# Feature Selection using Hybrid Dragonfly Algorithm in a Heart Disease Predication System

**Namariq Ayad Saeed, Ziyad Tariq Mustafa Al-Ta'i**

*Abstract: The heart disease considers as one of the fatal disease in many countries. The main reason is due to the approved methods of diagnostic are not available to the ordinary people. Many studies have been done to handle this case with the use of both methods of soft computing and machine learning. In this study, a hybrid binary dragonfly algorithm and mutual information proposed for feature selection, support vector machine and multilayer perceptron employed for classification. The Statlog dataset used for experiments. Out of a total of 270 instances of patient data, 216 employees for the purpose of practicing, 54 of them used for the purpose of examining. Maximum classification accuracy of 94.44% achieved with support vector machine and 92.59% with multilayer perceptron on features selected with binary dragonfly algorithm, whereas with features obtained from mutual information combined with binary dragonfly (MI_BDA) algorithm support vector machine and multilayer perceptron attained an accuracy of 96.29%. The time algorithm takes reduced from 15.4 with binary dragonfly algorithm to 6.95 seconds with MI_BDA.*

*Index Terms: About four key words or phrases in alphabetical order, separated by commas.*

## I. INTRODUCTION

The heart is one of the key parts of the human body. Due to heart diseases, many people lost their lives. The physical condition of the person is stated using several signs which determined if the person is normal or not. These signs are blood pressure, blood pressure, pulse rate and temperature [1].

There are many elements that cause a problem to heart. They are smoking, high pulse rate, poor eating methodology, high blood pressure, and cholesterol, etc. Heart disease diagnosing is one of the complicated missions in the medical field. So it is necessary to develop an efficient disease prediction system for the detection of diseases [2].

Data mining is a smart and inventive approach which is used in various fields and many places. Data mining tools may uncover important data patterns and perform data analysis. It's assisting extremely to knowledge bases, business strategies, and scientific medical researches [3].

Feature Selection, which is also known as an attribute or a variable selection, is the procedures of choosing a set of relevant attributes from a big extent of features. It is computationally less expensive to mining a model on a reduced attribute, which makes the model understandable. Therefore, selecting a relevant attribute subset is a problem of

**Revised Manuscript Received on August 25, 2019**.

**Namariq Ayad Saeed**, Department of computer science, University of Diyala/ computer science, Diyala, Iraq.

**Ziyad Tariq Mustafa AL-Ta'i**, Department of computer science, University of Diyala/ computer science, Diyala, Iraq.

significant consequence for applying or learning data mining models[4].

## II. RELATED WORK

In [5],a structure for heart disease diagnosis introduced by N. Cong Long, P. Meesad, and H. Unger (2015). They proposed the system by employing an interval type2 fuzzy logic system (IT2-FLS), and groups of rough on the basis of attribute reducing. IT2FLS utilized parameters tuning by genetic hybrid, and a chaos firefly algorithm applied on Statlog dataset and achieved 88.3% accuracy.

Marjia Sultana, Afrin Haider and Mohammad Shorif Uddin (2016) [6],developed a model to predict heart disease using SMO, Kstar, Bayes Network, J48, and Multi-layer Perceptron through Weka software implemented on a statlog dataset to classify patients as having heart disease or not. SMO provides maximum precision of 84.07%, whereas achieved 76.66% with J48, 81.11% with Bayes net, 75.18% with Kstar, and 77.04% with MLP (Multi-Layer Perceptron).

XinoLiu, XiaoliWang, QiangSu, MoZhang, YanhongZhu, Q. Wang, and Qian Wang(2017) [7], developed a system includes two subsystems: the Relief and Rough set (RFRS) the system of choosing features and categorization system with ensemble classifier on the basis of C4.5 divider. The first system have 3 steps: (i) data discretization, (ii) attribute extraction by Relief algorithm, (iii) attribute reducing by rough group reduction algorithm. The heart Statlog dataset was used for experiments and attained an accuracy of 92.59%.

