

IMGWO BASED ANN: A NEW HEART DISEASE DIAGNOSIS MODEL TO CLASSIFY REAL WORLD DATASET

Narender Kumar

Department of Computer Science and Engineering,
Guru Jambheshwar University of Science and Technology,
Hisar, Haryana Pin-125001, India
narenderster@gmail.com

Dharmender Kumar

Department of Computer Science and Engineering,
Guru Jambheshwar University of Science and Technology,
Hisar, Haryana Pin-125001, India
dharminia24@gmail.com

Abstract

Heart diseases cause most of the untimely and sudden deaths all over the world. Many of the lives can be saved by an early alarm with the means of expert diagnosis systems. Researchers developed such expert systems over time. However, there is still a need for more efficient methods for accurate and efficient diagnosis of heart disease. The accuracy and efficiency of the diagnosis system are highly dependent on the characteristics of data, feature selection (FS) algorithms and classification techniques. The artificial neural network (ANN) exhibits excellent performance on unseen data samples and metaheuristic optimization techniques have performed exceptionally well in training the ANN model and avoiding local minima problems. In this article, the authors have employed an artificial neural network (ANN), trained with a Grey Wolf Optimizer (GWO) and Particle Swarm Optimization (PSO) empowered, Inertia Motivated GWO (IMGWO) technique, to classify a real-world dataset. Three most popular feature selection techniques viz. Relief, mRMR and LASSO are employed to find out the most discriminative features from the collected dataset. After that, the selected features are utilized in the IMGWO trained ANN model for classification. The classification performance of the said technique is then compared with the results obtained from five most prevalent metaheuristic methods viz. Genetic Algorithm (GA), Firefly Algorithm (FF), Whale Optimization Algorithm (WOA), PSO and GWO used to train the ANN model. It is observed from the classification results that the performance of the IMGWO technique utilizing the mRMR feature selection algorithm supersedes all other techniques used. The results have been compared in terms of performance metrics that include accuracy, sensitivity, specificity, the area under the curve (AUC), F-measure, precision, kappa statistics and Mathew's correlation coefficients. The work will be a contribution towards the development of expert systems for heart disease diagnosis with an improved efficacy

Keywords: ANN, Heart diseases, Feature selection, Metaheuristic, GWO, IMGWO.

1. Introduction

Annually 17.9 million people die because of heart disease worldwide and it is the number 1 cause of deaths over all other diseases[1]. Cardiovascular disease (CVD) is another term used for heart disease. When the blood supply to any part of the heart is restricted by means of obstruction (atherosclerotic plaque), chest pain and other symptoms occur. Timely detection of such signs can save many lives. Machine learning (ML) methods are in use for a long time for the detection of CVDs.

Heart disease is detected by either an invasive modality where a catheter is inserted into the body to see the obstruction in an artery or by developing a machine learning model from the data set generated by non-invasive (NI) modalities. The NI modalities include detection of heart ailments by electrocardiography (ECG), echocardiography (ECHO), phonocardiography (PCG) or cardiovascular imaging. Many datasets are publically available for developing machine learning (ML) models that use different NI modalities. UCI ML repository is available publically that contains many datasets for developing and testing ML algorithms. It includes SPECT, Statlog(Heart), Heart Diseases(Cleveland, Hungary, Switzerland and VA long Beach) and Z-Alizadeh Sani [2] datasets for researchers and practitioners of ML who are interested in developing an expert system for the diagnosis of heart diseases. Classification is a supervised learning method which widely used for heart disease diagnosis.

A generic ML model for classification may be defined by Eq.(1).

$$P \xrightarrow{C} \min[Err\{C - Q[L(\sum_{j=1}^n \sum_{i=1}^m D(A_i, S_j))]\}] \quad (1)$$

Where L is an ML algorithm, Q is the quantity of accurate prediction that is compared with actual output C. This difference of the comparison, termed as an error (Err), is minimized during all learning samples of the dataset $D(A_i, S_j)$. The Err is minimized in w.r.t. performance measure P, which may be Accuracy, Sensitivity, Specificity, F-Score or Area under the curve[3]. In Eq.(1), the dataset $D(A, S)$ contains 'm' number of attributes (A) and 'n' number of samples(S). Dataset attributes are also known as features in ML literature.

Decision tree(DT)[4], [5], support vector machine (SVM)[6], k-nearest neighbor (K-NN)[7], artificial neural network (ANN)[8]–[11] are some popular classifiers which are used to develop expert systems for the diagnosis of heart problems. In a study, SVM, DT and ensemble machine learning methods were trained and tested using 10-fold cross-validation on the Heart Disease dataset (Hungary) on two different platforms for a comparative study of heart disease detection. [12]. Ayyapan and SivaKumar compared various correlation coefficient accuracies by applying different parameter pruning methods on the Cleveland heart disease dataset in 2018[13]. Many of the studies performed classification with their own dataset for coronary heart disease (CHD) diagnosis. For example, in 2007, Xing et al. collected 1000 cases in 6 months follow up; after that, three popular ML techniques were compared on the dataset. The SVM performed best with 92.1 % accuracy, ANN performed second best and DT performed worst with 89.6% accuracy[14]. ANN was found to be superseded in a comparative study with collected data of 1069 subjects of CHD in terms of accuracy, sensitivity, and specificity [15]. To diagnose CAD, rules extracted from DT were transformed and optimized with the application of a fuzzy system for 199 subjects in a study by [16].

ANNs performed better as compared to other classifiers while diagnosing a heart problem[17], [18] however, ANN often entrapped in local minima and poor convergence[19], [20]. The poor convergence and local minima entrapment may be handled by hybridizing ANN with suitable metaheuristic optimization methods[21], [22]. In this paper, classification via ANN-IMGWO model is performed preceded by feature selection on a real-world heart diseases dataset. Other contemporary metaheuristic optimization methods viz. Genetic Algorithm (GA)[23], FireFly Algorithm (FF)[24], [25], Whale Optimization Algorithm (WOA)[26], Particle swarm optimization (PSO)[27] and Grey Wolf Optimizer (GWO[28]) are also used as training algorithm for ANN in for the comparative analysis.

The key contribution of this paper is described in detail as follows:

- (1) A real-world heart disease dataset is collected from people of southern Haryana for developing an automated heart disease diagnosis model. Along with the necessary medical reports, this dataset includes some of the vital lifestyle information like alcohol consumption (AC), exercise, sitting hours (SH), air pollution (AP) exposure, and daily sleep hours (DSH). In total, our dataset includes demographic, lifestyle information, symptoms and examinations along with ECG information. All this information is required to detect a chance of heart disease.
- (2) Preprocessing and feature selection are performed before the classification. The three states of the art methods for feature selection viz. Relief, mRMR and LASSO are employed to attain the most discriminative features, thereby avoiding the complexity of developing the classification model. In each FS algorithm, the most discriminative features are selected based on good performances during classification. Missing values are handled by arithmetic means in numerical attributes and mode by categorical values. The normalization is performed with min-max method.
- (3) A new framework for heart disease diagnosis, i.e., ANN-IMGWO, is developed with the most discriminative features selected by FS algorithms. The ANN-IMGWO model's performance is then compared to other contemporary prevalent metaheuristic trained ANN Model in terms of performance metrics like accuracy, sensitivity, specificity, precision, area under the curve, kappa statistics, and Mathew's correlation coefficient (MCC). These other metaheuristic methods are GA, PSO, FF, WOA and GWO.

The paper's first and third contributions are the most appealing and demanding issues in the contemporary era of automated medical diagnosis. FS(the second contribution), is a naturally demanded issue while dealing with a new dataset with many dimensions. IMGWO has been proposed recently by the authors in 2021[22]. IMGWO improved the exploration and exploitation capabilities of GWO by combing the memory retention property of PSO and avoiding local optima and slow convergence of classic GWO. Therefore, it is better to train ANN with IMGWO to avoid local minima entrapment.

The remaining part of the papers is as follows: Section 2 describes the literature related to heart disease diagnosis, including classification through neural networks. Some studies showing improvement in ANN using metaheuristic computations are also elaborated in the literature review section. Section 3 discussed feature selection methods used in this paper. The methodology and the dataset are described in the 4th section. Experimental setup and results are discussed in the 5th section. Finally, the article got concluded in the last section.

2. Literature Review

Machine learning methodologies are being used for heart disease diagnosis for the last three decades, especially after introducing the Cleveland heart disease dataset presented by Detrano et al. in 1989[29]. In 1992, Long et al. developed a program to diagnose heart failure using a probability network and heuristic hypothesis generator with 242 cases of cardiac disease. They attained a 90% accuracy [30] as a result of formative analysis of the diagnostic capability of the program. Clinical input variables are used to train an ANN by an early study in 1994 by William G. Baxt[31] to detect myocardial infarction. A pioneer diagnosis with ANN was performed in 1999 by Azuaje and companions to detect a unique heart rate variability pattern that could cause coronary disease[32]. Heart rate variability (HRV) is a significant feature that helps to detect coronary disease. In 2000, Tkacz and Kostka presented an elegant application of a wavelet neural network to detect a heart abnormality using classification HRV signals. The best classification accuracy was achieved with a double layer ANN in which tangsoidal and linear activation functions were used. The process of classification proceeded with feature extraction with wavelet transform[33].

