pp. 304–310, 1989.
- [18] Z. Al-Makhadmeh and A. Tolba, "Utilizing IoT wearable medical device for heart disease prediction using higher order Boltzmann model: a classification approach," *Measurement*, vol. 147, p. 106815, 2019.
- [19] F. Ahmed, "An Internet of Things (IoT) application for predicting the quantity of future heart attack patients," *International Journal of Computers and Applications*, vol. 164, no. 6, pp. 36–40, 2017.
- [20] S. Shen, S. X. Han, P. Petousis et al., "A Bayesian model for estimating multi-state disease progression," *Computers in Biology and Medicine*, vol. 81, pp. 111–120, 2017.
- [21] G. Guidi, M. C. Pettenati, P. Melillo, and E. Iadanza, "A machine learning system to improve heart failure patient assistance," *IEEE journal of biomedical and health informatics*, vol. 18, no. 6, pp. 1750–1756, 2014.
- [22] G. Parthiban and S. K. Srivatsa, "Applying machine learning methods in diagnosing heart disease for diabetic patients," *International Journal of Applied Information Systems*, vol. 3, no. 7, pp. 25–30, 2012.
- [23] P. Melillo, N. D. Luca, M. Bracale, and L. Pecchia, "Classification tree for risk assessment in patients suffering from congestive heart failure via long-term heart rate variability," *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 3, pp. 727–733, 2013.
- [24] R. S. Singh, B. S. Saini, and R. K. Sunkaria, "Detection of coronary artery disease by reduced features and extreme learning machine," *Clujul Medical*, vol. 91, no. 2, pp. 166–175, 2018.
- [25] T. Karpagalingam and M. Karuppaiah, "Optimal feature subset selection based on combining document frequency and

- term frequency for text classification,” *Computing and Informatics*, vol. 39, no. 5, pp. 881–906, 2020.
- [26] S. Subbaraj, R. Thiagarajan, and M. Rengaraj, “Multi-objective league championship algorithm for real-time task scheduling,” *Neural Computing and Applications*, vol. 32, no. 9, pp. 5093–5104, 2020.
- [27] R. Rajagopal and V. Ranganathan, “Evaluation of effect of unsupervised dimensionality reduction techniques on automated arrhythmia classification,” *Biomedical Signal Processing and Control*, vol. 34, pp. 1–8, 2017.
- [28] S. Negi, Y. Kumar, and V. M. Mishra, “Feature extraction and classification for EMG signals using linear discriminant analysis,” in *2016 2nd International Conference on Advances in Computing, Communication, & Automation (ICACCA)(Fall)*, Bareilly, India, October 2016.
- [29] “UCI heart disease data set,” <http://archive.ics.uci.edu/ml/datasets/heartdisease/>.
- [30] S. K. Sarkar, H. Midi, and S. Rana, “Detection of outliers and influential observations in binary logistic regression: an empirical study,” *Journal of Applied Sciences*, vol. 11, no. 1, pp. 26–35, 2010.
- [31] P. Domingos, “A few useful things to know about machine learning,” *Communications of the ACM*, vol. 55, no. 10, pp. 78–87, 2012.
- [32] O. Bozorg-Haddad, *Advanced Optimization by Nature-Inspired Algorithms*, vol. 720, Springer, Singapore, 2018.
- [33] X. Dai and Y. Wei, “Application of improved moth-flame optimization algorithm for robot path planning,” *IEEE Access*, vol. 9, pp. 105914–105925, 2021.
- [34] H. Sikkandar and R. Thiyagarajan, “Deep learning based facial expression recognition using improved cat swarm optimization,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 2, pp. 3037–3053, 2021.
- [35] H. Sikkandar and R. Thiyagarajan, “Soft biometrics-based face image retrieval using improved grey wolf optimisation,” *IET Image Processing*, vol. 14, no. 3, pp. 451–461, 2020.
- [36] S. Haseena, S. Manoruthra, P. Hemalatha, and V. Akshaya, “Mining frequent item sets on large scale temporal data,” in *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, pp. 811–815, Coimbatore, India, March 2018.
- [37] S. Haseena, S. Bharathi, I. Padmapriya, and R. Lekhaa, “Deep learning based approach for gender classification,” in *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, pp. 1396–1399, Coimbatore, India, March 2018.
- [38] S. Subbaraj, R. Thiyagarajan, and M. Rengaraj, “A smart fog computing based real-time secure resource allocation and scheduling strategy using multi-objective crow search algorithm,” *Journal of Ambient Intelligence and Humanized Computing*, 2021.
- [39] S. Haseena, S. Saroja, and T. Revathi, “A fuzzy approach for multi criteria decision making in diet plan ranking system using cuckoo optimization,” *Neural Computing and Applications*, vol. 34, no. 16, pp. 13625–13638, 2022.
- [40] H. Miao Kathleen, H. Miao Julia, and G. J. Miao, “Diagnosing coronary heart disease using ensemble machine learning,” *International Journal of Advanced Computer Science and Applications*, vol. 7, no. 10, 2016.
- [41] M. Naidu and C. Rajendra, “Detection of health care using datamining concepts through web,” *International Journal of Advanced Research in Computer Engineering and Technology*, vol. 1, no. 4, pp. 45–50, 2012.
- [42] M. Shamosollahi, A. Badiie, and M. Ghazanfari, “Using combined descriptive and predictive methods of data mining for coronary artery disease prediction: a case study approach,” *Journal of AI and Data Mining*, vol. 7, no. 1, pp. 47–58, 2019.