Uma N. Dulhare (2018)[8], developed a model with Particle Swarm Optimization (PSO) as attribute selection to improves the predictive accuracy of the Naïve Bayes which is employed to evaluate the selected attributes by PSO algorithm. The model employed to classify heart disease patient applied on Statlog heart disease dataset and attained accuracy of 87.91%.

## III. METHODOLOGY

### A. Mutual Information

Mutual information considers a filter method. It is a measure of the dependence among the stochastic variables. The main idea of employing MI in choosing the features is the attributes must greatly correlate with the goal class. The variables are independent if the value of MI is zero. Where MI between two random variables E= (e1, e2 … e3) and F= (f1, f2 … f3) [9]:

$$MI(E,F) = \sum_e \sum_f p(e,f) \log \frac{p(e,f)}{p(e)p(f)} \qquad (1)$$

Where p(e) and p(f) are the marginal probability distribution functions for E and F. If the relationship between E and F are very far, The MI between both of them will reduce and vice versa.

### B. Dragonfly Algorithm

The Dragonfly Algorithm is a swarm-based algorithm. It simulates hunting mechanisms and migration of idealized dragonflies. The method of hunting is referred to as the static swarm (i.e. the feeding), where dragonflies fly as small groups above a small area for the sake of searching for sources of food. The mechanism of the migration is referred to as the dynamic swarm (i.e. the migratory). In this stage, dragonflies fly in one direction in bigger groups and as a result, the swarm migrates. Similar to other approaches that are nature-inspired, this algorithm is made up of 2 phases, which are: the exploration, which simulates the static behavior of the swarm, and the exploitation, which simulates the dynamic behavior of swarming [10].

For the sake of modeling the swarming behavior of the dragonflies, 5 distinct behaviors have been employed, in the following way. In the equations below, X denotes the current search agent's position, N is the neighborhood size, and Xj is the jth neighbor of the search agent of X [10]:

- Separation is the mechanism which is performed by the search agent in order to stay away from other adjacent search agents. That behavior has been mathematically implemented as Equation (2):

$$s_i = \sum_{j=1}^{n} X - X_i \qquad (2)$$

- Alignment points to the way by which one of the individuals matches their velocity with the velocities of other adjacent individuals. This behavior can be mathematically implemented as Equation (3):

$$A_i = \frac{\sum_{j=1}^{N} V_j}{N} \qquad (3)$$

- Cohesion indicates the individual's tendency to fly in the direction of the adjacent center of mass. This behavior mathematically modeled as Eq. (4):

$$C_i = \frac{\sum_{j=1}^{N} X_j}{N} - X \qquad (4)$$

- Attraction refers to the tendency of individuals to fly in the direction of the source of food. The attraction between the i$^{th}$ solution and the food source mathematically modeled as Eq. (5):

$$F_i = F_{max} - X \qquad (5)$$

Where F$_{loc}$ presents the food source position

- Distraction indicates the individuals' tendency to fly away from an enemy. The distraction among the i$^{th}$ solution and the enemy mathematically modeled as Eq. (6):

$$E_i = E_{loc} + X \qquad (6)$$

Where E$_{loc}$ represents the enemy's position.

When using DA, it assumed to update the food source fitness and location by means of the best candidate (search agent) yet. Moreover, the worst candidate is used to update the fitness and location on the enemy. This manage to convergence in the direction of promising areas of the search space and divergence outwards non-promising areas. According to the framework of the PSO algorithm, there are 2 vectors used for the purpose of updating the position of dragonfly: the step vector (ΔX) which is like to the position vector and the velocity vector in PSO. The step vector denotes the dragonflies' direction of movement. The step vector formulated as Eq. (7)

$$\Delta X_{i+1} = (sS_i + aA_i + cC_i + fF_i + eE_i) + wX_i \qquad (7)$$

Where s, w, a, c, f, and e present the weights of the separation S$_i$, alignment A$_i$, attraction in the direction of the food source F$_i$, cohesion C$_i$, and distraction from the enemy E$_i$ of the i$^{th}$ individual's individually. Those weights permit the Dragonfly Algorithm to reach various behaviors of intensification and exploration throughout optimization. The position of an individual updated as in Eq. (8):

$$X_{t+1} = X_t + \Delta X_{t+1} \qquad (8)$$

Where t is the current iteration.