Many researchers focused on optimizing network structure and weights to improve the classification performance of an ANN in a real-time scenario. In a comparative study of heart disease diagnosis, Gudadhe et al. shown that ANN attained an accuracy of 97.5 % and performed better than SVM (80.41%) on the Cleveland heart disease dataset[34]. In 2009, Resul Das et al. presented an ensemble of neural networks by combining posterior probabilities to detect heart disease with an accuracy of 89.01%, sensitivity of 80.95% and specificity of 95.91%[35]. Ari and Saha optimized the structure and weights of a neural network to classify heart sound signals and attained an accuracy of 99.279%. The numbers of nodes were optimized with singular value decomposition; the number of features selected by QR factorization with column pivoting (QRcp) and Fisher's F-ratio; weight optimization was performed with local relative sensitivity index[36].

Metaheuristic computations, which are usually nature inspired and swarm intelligence[37] based optimization techniques, play a vital role in many real-life applications, especially in medical diagnosis. Recently, particle swarm optimization (PSO) and firefly algorithm (FF) has been used to train an ANN to classify UCI datasets that include the heart disease dataset. More than 20 % enhancement in accuracies has been reported empirically when the two metaheuristic computations were clubbed with neural ANN[38]. Another recent study in 2017 used a hybrid metaheuristic that clubbed PSO and FF to train an ANN for detecting myocardial infarction. The study achieved an accuracy of 99.3%, sensitivity of 99.97%, and specificity of 98.7% on the MIT-BIH database[39]. A hybrid model to detect heart disease is proposed by Priynaga et al. in 2020 that optimizes recurrent neural network with whale optimization algorithm (WOA). A hybrid WOA is used in the training of an ANN by Zhou et al. [40]. Grey wolf optimization, a famous swarm optimization algorithm also used to train an ANN by Mirjalili in [28] and its hybrid version was used to train an ANN Agrawal et al. in[41]. Both the studies used the heart disease dataset for training and testing purposes.

Arabasadi et al. designed a diagnosis model with ANN and genetic algorithm for detecting heart disease in 2017. The ANN model was optimized with GA and feature selection was performed with Gini index, information gain, SVM and PCA. The model attained an average accuracy of 93.85%, sensitivity of 97 % and specificity of 93 % on the Alizadehsani heart disease dataset. The model achieved notable differences in the performance metrics compared to the classic neural network model [42]. Cherian et al. proposed a model for heart disease prediction in 2020 that used ANN being optimized with hybridization of PSO and lion algorithm. Before building the model, PCA was used to minimize the number of dimensions. The model achieved improvements in accuracy as 3.85%, 12.5%, 12.5%, 3.85% and 9.41% over the ANN- Levenberg–Marquardt, ANN-WOA, ANN-FF, ANN-PSO and ANN-Lion algorithm.

The primary issue in heart disease diagnosis is the complexity involved in designing an expert system involving many dimensions. Other issues are low performance regarding accuracy, recall, specificity, precision, F-measures, area under the curve and some essential statistics like kappa and MCC. For the precise and efficient diagnosis system for heart disease, it is very much required that the highlighted issues must be addressed. The work in the paper put a massive endeavor to address the issues with a real-world dataset and increased efficacy.

3. Feature Selection

Understanding the quality of data is essential before applying any ML technique to find patterns out of it. Oreski et al.(2016) studied exhaustively and found that there exists a correlation between the dataset characteristics and performance of feature selection and classification[43]. Characteristics or features are usually examined and handled in data preprocessing, consuming about 50-70 % of the computational effort during any data mining task[44].

In CVD detection, the dataset characteristics have a significant impact on the performance of classification. Researchers are making efforts on data preprocessing continuously to improve the diagnosis accuracy for heart diseases[45]. Data reduction is a major task that is performed in data preprocessing. Data reduction can further be

performed in two ways: dimensional reduction and attribute selection or feature selection (FS). Dimensional reduction often requires transformational compression techniques like Principal Component Analysis(PCA) and Discrete Wavelet Transformation(DWT), while FS involves direct consideration of the attribute without transformation or compression [46].

Diagnosis via classification without considering the significance of the attributes often results in poor efficiency and accuracy. It happens because irrelevant attributes make the classification process a complicated one, and as a result performance, the diagnosis model degrades. FS accelerates the process of building the classification model. A rich literature of FS is available for heart disease diagnosis. Relief is a weight-based domain independent feature selection method used in many recent studies, including[47]. mRMR is a kind of feature selection that handles redundancy and correlation very efficiently. It has been used very recently in medical data, especially in heart disease diagnosis [48], [49]. Another approach that has been considered for feature selection is LASSO[50], which maintains the model parameters by putting a threshold. Hence, regularization and feature selection are performed simultaneously. Relief, MRMR and LASSO are considered feature selection algorithm in this study.

3.1 Relief

Relief is a weight-based feature selection method where weights of the features are noted and updated. Weights are updated in nearest neighbor fashion. The Equation for the weight adjustment is as follows:

$$Weight(w_i) = \frac{1}{2} \sum_{i=1}^l d(X(j, i) - X(NM(j), i)) - d(X(j, i) - X(NH(j), i)) \quad (2)$$

In Eq.(2), w_i is the feature score of an attribute. Data samples l are randomly selected among all n instances, $NM(j)$ is the nearest instance to $X(j)$ with the same class label, $NH(j)$ nearest instance with a different class, $d(.)$ is the distance function for binary classification problem[47]. The feature with greater weights will be selected. It is a kind of supervised filter-based algorithm.

3.2 Minimal-redundancy-maximal-relevance (mRMR)

mRMR is based on the concept of minimum redundancy and maximum relevance theory, which is further based on a seminal concept called *mutual information (MI)* between two attributes. In the theory of feature selection, the *MI* is calculated between the feature under consideration and the class variable. If x and y are two random variables (representing features), *MI* can be expressed in probability distribution functions $p(x), p(y)$ and $p(x, y)$ by the Eq.(3) and (4).

$$I(x, y) = \iint \frac{p(x, y) \log(p(x, y))}{p(x)p(y)} dx dy \quad (3)$$

$$I(x, y) = \sum_{x \in X} \sum_{y \in Y} \frac{p(x, y) \log(p(x, y))}{p(x)p(y)} \quad (4)$$

Eq.(3) represents the *MI* for continuous variables, while Eq.(4) represents the *MI* for discrete-valued variables[51]. If we further assume a feature set S with m number of features $\{x_i\}$ that are jointly dependent on the class label c , then maximal dependency may be expressed as per Eq.(5).

$$\max(dep(S, c)), dep = I(\{x_i, i=1, 2, \dots, m\}; c) \quad (5)$$

Finding maximal dependency is an incremental search and one has to find the feature that has contributed the largest increase for $I(S, c)$. It is a very hard implementation. An alternative solution is to select the feature based on maximal relevance, expressed as the Eq.(6).

$$\max(dep(S, c)), dep = \frac{1}{|S|} \sum_{x_i \in S} I(x_i; c) \quad (6)$$

Mutually dependent features generate redundancy if they exist together in a feature set. Therefore, minimal redundancy constraints must be added to select mutually exclusive features. Minimal redundancy is added as per Eq.(7).

$$\min(rel(S)), rel = \frac{1}{|S|^2} \sum_{x_i, x_j \in S} I(x_i, x_j) \quad (7)$$

Finally minimal-redundancy-maximal-relevance(mRMR) is obtained through combination of both the Equations (Eq.(8)) i.e. Eq.(6) and (7)[52].

$$\text{Max } RR(dep, rel), RR = dep - rel \quad (8)$$

3.3 LASSO

Robert Tibshirani introduced the least absolute shrinkage and selection operator (LASSO) to estimate linear models[53]. There are certain absolute coefficient values associated with features. LASSO puts a limit on the sum of the values. It is achieved by a shrinking or regularization process. The shrinking process may result in such

conditions where we need to put some coefficients value as zero. The leftover coefficients, which are non-zero values, are the selected features. LASSO is assumed to be a good combination of ridge regression and subset selection. Among highly correlated features, LASSO selects one and puts the other to zero. Setting a zero value to a variable is known as shrinking. Putting zero values to most of the variables makes the interpretability of the process an easier one.

4. Dataset and Methodology

4.1 Dataset Collection

The data set was collected from people of southern Haryana during their visits to diagnostic centers. The patients are gone under the examination of the physician, radiologist, and then cardiologist. A detailed description of the collected data is given in section 5. The dataset consists of 500 samples and 42 attributes collected from subjects of different ages and gender. The dataset further consists of demographics, lifestyle, symptoms, lab tests, and ECG information. In this dataset, lifestyle information is added along with symptoms, lab tests, and ECG information since it contributes a lot towards heart disease. In the experimental section, lifestyle information like AC, exercise and smoker, and Ex-smoker are some features that are proved to be significant.

4.2 Methodology

Metaheuristic optimization techniques have optimized an ANN in multiple ways for more than two decades[21]. Optimizing network structure, weights and biases, learning rate and activation function through metaheuristic methods are popular. In this study, the weights and biases are optimized with a variant of grey wolf optimization (GWO)[54]. This variant is named inertia motivated grey wolf optimization (IMGWO) and proposed by Kumar et al. in 2021[22]. IMGWO was proposed to increase the exploration and exploitation capabilities of GWO since it is usually entrapped in local minima and slow convergence during the later part of evolution[55]. IMGWO is explained in Algorithm1.

Algorithm1: Inertia Motivated Grey Wolf Optimization

Variable declaration:

total_pop: total number of solutions

t=0; variable *t* denoting current iteration

max_itr : maximum iteration

A and *C* are some coefficients in the original GWO that managed exploration and exploitation ($(A=2\vec{a}r_1-\vec{a})$ and $(C=2r_2)$)

\vec{a} is a control variable and r_1, r_2 are some random numbers ranging between 0 to 1.

\vec{a} is modified as Eq.(9):

$$\vec{a}(t) = \vec{a}_{ini} - (\vec{a}_{ini} - \vec{a}_{fin}) \times \left(\left(\frac{1}{e-1} \right) \times \left(e^{\frac{t}{max_itr}} - 1 \right) \right) \quad (9)$$

Resultant positions *X* of the alpha search agents is updated as per Eq.(10):

$$X(t+1) = w \cdot \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} + c_1 \cdot r_3 \cdot (\vec{X}_{pbest} - \vec{X}) + c_2 \cdot r_4 \cdot (\vec{X}_1 - \vec{X}) \quad (10)$$

Control variable \vec{a} changes its value from \vec{a}_{ini} to \vec{a}_{fin} in a non-linear way during all iteration. Positions \vec{X}_1, \vec{X}_2 and \vec{X}_3 are individual leading solutions as per classic GWO and r_3, r_4 are random numbers in the range 0 to 1. Inertia factor *w* is a function of time, and it is also a component of particle swarm optimization (PSO) [56]. Coefficients c_1 and c_2 are individual memory components and personal communication components, respectively, with the original PSO range [0,1][55]. \vec{X}_{pbest} is the historical personnel best solution.

Input: *total_pop*, *max_itr*

Calculate the fitness of all of the search agents considering MSE of ANN as a fitness function.

Alpha (α) is the best solution, while beta (β) and delta (δ) are the second and third best solutions, respectively.

Rests of the solutions are omega (ω).

while $t < max_itr$

for $i = 1$ to $total_pop$

 Calculate and update the position of the search agent as per Eq.(10)

end for

 Update *A*, *C* and \vec{a}

 Recalculate fitness

 Update positions of α, β , and δ search agents

end while

Output: Return the position of α as an optimal solution

The position of the alpha search agent is the optimal solution. The details of the IMGWO algorithm is given in [22]. The current research study investigated the performance of IMGWO in terms of better convergence by avoiding the local minima entrapment in a large search space of weights and biases for an ANN. The ANN-IMGWO model was built and evaluated with a 10-fold cross-validation methodology, as shown in Fig. 1. The fitness function for IMGWO is the average mean square error (AVG_{MSE}) which is to be minimized over the course of iterations. The expression for the vector representing weights and biases is given in Eq.(11), and that is for AVG_{MSE} is given by Eq.(12).

$$\vec{V}_{wt-b} = \{\vec{wt} | wt: weights, \vec{b} | b: biases\} \quad (11)$$

$$AVG_{MSE} = \sum_{p=1}^q \frac{\sum_{i=1}^m (ob_i^p - exp_i^p)^2}{q} \quad (12)$$

In Eq.(12), the total numbers of samples are q , and the total numbers of classes are m . The terms ob_i^p exp_i^p are observed and expected outputs for the input unit i w.r.t. p th training samples. Ultimately training of ANN may now be formulized as minimization of a function F as per Eq.(13).

$$AVG_{MSE} = Min. (F(\vec{V}_{wt-b})) \quad (13)$$

Connection weights and biases are modified over the iteration to minimize the AVG_{MSE} for all training instances. The flow of the process of the proposed study is given in Fig.1.

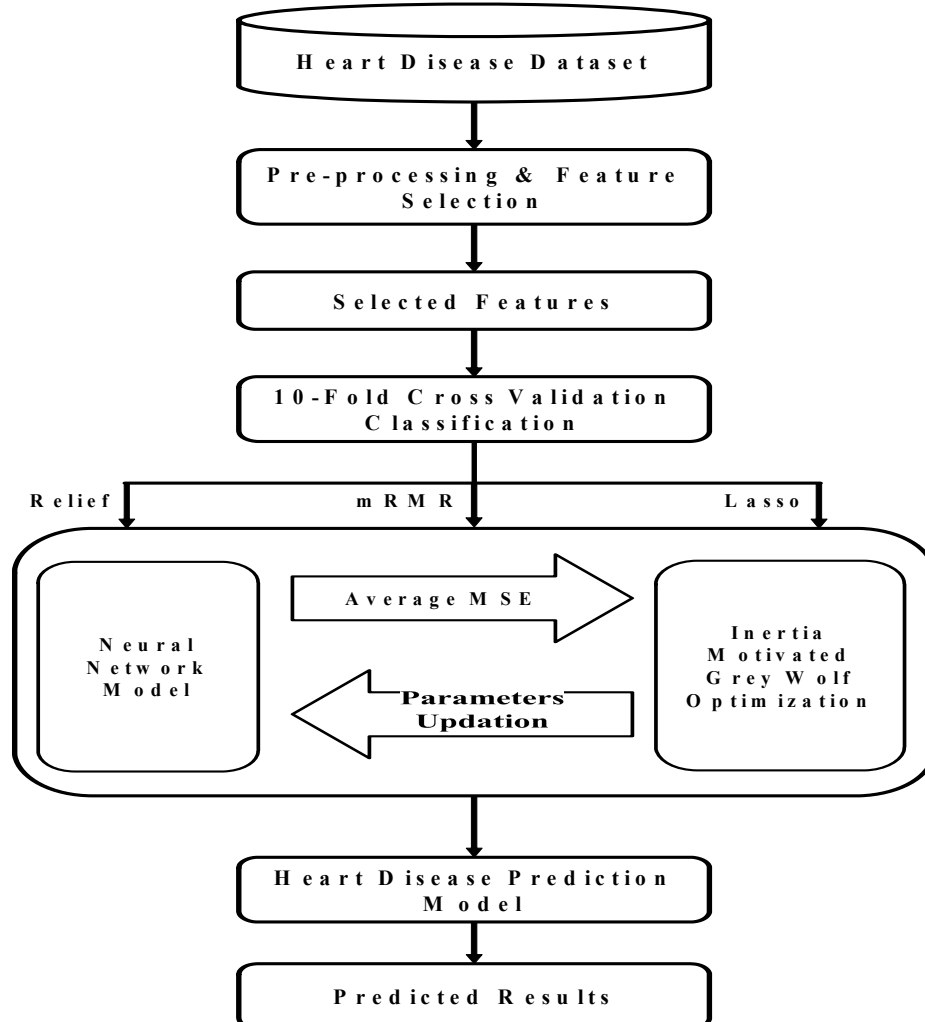


Fig.1. Process Flow of ANN-IMGWO model preceded by feature selection for heart disease diagnosis.

5.Experimental Results and Discussion

The experiment was conducted on core i5 Intel processor with a 4 GB RAM capacity and window10 based desktop system. The whole experiment has been performed with python 3.6 with scikit-learn, NumPy and Pandas libraries. Description and domains of the dataset attributes have been detailed in Table 1.Total number of 500 samples includes 150 as positive samples and 350 as negative samples. Positive samples are the patients having heart disease. Numeric and nominal (categorical) features are mentioned explicitly in the table under the column 'description and domain.' In addition, some features have binary domains, including Gender, HHTN, Smoker, Ex-Smoker, FHHD, Thyroid, AD, SA, VD, Dyspnea, DM, SM, ESTS, ESTD, TI, LVH and PRP. The motivation behind the experiment is to judge the classification performance of the IMGWO based ANN on a heart disease dataset that covers all kinds of features, including demographics, lifestyle, symptoms, lab tests and ECG information.

Sr. No.	Type of information	Name of the feature	Description and domain
1	Demographic and life style	Age	35-85 in years; Numeric
2		Gender	Male/Female
3		Weight	40-120 in kg; Numeric
4		Height	1.55-1.82 meters; Numeric
5		BMI	Body mass index(Kg/m ²); Numeric
6		HHTN	History of hypertension; Yes /No
7		Smoker	1 if yes 0 if no
8		Ex-Smoker	1 if yes 0 if no
9		FHHD	family history of heart disease ;Yes /No
10		Thyroid	1 if yes 0 if no
11		Exercise	3: more than 2.5 hours per week 2: less than 2.5 but more than 1 hours per week 1: less than 1 but more than 30 minutes per week 0: less than 30 minutes per week
12		AC	Alcohol consumption; nominal 3: regularly drink 2: moderately drink and 1: rarely drink 0: never drink
13		AD	Anxiety and Depression; 1 if Yes 0 if No
14		DSH	Daily sleep hours; nominal 3: more than 8 hours 2: less than 8 but more than 6 hours 1: less than 6 hours but more than 5 hours 0: less than 5 hours
15		SH	Sitting hours; Nominal 3: more than 8 hours 2: less than 8 but more than 6 hours 1: less than 6 but more than 4 hours 0: less than 4 hours
16		AP	Air pollution; Nominal 3: air quality index(AQI) is more than 200 2: AQI is more than 100 but less than 200 1: AQI is more than 50 but less than 100 0: AQI is less than 50
17	Symptoms Lab Test	SA	Sleep apnea ;1 if Yes 0 if No
18		VD	Vitamin D deficiency ;1 if Yes 0 if No
19		SC	Serum cholesterol (mg/dl); numeric
20		FBS	Fasting Blood Sugar (mg/dl); numeric
21		Dyspnea	1 if Yes 0 if No
22		DM	Diastolic mummer ;1 if Yes 0 if No
23		SM	Systolic murmur ;1 if Yes 0 if No
24		CPT	Chest pain type 1 : Typical angina 2: Atypical angina 3: Non-anginal pain, 4: Asymptomatic
25		HR	Heart rate at rest (beats per minute); Numeric
26		SBP	Systolic BP (mmHg); Numeric
27		DBP	Diastolic BP (mmHg);Numeric
28		LDL	Low density lipoprotein(mg/dl);Numeric
29		HDL	High density lipoprotein (mg/dl);Numeric
30		TRIG	Triglyceride (mg/dl);Numeric
31		Uric Acid	(mg/dl);Numeric
32		HB	Hemoglobin (g/dl): Numeric