The beginning of algorithm is with the creation of a random initial population. The definition of both step vectors and positions of dragonflies will be random. The algorithm is repeated in every iteration in order to implement the steps listed below. This process is continuous till a termination criterion is satisfied. The first step is evaluating every one of the individuals in the population by utilizing a fitness function. The second step is updating the primary coefficients. The next step is updating the alignment (A), separation (S), food source (F) cohesion (C), enemy (E) with the use of Equations (2)–(6). The final step is updating the position and the step vectors by Eqs. (8) and (7), individually [10].

❖ Binary Dragonfly Algorithm

The search space presented as a hypercube in a binary optimization problem, in which an individual can alter their position from one place to another via altering one or more bits of its position vector {x= x1…x2xd}. Due to the fact that the original Dragonfly Algorithm has been designed for dealing with the issues of continuous optimization, the location of an individual is changing when the current position added to the step vector [10].

Normally, production of the probability of changing a position's elements to 1 or 0 is done by transfer functions, this changing is based on how much the velocity (step vector) of the ith search agent in the d-th dimension in the actual iteration (t) as a parameter of input. In a former research by Seyedali Mirjalili, calculating the likelihood of changing the continuous positions into binary is done using the transfer function of Eq. (9) [10].

$$T(V_d^t(t)) = \left| V_d^t(t) / \sqrt{1 + (V_d^t(t))^2} \right| \qquad (9)$$

$T(V_d^t(t))$ is afterwards employed for converting the ith element of the position vector to 0 or 1 using Equation (10)

$$X(t+1) = \begin{cases} \neg X_i & r < T(V_k^i(t)) \\ X_i & r \geq T(V_k^i(t)) \end{cases} \qquad (10)$$

### C. Naïve Bayes

algorithm is based on Bayesian theorem along to some presumptions. Where it assumes that the presence any class is not related of another class. Using Bayes theorem, one can get the likelihood of happening an event regarding to another event which has happened previously. NB based on conditional probability.

The Bayes theorem is in the following way [11]:
Let X={x1, x2, .....,xn} be a group of n attributes.
In Bayesian, X is treated as an evidence and H be some hypothesis, meaning that the data of X is a part of a certain class C.

P(H|X) must be calculated, the likelihood that hypothesis H holds considering the evidence, which is the data sample X. Based on Bayesian theorem the P (H|X) can be stated as

$$P(H \mid X) = \frac{P(X \mid H)P(H)}{P(X)} \qquad (11)$$

### D. Support Vector Machine

SVM is a classification approach used the hyperplanes, where the chosen hyperplanes is the one with the superior distance between 2 classes. The points falling on the boundaries are referred to as support vectors. One of the advantages of this method is it overcomes the high dimensionality problem[12].

SVM is a class of universal feedforward networks as Radial-basis function networks. SVM is an estimate application to the technique of reducing structural risk. It relies on the truth which said that the average of error in a learning machine on test data bounded by term which is dependent on Vapnik-Chervonenkis (VC) dimension and the total of the average of the training-error. In case of pattern classification problem, a good generalization operation can get from the support vector machine [13].

Best Hyperplane for patterns: Training sample where $y_i$ represents the output goal for the i-[th] sample while $x_i$ represents the equivalent input pattern. With pattern introduced by the sub-set $y_i$= +1 and the pattern introduced by the sub-set $y_i$=-1 are linearly separable [13]. The hyperplane that can does the separation, can be expressed in the following equation:

$$w^T x + b = 0 \qquad (12)$$

Where w is an adjustable weight vector, x is an input vector, and b is a bias. Thus,

$$w^T x_i + b \geq 0 \quad \text{for} \quad y_i = +1 \qquad (13)$$

$$w^T x_i + b < 0 \quad \text{for} \quad y_i = -1 \qquad (14)$$

In case of known weight vector w and a bias b, Eq. 3 defines the separation between the closest data point and hyperplane which is referred to as the separation.