33		WBC	White blood cells (per micro liter); Numeric
34		Platelet	per micro liter; Numeric
35		EP	Ejection percentage (15-60); Numeric
36		NST	Nuclear stress test ; Nominal 1: Normal 2: fixed defect 3: reversible defect
37		CCS	Coronary calcium score; Numeric
38	ECG	ESTD	Exercise ST depression; Yes/No
39		ESTS	Exercise ST slope; Yes/NO
40		TI	T-inversion; Yes/NO
41		LVH	Left ventricles hypertrophy; Yes/No
42		PRP	Poor R wave progression; Yes/No
43		Heart disease	Class Attribute; Yes /NO

Table 1. Heart disease dataset

The preprocessing and FS was done prior to building of the diagnosis model. Mean values fill missing values; normalization was performed with min-max normalization and feature selection was performed on three states of the art methods: Relief, mRMR and LASSO. The performance of the proposed methodology was judged on accuracies, sensitivities or recall, specificities, precision, f-measure, area under the curve (AUC), Kappa statistics and Mathew's correlation coefficient (MCC). The comparisons were performed with contemporary metaheuristics-based ANN models that include GA, PSO, GWO, FF and WOA. The results are obtained with 10-fold cross-validation.

FS was performed with three popular methods explained in section 3. The filter methodology was adopted in FS to avoid overfitting. In Tables 2 to 4, the different attributes got selected for classification. Relief determines CPT as the most important, NST as the second most important feature. All selected features are shown in table 2. Most of the ANN models trained over selected metaheuristic techniques got improved when features selected by Relief. The other 28 features have insignificant scores and don't contribute much towards accuracies and other performance metrics.

Rank	Feature name	Score
1	CPT	0.255
2	NST	0.245
3	ESTD	0.220
4	FBS	0.217
5	CCS	0.213
6	HR	0.195
7	LVH	0.189
8	Age	0.175
9	Gender	0.171
10	TG	0.165
11	Smoker	0.161
12	TI	0.137
13	HHTN	0.115
14	Exercise	0.111

Table 2. Feature selected by Relief with the rank and score

It is interesting to observe that the features like Age, Gender, Smoker, HHTN and Exercise that contains demographic and lifestyle information have significant but lesser scores than the features that include symptoms, lab test, and ECG information. Table 3 contains the features selected by mRMR with the corresponding score and rank. Age is having the highest score, succeeded by CPT. Age is a demographic feature, CPT is a symptom and ESTD is ECG information. CCS is a lab test that is obtained after fluoroscopy. Dyspnea, FHHD, ESTS are some significant features considered in mRMR but couldn't be selected by Relief.

Rank	Feature name	Score
1	Age	0.637
2	CPT	0.597
3	ESTD	0.567
4	CCS	0.558
5	LVH	0.539
6	HR	0.522
7	Smoker	0.511
8	Dyspnea	0.501
9	TI	0.445
10	FHHD	0.420
11	TG	0.352

12	ESTS	0.271
13	Gender	0.258
14	FBS	0.246

Table 3. Feature selected by mRMR with the rank and score

Table 4 contains the features selected by LASSO, which works based on the shrinking and regularization process to avoid overfitting. LASSO selected CPT as the most significant feature, like Relief. The second and third most significant features are ESTD and NST. ESTD is an ECG based feature while NST is a lab test based feature. It is important to note that AC is a lifestyle-based feature selected only by LASSO. Age, HHTN and exercise are other demographic and lifestyle-based features found significant in the classification process that ultimately diagnoses a heart ailment.

Rank	Feature name	Score
1	CPT	0.149
2	ESTD	0.147
3	NST	0.143
4	CCS	0.138
5	LVH	0.133
6	HR	0.125
7	Gender	0.116
8	Dyspnea	0.114
9	Age	0.107
10	TG	0.103
11	TI	0.093
12	AC	0.091
13	Exercise	0.085
14	HHTN	0.081

Table 4. Feature selected by LASSO with the rank and score

Table 5 contains the performance metrics for all the ANN models trained by different metaheuristic methods. These performance metrics include Accuracy, Sensitivity, Specificity Precision, F-measure and AUC. Accuracy is the primary measure that gives raw performance without considering the class distribution. Sensitivities and specificities are specific measures that are associated with true positive rate and true negative rate. An important reason to select sensitivity or recall as a performance measure is that it describes completeness. 100 % sensitivity ensures that all the subjects that have heart disease problems are classified as positive. It can be observed from table 5 that the ANN-IMGWO classification model outperforms as compared to ANN-GA, ANN-PSO, ANN-FF, ANN-WOA and ANN-GWO on almost all performance metrics. The average accuracy of ANN-IMGWO is 94.68 % which is the highest among all other ANN models, build with features selected Relief. A score of 93.17 % for recall (sensitivity) is achieved, which is the highest among all other ANN models. ANN-IMGWO also achieved the highest specificity that 96.30 %. The highest precision (0.96) and F-measure (0.94) are also achieved by IMGWO based ANN model succeeded by feature selection with Relief. In classification, precision gives the exactness for the positive classes and F-measure combined exactness and completeness for the positive classes. Therefore, in other words, ANN-IMGWO provides the 96 % exact and 94 % complete classification w.r.t. positive cases, which is much essential in medical diagnosis, especially in heart disease diagnosis.

Algorithm	Acc.	Sens./recall	Spec.	Precision	F-measure	AUC.
ANN-GA	90.29	90.06	90.48	0.89	0.90	0.931
ANN-PSO	92.29	91.93	92.59	0.91	0.92	0.940
ANN-FF	88.00	89.44	86.80	0.85	0.87	0.900
ANN-WOA	88.29	88.82	87.83	0.86	0.87	0.899
ANN-GWO	92.57	92.55	92.59	0.91	0.92	0.960
ANN-IMGWO	94.86	93.17	96.30	0.96	0.94	0.965

Table 5 Performance metrics of all classifiers with the selected features by Relief

AUC gives the true positive classification at the rate of false positive classification that is highly demanded in medical data classification since it reduces the cost of diagnosis. ANN-IMGWO achieved a score of 0.965 for AUC, which is the highest among all other models.

Performance measures achieved by different ANN-metaheuristic models are shown in Table 6. All of these models are preceded with the mRMR feature selection method. The performance of the classification model ANN-IMGWO is best among all the other ANN models.

Algorithm	Acc.	Sens./Recall	Spec.	Precision	F-measure	AUC.
ANN-GA	90.29	90.06	90.48	0.89	0.90	0.931
ANN-PSO	92.57	92.55	92.59	0.92	0.92	0.958
ANN-FF	90.00	91.30	88.89	0.91	0.91	0.928
ANN-WOA	89.43	89.44	89.42	0.87	0.88	0.910
ANN-GWO	92.86	92.55	93.12	0.93	0.93	0.955
ANN-IMGWO	95.43	93.79	96.83	0.97	0.95	0.969

Table 6 Performance metrics of all the classifiers with the selected features by mRMR

The performance metrics of all the classification models attained with features selected by LASSO is shown in Table 7. It is further notable that features selected by mRMR gave a maximum performance in terms of accuracy and all other measures mentioned in Tables 5 to 7.

Algorithm	Acc.	Sens./Recall	Spec.	Precision	F-measure	AUC.
ANN-GA	89.71	89.44	89.95	0.88	0.89	0.905
ANN-PSO	90.29	88.20	92.06	0.89	0.89	0.921
ANN-FF	87.71	86.34	88.89	0.86	0.86	0.889
ANN-WOA	90.00	88.82	91.01	0.89	0.89	0.917
ANN-GWO	92.00	90.68	93.12	0.92	0.91	0.958
ANN-IMGWO	94.00	91.93	95.77	0.93	0.92	0.959

Table 7 Performance of all classifiers with the selected features by LASSO

Figures 2 to 4 show the two vital statistics much needed for classification, especially for statistical validation. These statistics are Mathew's correlation coefficient (MCC) and Kappa statistics. MCC includes all of the corner values of the confusion matrix and has a range between -1 to +1. It quantifies the quality of the classification. The expression of MCC is given as in Eq.(14).

$$MCC = \frac{TP*TN-FP*FN}{\sqrt{(TP+FN)*(TP+FP)*(TN+FN)*(TN+FP)}} \quad (14)$$

The Kappa statistics evaluate the model's consistency and express the agreement between the expected and actual outcomes by the model. Its values also lie between -1 to +1. Values near 0 are considered by chance, and values near +1 are considered statically validated.