### E. Multilayer Perceptron

The multilayer perceptron neural network (MLP), demonstrates that it contains multiple layers. Linearly separable problems solved by single layer perceptron, in other hand, not all the complicated problems are linearly separable, in this case one or more layers are used besides the single layer perceptron in order to get the solutions to such problems, which is called multilayer perceptron.

MLP can be considered as a feed-forward NN that have one or more hidden layers, as demonstrated in Figure 1. Usually they can be utilized for pattern recognition, a prediction depends on the approximation and input information, classification of input patterns[14].
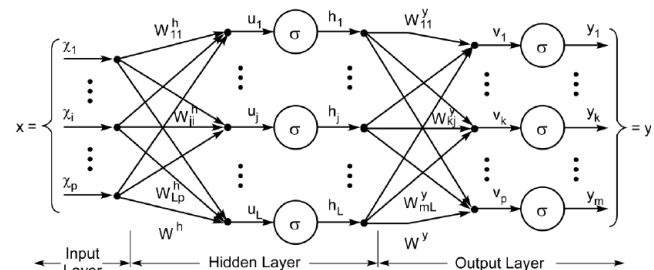

Fig.1 The Architecture of the Multi-layer Perceptron NN.

### F. Dataset

The Statlog dataset (UCI) mentioned in this study was achieved by the University of California Irvine (UCI) Machine Learning Repository. The Statlog dataset includes 270 instances of patient data. Table I illustrates the Statlog dataset features with their definitions[15].

TABLE I: STATLOG DATASET

| NO. | Name | Description |
|---|---|---|
| 1 | age | Age in years |
| 2 | Cp | Chest pain type |
| 3 | sex | Sex |
| 4 | trestbps | Resting blood pressure in mm Hg |
| 5 | Fbs | Fasting blood sugar is larger than 120 mg/dl or not |
| 6 | Chol | Serum cholesterol in mg/dl |
| 7 | Thalach | Maximum heart rate achieved |
| 8 | restecg | Resting electrocardiographic results |
| 9 | exang | Exercise induced angina |
| 10 | Slope | The slope of the peak exercise |
| 11 | oldpeak | ST depression induced by exercise relative to rest |
| 12 | Thal | The heart status |
| 13 | Ca | Number of major vessels (0-3) colored by fluoroscopy |
| 14 | Num | Diagnosis of heart disease (1 = Absence, 2 = presence) |

## IV. PROPOSED METHOD

The suggested approach used mutual information to determine the feature-class mutual information to get an extreme relevance between selected attributes and the target class. This hybrid algorithm consists of a wrapper and filter type for attribute subset selection employed to enhance the operation of heart disease prediction system. Fig.2 shows the suggested hybrid framework.
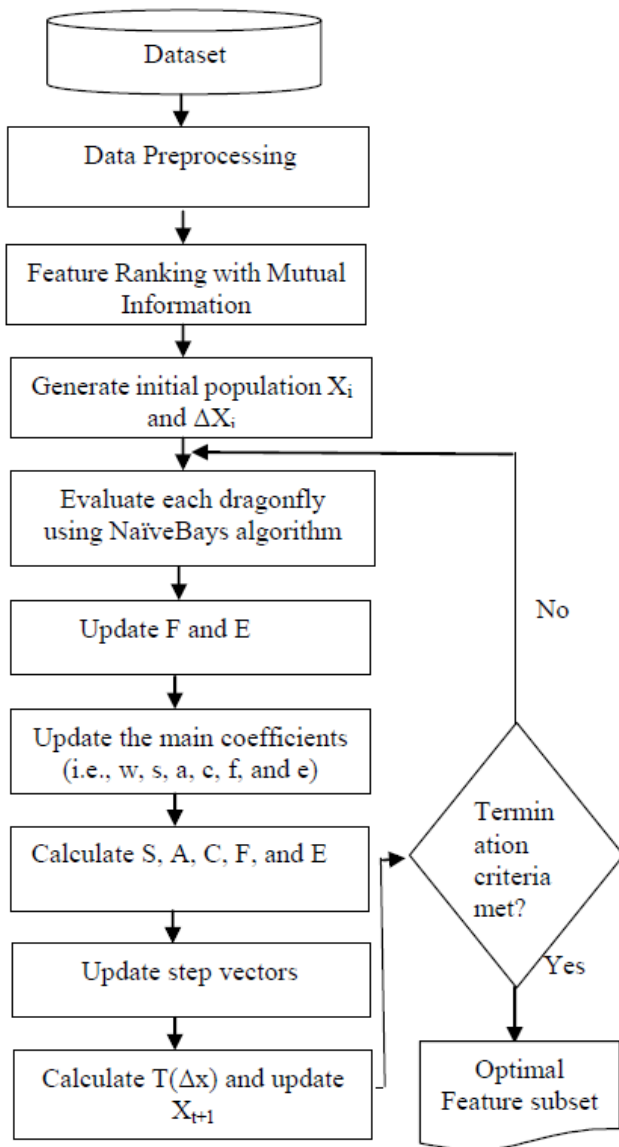
**Fig.2.The suggested hybrid framework for attribute selection**

Based on Fig. 2, attribute choosing process was done using the steps listed below:

Step 1: Applying mutual information on the dataset and remove features less than or equal to zero.

Step 2(Initialization): Several attribute subsets were generated in a random way as the initial population of dragonfly. Where every dragonfly consists of a binary string, its length is m (the number of features in the dataset which obtained from the first step), where the presence of a corresponding feature in a dragonfly is expressed by the value of every bit.

Step 3(Evaluation): In this step, the fitness function used to determine the dragonfly. Where for each dragonfly (feature subset), training instances in the dataset which have only attributes in the equivalent attribute subset is used to form the NB model. Furthermore, it was measured roughly using testing instances. The optimal position of dragonfly, i.e. best attribute subset is determined using a fitness function, and it's normally located between the optimal position in the last iterations, and the present position of dragonflies.

Fitness=Accuracy of NB algorithm                    (15)

Step 4 (Updating): Updates the primary coefficients. The alignment (A), separation (S), food source (F) cohesion (C), enemy (E) are updated with the use of Equations (2 to 6). Updates the positions of elements and step vector using Eqs.(8) and (7).

---

**Algorithm I. Mutual Information-Binary Dragonfly Algorithm (MI_BDA) Feature selection method**

**Input:** Dataset.
**Output:** Chosen features F`.
**Steps:**
1. Initialize n (number of dragonflies), m (number of attributes), and $t_{max}$ (maximum number of iteration).
2. Initialize t=0 , $X_i$ (positions),$\Delta X_i$ (step vectors) for i=1,2, …,n randomly with a binary value.
3. Initialize the fitness of dragonflies fitness(i)=$[0]_{nx1}$
4. Initialize the velocity of dragonflies $V_i(t)$ )=$[0]_{1xn}$ for i=1,2,…,n
5. Set best_v=0 as the best fitness and as the best position best_p=$[0]_{1xn}$
6. Initialize enemy_fit=$[0]_{1xn}$ ,enemy_pos=$[0]_{1xn}$ , food_fit=$[0]_{1xn}$ , food_pos=$[0]_{1xn}$ for i=1,2,…,n
7. **for** each attribute from the original attribute set(d=1,2,…,m)
8.     Compute MI($f_d$ ;y)
9.     If MI($f_d$ ;y)>0
10.       ff**a = ffa** ∪ $f_d$
11.     end if
**12. end for**
**13. while** t<$t_{max}$
14.     Calculate w, e, s, a, c, f from following equations
15.     $e = 0.1 - t$
16.     $e = 0.1 - t * (0.1/(t_{max} / 2))$
17.     If e<0
18.       e=0
19.     end if
20.     s=2*rand()*e
21.     a=2*rand()*e
22.     f=2*rand()*e
23.     If t>(3*$t_{max}$/3)
24.       E=0
25.     end if
26.     Determine Fitness (t) by determining position of dragonflies by the suggested MI_BDA fitness function
27.     for I =1 to n
28.       If fitness[i]>food_fit[i]
29.         food_fit[i]=fitness[i]
30.       end if
31.       If fitness[i]<enemy_fit[i]
32.         enemy_fit[i]=fitness[i]
33.       end if
34.     Calculate S using eq.[2]
35.     Calculate A using eq.[3]
36.     Calculate C using eq.[4]
37.     Calculate F using eq.[5]
38.     Calculate E using eq.[6]
39.     for j=1 to d
40.       Update ΔX using eq.[7]
41.       Calculate T using eq.[9]
42.       Update X using eq.[10]
43.     end for
44.     t=t+1
45. end while
46. Return food_pos  ( as best features subset)