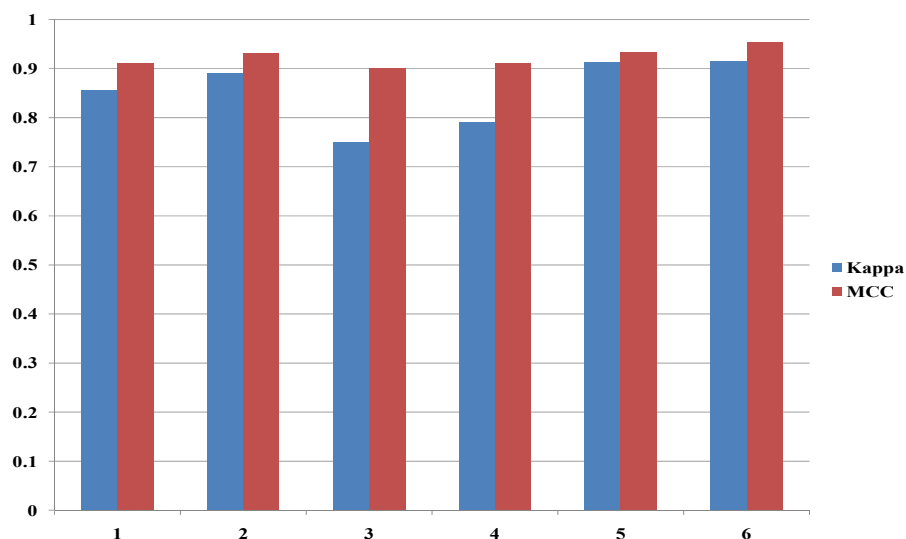


Fig. 2. Kappa statistics and MCC with features selected by Relief

It can be observed from Fig. 2 that MCC for the ANN-IMGWO is highest among all the other models, and so is the Kappa statistics. ANN-IMGWO attained a value of MCC 0.9496, which is very close to 1. So the model could be considered a reliable one. MCC values near 0 are an indicator of random prediction. Kappa for the ANN-IMGWO model was noted as 0.9012, which is near to +1 means the performance of ANN-IMGWO is not by chance or random, but the performance is reliable.

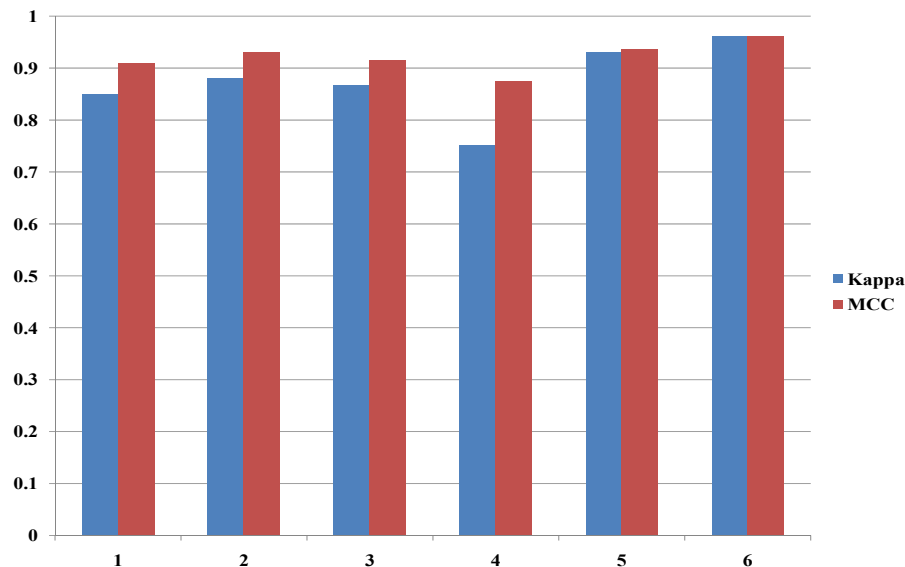


Fig. 3. Kappa statistics and MCC with features selected by mRMR

Fig. 3 shows that the highest MCC value achieved (0.9614) in the ANN-IMGWO model is near 1. In this figure, Kappa for the ANN-IMGWO is 0.9615. It is important to note that with features of mRMR the both of the statistical measures favor the IMGWO based ANN model for heart disease prediction.

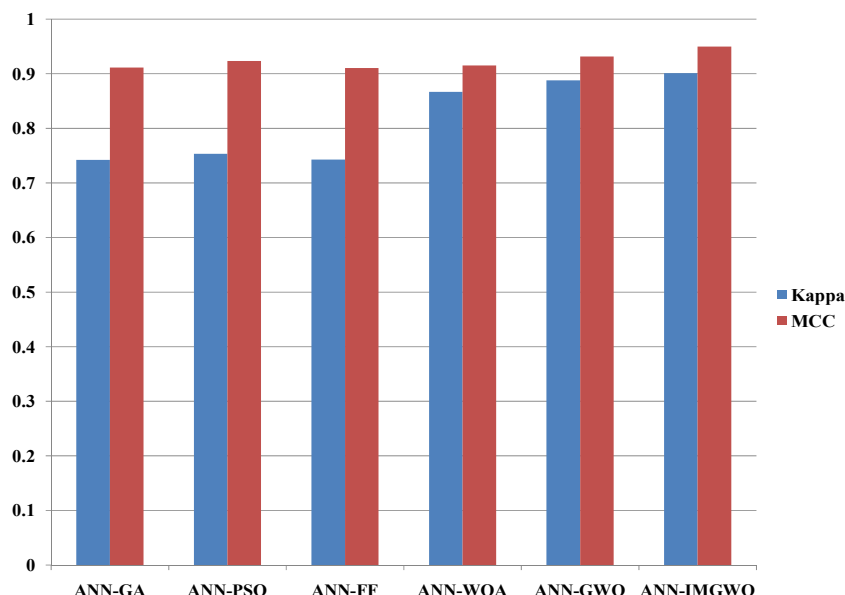


Fig. 4. Kappa statistics and MCC with features selected by LASSO

The values of MCC and Kappa for the models described by the feature of LASSO are again highest for IMGWO, which are 0.9541 and 0.9155, respectively.

Fig. 5 to 10 show a visual representation of the individual performance measures achieved by different ANN models with features selected by the 3 FS techniques.

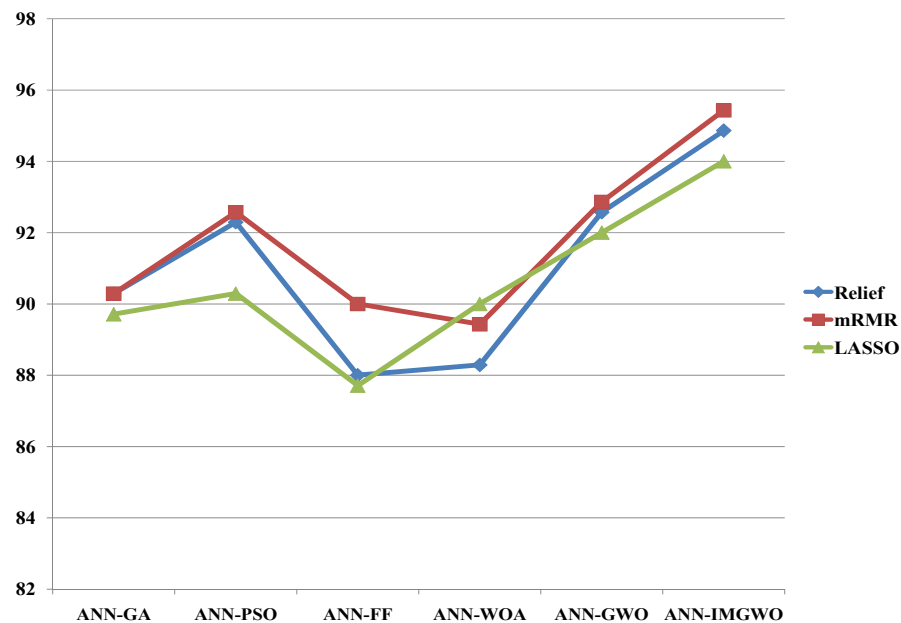


Fig. 5. Accuracies of different classifiers with selected features

Fig. 5 highlighted the improvement in ANN-IMGWO and that too with mRMR FS technique. It is also observed that Relief comes second and LASSO comes last towards improvement in accuracy.