## V. EXPERIMENTAL RESULTS

### A. Evaluation Criteria

The evaluating of the suggested system was executed depending on accuracy, recall, precision, and F-measure tests which use the true negative (TN), true positive (TP), false negative (FN), false positive (FP) terms [16], and these criteria are computed as follows:

$$F_{measure} = \frac{2 * TP * precision}{TP + precision} \quad (16)$$

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (17)$$

$$precision = \frac{TP}{TP + FP} \quad (18)$$

$$recall = \frac{TP}{TP + FN} \quad (19)$$

### B. Results

In proposed method samples divided into two independent dataset that randomly selected 80% of the dataset is used to construct and train the classifier while the remaining 20% of the dataset used for testing.

First, we used Z-score normalization to normalize the original data. In the feature ranking with MI the equation implemented on the dataset and the features less or equal to the threshold removed, threshold was set to zero. Based on these Age, Fbs and Restecg were removed. The output of feature ranking model is fed to BDA. Table II, Figure 3 and 4 illustrate the result of proposed system.

Parameters of BDA are population size is 20, maximum iteration number is 100.

TABLE II: THE RESULTS OF PROPOSED SYSTEM

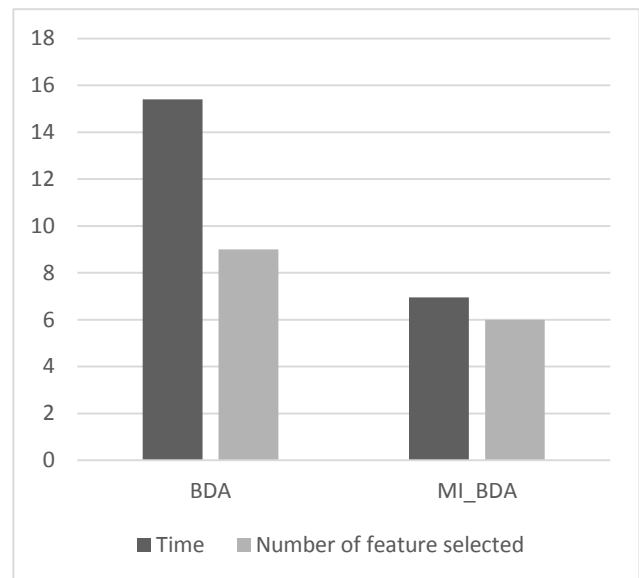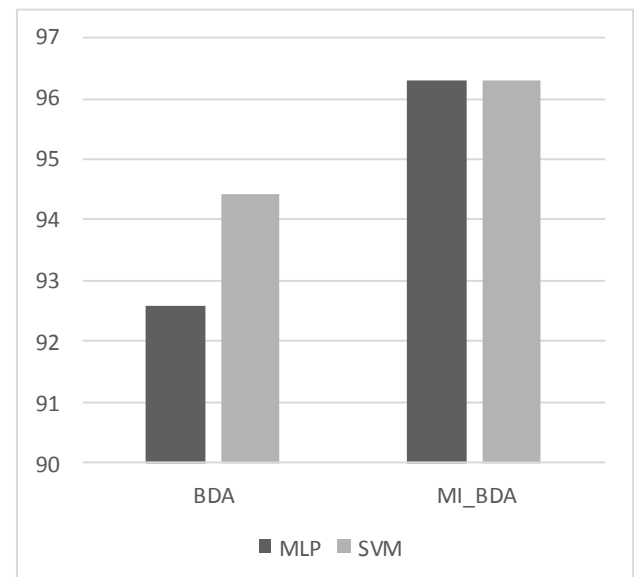| Algorithm | Selected features | Time | Metrics | SVM | MLP |
|---|---|---|---|---|---|
| MI_BDA | 1000111011 | 6.95 Sec | Accuracy | 96.29 | 96.29 |
| | | | Precision | 96.87 | 96.87 |
| | | | Recall | 95.83 | 95.83 |
| | | | F-measure | 96.21 | 96.21 |
| BDA | 0011011111011 | 15.4 Sec | Accuracy | 94.44 | 92.59 |
| | | | Precision | 94.6 | 93.03 |
| | | | Recall | 94.16 | 92.08 |
| | | | F-measure | 94.34 | 92.42 |