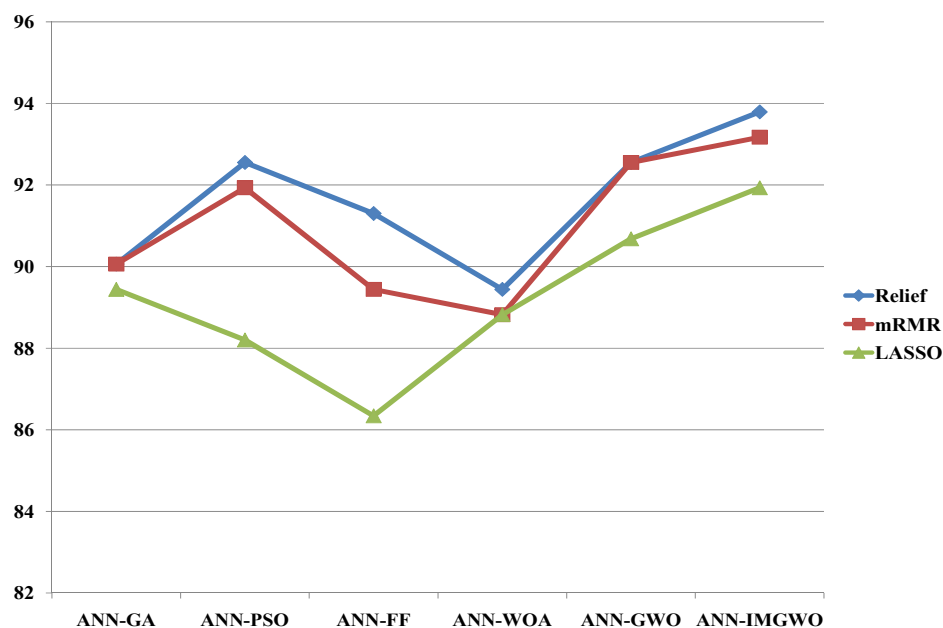


Fig. 6. Sensitivities/recall of different classifiers with selected features

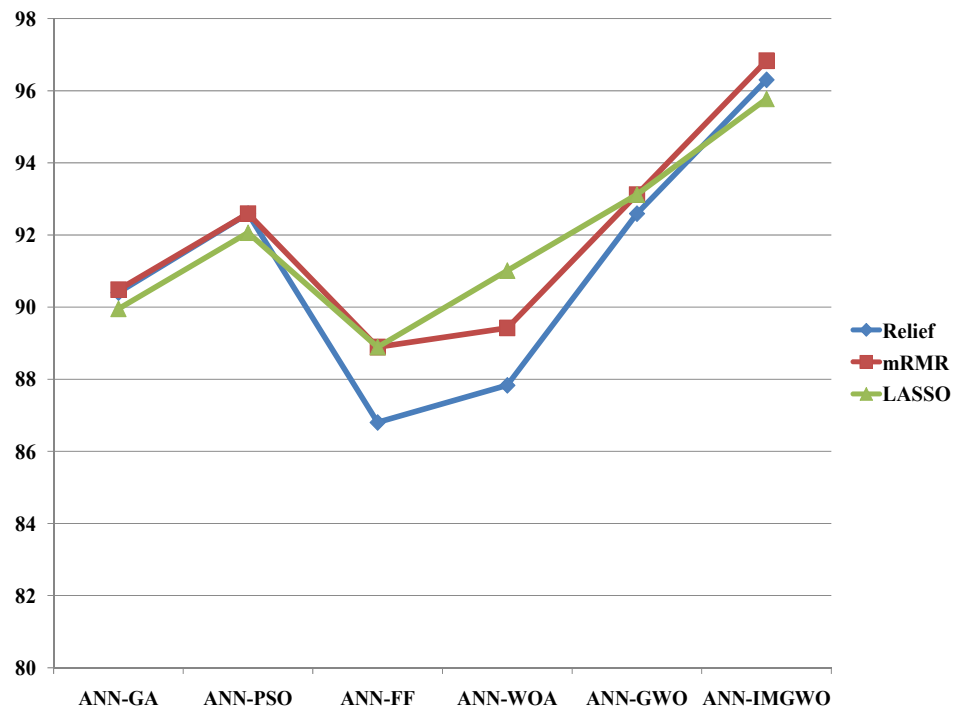


Fig. 7. Specificities of different classifiers with selected features

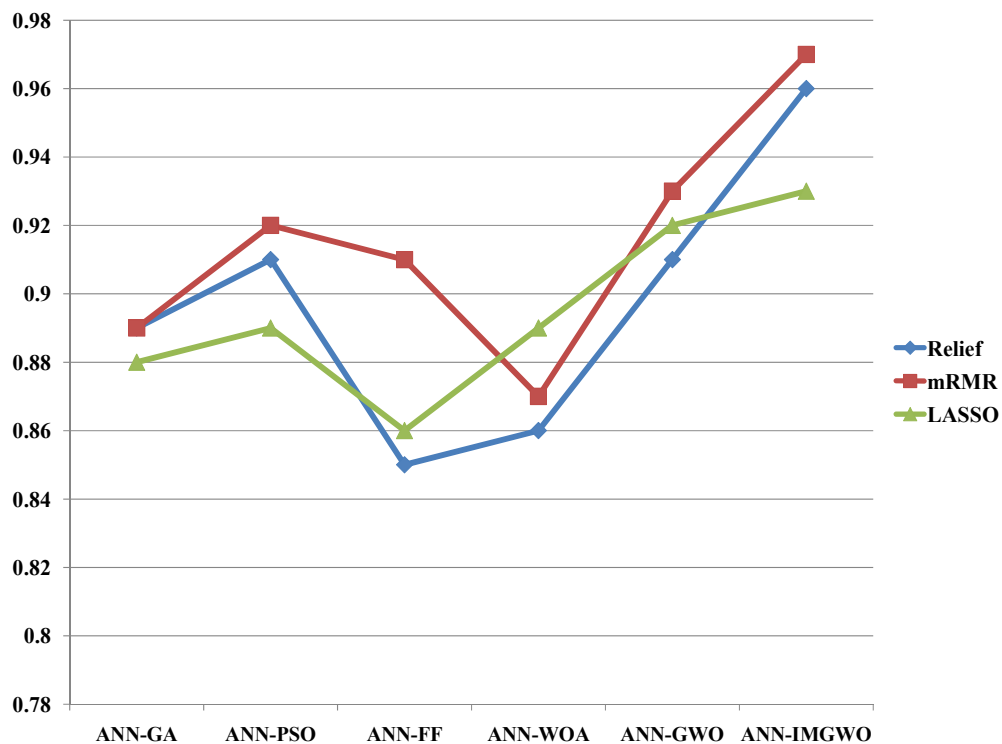


Fig.8. Precisions of the classifiers with selected features

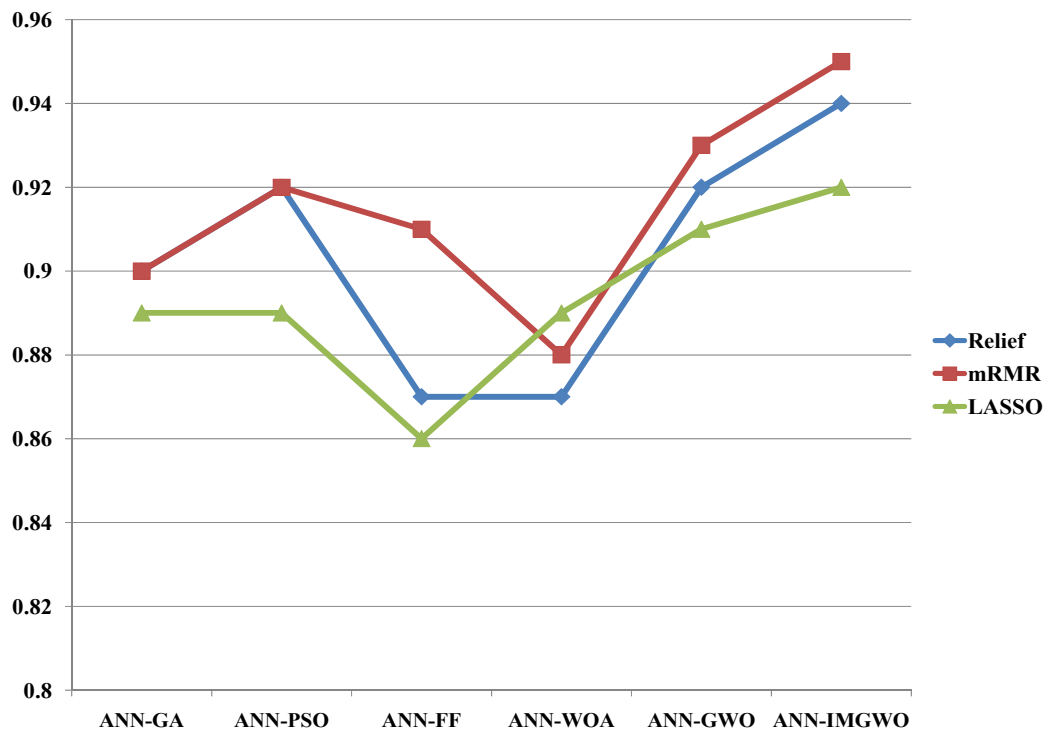


Fig. 9. F-Measures of the all the classifiers with the selected features

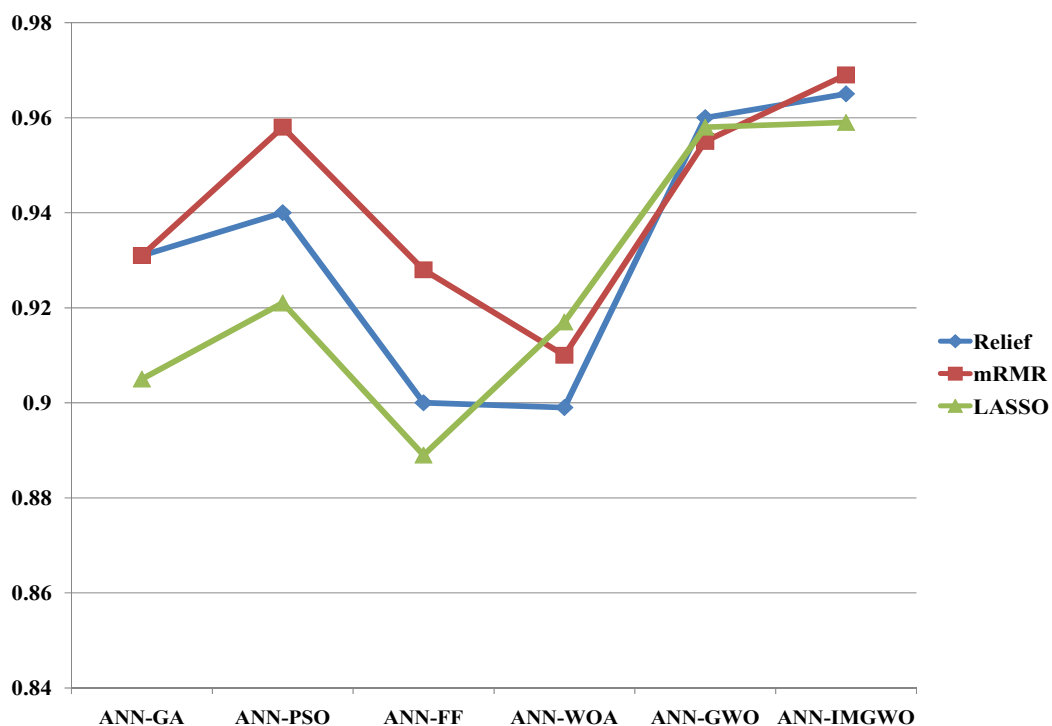


Fig. 10. Area under the curve (AUC) for different classifiers with the selected features

The better performance of IMGWO based learning of ANN might be associated with the following reasons:

- (1) Metaheuristic computations were better alternatives to the classical gradient descent or backpropagation training algorithms[21]. Moreover, in IMGWO, the essence of PSO and GWO were combined to combat

the local minima entrapment and poor convergence. Better balance between exploration and exploitation by the two metaheuristic techniques resulted in excellent search capability.