**Fig.3. Results of Statlog dataset**



**Fig.5. Accuracy of Statlog dataset**

The results which we get is compared with other studies are listed in table III.

TABLE III: CDMPARSION OF OUR RESULTS WITH OTHER STUDIES

| Author | Method | Classification accuracy |
|---|---|---|
| This study | MI_BDA | 96.29% |
| Nguyen Cong Long, Phayung Meesad, and Herwig Unger[5] | Interval type-2 fuzzy logic system (IT2FLS), and rough sets based attribute reduction. | 88.3% |
| XinoLiu, XiaoliWang, QiangSu, MoZhang, YanhongZhu, Qiugen Wang, and Qian Wang[7] | RFRS classification system | 92.59% |
| Uma N. Dulhare[8] | PSO+ Naïve Bayes | 87.91% |
| Marjia Sultana, Afrin Haider and Mohammad Shorif Uddin[6] | Bayes net | 81.11% |

The MI_BDA consist of two subsystems: Mutual Information and Binary Dragonfly Algorithm for feature selection and classification sub-system using multilayer perceptron and support vector machine. The Statlog dataset from UCI machine learning repository was chosen for the sake of examining the system. The test results demonstrate that the mixture (Sex, Thalach, Exang, Oldpeak, Ca, Thal) reaches the peak classification accuracy (96.29%) with SVM and MLP. The results show that classification accuracy of support vector machine increased from (94.44% to 96.29%) when using MI_BDA as feature selection approach as well as the accuracy of multilayer perceptron increased from (92.59% to 96.29%). The time algorithm takes is minimized from 15.4 sec with binary dragonfly algorithm to 6.95 sec with suggested approach

## VI. CONCLUSION

In this paper, Mutual Information and binary dragonfly algorithm-based classification system is suggested for heart disease predication system. In the suggested method MI and BDA are combined to classify problems of heart disease in rapid and an effective way.

Depending on the experimented result combining mutual information with BDA enhanced the performance of the proposed method. This method can be employed to detect if the patient suffering from heart disease to help specialist makes a decision. Other optimization methods can be used instead of BDA for feature selection may enhance the accuracy of prediction. The performance of the suggested system is better than the present system in the literature.