- (2) The features selected through the FS algorithms played a significant role in building the classification model and, ultimately heart diseases diagnosis model. The FS algorithm mRMR outperformed, as confirmed by the results. The irrelevant feature could have made the process of training complicated. Avoidance of irrelevant features increases the performance measures. Further, too many dimension creates the problem of overfitting which the features selected by mRMR efficiently handles.

Conclusion

This paper presented a heart disease diagnosis model that used ANN as a base classifier and IMGWO (a variant of GWO) as a training algorithm. Before the classification process, FS was performed with three prominent FS methods: Relief, mRMR and LASSO. CPT, ESTD, CCS, LVH, age, gender, TG, and TI are the features that are selected by all of the FS algorithms as the most discriminative feature. It is worth noting that features like AC and Exercise have also contributed to heart disease diagnosis. As per our knowledge and belief, AC and Exercise are the features that are included the first time in a heart disease dataset. The current research work has been carried out to investigate the performance of the ANN-IMGWO model preceded by the aforementioned feature selection algorithms on a newly collected dataset of 500 people from. During FS 14 most relevant features were selected by the three algorithms, and classification was performed. The ANN-IMGWO proved its competence in accuracy, sensitivity or recall, specificity, precision, f-measure and area under the curve over some contemporary metaheuristic based learning models like ANN-GA, ANN-PSO, ANN-FF, ANN-WOA and ANN-GWO. The proposed ANN-IMGWO model outperformed all other models w.r.t. all the performance metrics. The performance of the proposed model was further compared in terms of Kappa statistics and MCC to validate the results statistically. It is to note further that the highest accuracy was achieved with the features selected by mRMR. The mRMR-ANN-IMGWO framework achieved 95.43 % accuracy, 93.79% recall, 96.89% specificity, 0.97precision, 0.95 F-measure and 0.969 AUC. These performance measures are the highest among all the considered ANN-metaheuristic models (ANN-GA, ANN-PSO, ANN-FF, ANN-WOA and ANN-GWO) preceded by all the considered FS algorithms (Relief, mRMR and LASSO). The proposed hybrid model could assist in decision making during the diagnosis process of heart ailments.

References

- [1] "Cardiovascular diseases." <https://www.who.int/westernpacific/health-topics/cardiovascular-diseases> (accessed Mar. 29, 2021).
- [2] R. Alizadehsani *et al.*, "A database for using machine learning and data mining techniques for coronary artery disease diagnosis," *Sci Data*, vol. 6, no. 1, p. 227, Dec. 2019, doi: 10.1038/s41597-019-0206-3.
- [3] A. Gupta, R. Kumar, H. Singh Arora, and B. Raman, "MIFH: A Machine Intelligence Framework for Heart Disease Diagnosis," *IEEE Access*, vol. 8, pp. 14659–14674, 2020, doi: 10.1109/ACCESS.2019.2962755.
- [4] D. Kumar, "Decision tree classifier: a detailed survey," p. 24.
- [5] M. M. Ghiasi, S. Zendehboudi, and A. A. Mohsenipour, "Decision tree-based diagnosis of coronary artery disease: CART model," *Computer Methods and Programs in Biomedicine*, vol. 192, p. 105400, Aug. 2020, doi: 10.1016/j.cmpb.2020.105400.
- [6] E. Avci, "A new intelligent diagnosis system for the heart valve diseases by using genetic-SVM classifier," *Expert Systems with Applications*, vol. 36, no. 7, pp. 10618–10626, Sep. 2009, doi: 10.1016/j.eswa.2009.02.053.
- [7] I. Saini, D. Singh, and A. Khosla, "QRS detection using K-Nearest Neighbor algorithm (KNN) and evaluation on standard ECG databases," *Journal of Advanced Research*, vol. 4, no. 4, pp. 331–344, Jul. 2013, doi: 10.1016/j.jare.2012.05.007.
- [8] R. Das, I. Turkoglu, and A. Sengur, "Diagnosis of valvular heart disease through neural networks ensembles," *Computer Methods and Programs in Biomedicine*, vol. 93, no. 2, pp. 185–191, Feb. 2009, doi: 10.1016/j.cmpb.2008.09.005.
- [9] K. Mathan, P. M. Kumar, P. Panchatcharam, G. Manogaran, and R. Varadharajan, "A novel Gini index decision tree data mining method with neural network classifiers for prediction of heart disease," *Des Autom Embed Syst*, vol. 22, no. 3, pp. 225–242, Sep. 2018, doi: 10.1007/s10617-018-9205-4.
- [10] B. Zebardast, R. Rashidi, T. Hasanpour, and F. S. Gharehchopogh, "Artificial neural network models for diagnosing heart disease: a brief review," *IJAR*, vol. 6, no. 3, pp. 73–78, May 2014, doi: 10.7813/2075-4124.2014/6-3/A.11.
- [11] "High-Performance Personalized Heartbeat Classification Model for Long-Term ECG Signal," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 1, pp. 78–86, Jan. 2017, doi: 10.1109/TBME.2016.2539421.
- [12] S. Ekiz and P. Erdogmus, "Comparative study of heart disease classification," in *2017 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT)*, Istanbul, Turkey, Apr. 2017, pp. 1–4. doi: 10.1109/EBBT.2017.7956761.
- [13] Dr. A. G. and S. K., "HEART DISEASE DATA SET CLASSIFICATIONS: COMPARISONS OF CORRELATION CO EFFICIENT BY APPLYING VARIOUS PARAMETERS IN GAUSSIAN PROCESSES," *INDJCSE*, vol. 9, no. 4, pp. 130–134, Oct. 2018, doi: 10.21817/indjcse/2018/v9i5/180905019.
- [14] Y. Xing, J. Wang, Z. Zhao, and andYonghong Gao, "Combination Data Mining Methods with New Medical Data to Predicting Outcome of Coronary Heart Disease," in *2007 International Conference on Convergence Information Technology (ICCIT 2007)*, Gyeongju-si, Gyeongbuk, Korea, Nov. 2007, pp. 868–872. doi: 10.1109/ICCIT.2007.204.
- [15] J. Chen *et al.*, "A Comparison of Four Data Mining Models: Bayes, Neural Network, SVM and Decision Trees in Identifying Syndromes in Coronary Heart Disease," in *Advances in Neural Networks – ISNN 2007*, vol. 4491, D. Liu, S. Fei, Z.-G. Hou, H. Zhang, and C. Sun, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 1274–1279. doi: 10.1007/978-3-540-72383-7_148.
- [16] M. G. Tsipouras *et al.*, "Automated Diagnosis of Coronary Artery Disease Based on Data Mining and Fuzzy Modeling," *IEEE Trans. Inform. Technol. Biomed.*, vol. 12, no. 4, pp. 447–458, Jul. 2008, doi: 10.1109/TITB.2007.907985.
- [17] R. Alizadehsani *et al.*, "Machine learning-based coronary artery disease diagnosis: A comprehensive review," *Computers in Biology and Medicine*, vol. 111, p. 103346, Aug. 2019, doi: 10.1016/j.combiomed.2019.103346.
- [18] S. Safdar, S. Zafar, N. Zafar, and N. F. Khan, "Machine learning based decision support systems (DSS) for heart disease diagnosis: a review," *Artif Intell Rev*, vol. 50, no. 4, pp. 597–623, Dec. 2018, doi: 10.1007/s10462-017-9552-8.

- [19] R. P. Brent, "Fast training algorithms for multilayer neural nets," *IEEE Trans. Neural Netw.*, vol. 2, no. 3, pp. 346–354, May 1991, doi: 10.1109/72.97911.
- [20] M. Gori and A. Tesi, "On the problem of local minima in backpropagation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 1, pp. 76–86, Jan. 1992, doi: 10.1109/34.107014.
- [21] V. K. Ojha, A. Abraham, and V. Snášel, "Metaheuristic design of feedforward neural networks: A review of two decades of research," *Engineering Applications of Artificial Intelligence*, vol. 60, pp. 97–116, Apr. 2017, doi: 10.1016/j.engappai.2017.01.013.
- [22] N. Kumar and D. Kumar, "AN IMPROVED GREY WOLF OPTIMIZATION-BASED LEARNING OF ARTIFICIAL NEURAL NETWORK FOR MEDICAL DATA CLASSIFICATION," *JICT*, vol. 20, no. Number 2, pp. 213–248, Feb. 2021, doi: 10.32890/jict2021.20.2.4.
- [23] K. Uyar and A. İlhan, "Diagnosis of heart disease using genetic algorithm based trained recurrent fuzzy neural networks," *Procedia Computer Science*, vol. 120, pp. 588–593, 2017, doi: 10.1016/j.procs.2017.11.283.
- [24] M. Alweshah and S. Abdullah, "Hybridizing firefly algorithms with a probabilistic neural network for solving classification problems," *Applied Soft Computing*, vol. 35, pp. 513–524, Oct. 2015, doi: 10.1016/j.asoc.2015.06.018.
- [25] A. Lamba and D. Kumar, "Optimization of KNN with Firefly Algorithm," vol. 8, no. 2, p. 7.
- [26] S. Mirjalili and A. Lewis, "The Whale Optimization Algorithm," *Advances in Engineering Software*, vol. 95, pp. 51–67, May 2016, doi: 10.1016/j.advengsoft.2016.01.008.
- [27] A. A. Hameed, N. Ajlouni, A. Özyavaş, Z. Orman, and A. Güneş, "An Efficient Medical Diagnosis Algorithm Based on a Hybrid Neural Network with a Variable Adaptive Momentum and PSO Algorithm," in *International Congress on Human-Computer Interaction, Optimization and Robotic Applications Proceedings*, Jul. 2019, pp. 152–157. doi: 10.36287/setsci.4.5.030.
- [28] S. Mirjalili, "How effective is the Grey Wolf optimizer in training multi-layer perceptrons," *Appl Intell*, vol. 43, no. 1, pp. 150–161, Jul. 2015, doi: 10.1007/s10489-014-0645-7.
- [29] R. Detrano *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, Aug. 1989, doi: 10.1016/0002-9149(89)90524-9.
- [30] W. J. Long, S. Naimi, and M. G. Criscitiello, "Development of a knowledge base for diagnostic reasoning in cardiology," *Computers and Biomedical Research*, vol. 25, no. 3, pp. 292–311, Jun. 1992, doi: 10.1016/0010-4809(92)90044-B.
- [31] W. G. Baxt, "A Neural Network Trained to Identify the Presence of Myocardial Infarction Bases Some Decisions on Clinical Associations That Differ from Accepted Clinical Teaching," *Med Decis Making*, vol. 14, no. 3, pp. 217–222, Aug. 1994, doi: 10.1177/0272989X9401400303.
- [32] F. Azuaje *et al.*, "Predicting coronary disease risk based on short-term RR interval measurements: a neural network approach," *Artificial Intelligence in Medicine*, vol. 15, no. 3, pp. 275–297, Mar. 1999, doi: 10.1016/S0933-3657(98)00058-X.
- [33] E. J. Tkacz, "An Application of Wavelet Neural Network for Classification Patients with Coronary Artery Disease Based on HRV Analysis," p. 3.
- [34] M. Gudadhe, K. Wankhade, and S. Dongre, "Decision support system for heart disease based on support vector machine and Artificial Neural Network," in *2010 International Conference on Computer and Communication Technology (ICCCCT)*, Allahabad, Uttar Pradesh, India, Sep. 2010, pp. 741–745. doi: 10.1109/ICCCCT.2010.5640377.
- [35] R. Das, I. Turkoglu, and A. Sengur, "Effective diagnosis of heart disease through neural networks ensembles," *Expert Systems with Applications*, vol. 36, no. 4, pp. 7675–7680, May 2009, doi: 10.1016/j.eswa.2008.09.013.
- [36] S. Ari and G. Saha, "In search of an optimization technique for Artificial Neural Network to classify abnormal heart sounds," *Applied Soft Computing*, vol. 9, no. 1, pp. 330–340, Jan. 2009, doi: 10.1016/j.asoc.2008.04.010.
- [37] M. Kumar and S. -, "META-HEURISTICS TECHNIQUES IN CLOUD COMPUTING: APPLICATIONS AND CHALLENGES," *INDJCSE*, vol. 12, no. 2, pp. 385–395, Apr. 2021, doi: 10.21817/indjcse/2021/v12i2/211202055.
- [38] S. Nandy, A. Mitra, and T. Mukherjee, "Study of PSO and Firefly algorithm-based Feed-forward neural network training algorithms," in *2020 7th International Conference on Signal Processing and Integrated Networks (SPIN)*, Noida, India, Feb. 2020, pp. 908–913. doi: 10.1109/SPIN48934.2020.9070809.
- [39] P. Kora, "ECG based Myocardial Infarction detection using Hybrid Firefly Algorithm," *Computer Methods and Programs in Biomedicine*, vol. 152, pp. 141–148, Dec. 2017, doi: 10.1016/j.cmpb.2017.09.015.
- [40] Y. Zhou *et al.*, "Teaching learning-based whale optimization algorithm for multi-layer perceptron neural network training," *Mathematical Biosciences and Engineering*, vol. 17, no. 5, pp. 5987–6025, 2020, doi: 10.3934/mbe.2020319.
- [41] U. Agrawal, J. Arora, R. Singh, D. Gupta, A. Khanna, and A. Khamparia, "Hybrid Wolf-Bat Algorithm for Optimization of Connection Weights in Multi-layer Perceptron," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 16, no. 1s, p. 37:1-37:20, Apr. 2020, doi: 10.1145/3350532.
- [42] Z. Arabasadi, R. Alizadehsani, M. Roshanzamir, H. Moosaei, and A. A. Yarifard, "Computer aided decision making for heart disease detection using hybrid neural network-Genetic algorithm," *Computer Methods and Programs in Biomedicine*, vol. 141, pp. 19–26, Apr. 2017, doi: 10.1016/j.cmpb.2017.01.004.
- [43] D. Oreski, S. Oreski, and B. Klicek, "Effects of dataset characteristics on the performance of feature selection techniques," *Applied Soft Computing*, vol. 52, pp. 109–119, Mar. 2017, doi: 10.1016/j.asoc.2016.12.023.
- [44] H. Benhar, A. Idri, and J. L. Fernández-Alemán, "Data preprocessing for heart disease classification: A systematic literature review," *Computer Methods and Programs in Biomedicine*, vol. 195, p. 105635, Oct. 2020, doi: 10.1016/j.cmpb.2020.105635.
- [45] H. Benhar, A. Idri, and J. L. Fernández-Alemán, "A Systematic Mapping Study of Data Preparation in Heart Disease Knowledge Discovery," *J Med Syst*, vol. 43, no. 1, p. 17, Jan. 2019, doi: 10.1007/s10916-018-1134-z.
- [46] J. Han, "Data Mining: Concepts and Techniques," p. 129, 2013.
- [47] J. Li *et al.*, "Feature Selection: A Data Perspective," *ACM Comput. Surv.*, vol. 50, no. 6, pp. 1–45, Jan. 2018, doi: 10.1145/3136625.
- [48] M. Radovic, M. Ghalwash, N. Filipovic, and Z. Obradovic, "Minimum redundancy maximum relevance feature selection approach for temporal gene expression data," *BMC Bioinformatics*, vol. 18, no. 1, p. 9, Jan. 2017, doi: 10.1186/s12859-016-1423-9.
- [49] C. Puri *et al.*, "Classification of normal and abnormal heart sound recordings through robust feature selection," in *2016 Computing in Cardiology Conference (CinC)*, Sep. 2016, pp. 1125–1128.
- [50] P. Ghosh *et al.*, "Efficient Prediction of Cardiovascular Disease Using Machine Learning Algorithms With Relief and LASSO Feature Selection Techniques," *IEEE Access*, vol. 9, pp. 19304–19326, 2021, doi: 10.1109/ACCESS.2021.3053759.
- [51] A. Unler, A. Murat, and R. B. Chinnam, "mr2PSO: A maximum relevance minimum redundancy feature selection method based on swarm intelligence for support vector machine classification," *Information Sciences*, vol. 181, no. 20, pp. 4625–4641, Oct. 2011, doi: 10.1016/j.ins.2010.05.037.
- [52] H. Peng, F. Long, and C. Ding, "Feature Selection Based on Mutual Information: Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy," *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, vol. 27, no. 8, p. 13, 2005.
- [53] R. Tibshirani, "Regression Shrinkage and Selection Via the Lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996, doi: https://doi.org/10.1111/j.2517-6161.1996.tb02080.x.

- [54] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," *Advances in Engineering Software*, vol. 69, pp. 46–61, Mar. 2014, doi: 10.1016/j.advengsoft.2013.12.007.
- [55] W. Long, J. Jiao, X. Liang, and M. Tang, "Inspired grey wolf optimizer for solving large-scale function optimization problems," *Applied Mathematical Modelling*, vol. 60, pp. 112–126, Aug. 2018, doi: 10.1016/j.apm.2018.03.005.
- [56] A. Chatterjee and P. Siarry, "Nonlinear inertia weight variation for dynamic adaptation in particle swarm optimization," *Computers & Operations Research*, vol. 33, no. 3, pp. 859–871, Mar. 2006, doi: 10.1016/j.cor.2004.08.012.

Author's Profile



Mr. Narender Kumar, is currently pursuing his Ph.D. in department of Computer Science and Engineering(CSE) in Guru Jambheshwar University of Science and Technology(GJUS&T), Hisar, Haryana(India). He is an Assistant Professor in CSE department in GJUS&T Hisar. He has received his B Tech degree in Computer Science and Engineering from National Institute of Technology Hamirpur, H.P. (India) in 2005 and M Tech degree from Deenbandhu Chhotu Ram University of Science and Technology Murthal, Sonapat, Haryana (India) in 2011. His areas of interest are Data Mining and Machine Learning.



Prof. Dharmender Kumar, is currently working as a Chairperson and Professor in Department of Computer Science and Engineering in GJUS&T, Hisar, Haryana, India. He is has completed his Bachelor of Technology in Computer Science & Engineering (CSE) from GJUS&T, Hisar, Haryana, Master of Technology in CSE from Kurukshetra University, Kurukshetra and Ph.D in CSE from GJUS&T, Hisar. His main research work focuses on Artificial Intelligence, Data Mining and Big Data Analytics.