## REFERENCES

1. Gnaneswar B, Ebenezar Jebarani M.R.(2017)," A REVIEW ON PREDICTION AND DIAGNOSIS OF HEART FAILURE", Int. Conf. Innov. Information, Embed. Commun. Syst.,11-13.
2. Keerthana T K(6 may 2017)," Heart Disease Prediction System using Data Mining Method", International Journal of Engineering Trends and Technology (IJETT),(Vol.47,no.6), 361-363.
3. C.Sowmiya, Dr.P.Sumitra(2017)," Analytical Study of Heart Disease Diagnosis Using Classification Techniques C.Sowmiya ", IEEE INTERNATIONAL CONFERENCE ON INTELLIGENT TECHNIQUES IN CONTROL,OPTIMIZATION AND SIGNAL PROCESSING.
4. Ruba Arockia Archana, Dr.M.S. Thanaba l(2016 April 10(4))," Feature Selection for Heart Disease using Enhanced Cuckoo Search ", ADVANCES in NATURAL and APPLIED SCIENCES,50-56.
5. Nguyen Cong Long, Phayung Meesad, Herwig Unger(29 June 2015)," A highly accurate firefly based algorithm for heart disease prediction", Expert Systems with Applications.
6. Marjia Sultana, Afrin Haider, Mohammad ShorifUddin(2016), " Analysis of Data Mining Techniques for Heart Disease Prediction"
7. Xiao Liu, Xiaoli Wang, Qiang Su,Mo Zhang, Yanhong Zhu, Qiugen Wang, QianWang(3 January 2017)," A Hybrid Classification System for Heart Disease Diagnosis Based on the RFRS Method", Computational and Mathematical Methods in Medicine ,Volume 2017.
8. Uma N Dulhare(2018)," Prediction system for heart disease using Naive Bayes and particle swarm optimization", Biomedical Research, 29 (12): 2646-2649.
9. Shuang Cang (2011)," A Mutual Information based Feature Selection Algorithm", 4th International Conference on Biomedical Engineering and Informatics (BMEI), 2241-2245.
10. Majdi Mafarja , Ibrahim Aljarah, Ali Asghar Heidari, Hossam Faris, Philippe Fournier-Viger, Xiaodong Lie, Seyedali Mirjalili (2018)," Binary dragonfly optimization for feature selection using time-varying transfer functions", Knowledge-Based Systems journal.
11. K.Sudhakar, Dr. M. Manimekalai(January 2014)," Study of Heart Disease Prediction using Data Mining", International Journal of Advanced Research in Computer Science and Software Engineering , Volume 4, Issue 1.
12. Ali Al Bataineh (June 2019)," A Comparative Analysis of Nonlinear Machine Learning Algorithms for Breast Cancer Detection", International Journal of Machine Learning and Computing, Vol. 9, No. 3.
13. CHRISTOPHER J.C. BURGES(1998)," A Tutorial on Support Vector Machines for Pattern Recognition", Data Mining and Knowledge Discovery, 2, 121–167.
14. Jayshril S. Sonawane, D. R. Patil(2014) ," Prediction of Heart Disease Using Multilayer Perceptron Neural Network", ICICES2014 .
15. DATASETURL: http://archive.ics.uci.edu/ml/datasets/statlog+
a. (heart).
16. Alireza Baratloo, Mostafa Hosseini, Ahmed Negida, Gehad El Ashal (2015)," Part 1: Simple Definition and Calculation of Accuracy, Sensitivity and Specificity", Emergency : 3 (2): 48-49.

## AUTHORS PROFILE

**Namariq Ayad Saeed** is currently a M.S student at the department of Computer Science. Diyala University, Computer science, Iraq/Diyala. She had received a B.Sc. degree in computer science from Diyala University, Computer science, Iraq, Diyala (2011). Her research interest include data mining, and metaheuristic optimization.

**Ziyad Tariq Mustafa Al-Ta'i** received a B.Sc. degree in Electrical Engineering from Baghdad University (1987). In (1995) he obtained an M.S. degree in Computer Science from the University of Technology / Iraq. In (2002) got a Ph.D. in Computer Science from the University of Technology / Iraq. He had been a Professor in (2014) at the University of Diyala / Iraq. His research interests include network security and AI techniques.