Medical Data Classification using Fuzzy Min Max Neural Network Preceded by Feature Selection through Moth Flame Optimization

Ashish Kumar Dehariya[1]

Computer Engineering Department, Institute of Engineering and Technology
Devi Ahilya Vishwavidyalaya, Indore, India

Pragya Shukla[2]

Professor, Computer Engineering Department, Institute of Engineering and Technology
Devi Ahilya Vishwavidyalaya, Indore, India

Abstract—**Prediction of the diseases are possible using medical diagnosis system. This type of health care model can be developed using soft computing techniques. Hybrid approaches of data classification and optimization algorithm increases data classification accuracy. This research proposed applications of Moth Flame optimization (MFO) and Fuzzy Min Max Neural Network (FMMNN) for the development of medical data classification system. Here MFO algorithm considers bulk of features from the disease dataset and produces optimized set of features based on fitness function. MFO is able to avoid local minima problem and this is the main cause behind production of optimal set of features. Optimized features are then passed to FMMNN for classification of malignant and benign cases. As classification is concerned, model experiment achieved 97.74% accuracy for Liver Disorders and 86.95 % accuracy for Pima Indian Diabetes dataset. Improving the medical data classification accuracy is directly related to attain good human health.**

Keywords—*Moth flame optimization; nature inspired optimization; feature selection; fitness function; fuzzy min-max neural network*

I. INTRODUCTION

Nowadays, specialized computer software is very popular for medical data classification. For the development of smart healthcare system in smart cities use medical diagnosis software plays a vital role. In this type of medical data classification tool/software, patient's disease symptoms are generally learned then this learned knowledge is used for deciding the facts whether patients are suffering from disease or not. Here diagnosis of the disease is performed by applying soft computational techniques for classifying the disease datasets. Medical data classification tool needed and demanded in medical field because it overcomes the problem like difficulties in huge medical data analysis by medical professionals means they hardly processes the large amount of medical data, sometimes they may suffer with fatigue or other tensions resulting classification error. So they need a tool/ expert software that can help them to make a justified decision. Aim of our proposed model is to develop a robust medical data classification system for the numerical type medical dataset.

In our model we are using Moth Flame Optimization (MFO) for feature selection and Fuzzy Min Max Neural Network (FMMNN) for Classification of disease datasets.

MFO algorithm can be used for selecting best features hence it can be used with classification algorithms to get efficient result [9]. MFO is a genetic algorithm which is inspired from nature. It discovers optimal features to be supplied to data sets classifier. It simulates the behavior of the moths moving towards light source [20]. FMMNN is a supervised classifier. It is inspired from neural structure of the human brain and capable to learn data, accordingly conclude intelligent decisions. FMMNN is a bundle comprising fuzzy logic and min max neural network. FMMNN consist three layers. First (Input) layer contains the number of input nodes which accept input features of supplied datasets. Based on the input features, Different logically separated regions are created by middle layer or hyperbox layer. Output layer comprises the number of nodes equals to the number of output classes [18].

In our proposed model, MFO generates optimal set of features to be supplied to classifier FMMNN for achieving classification accuracy with less error.

II. REVIEW OF LITERATURE

Optimization algorithms are assorted into a number of candidate solutions and dependency criteria. Here number of candidate solutions are divided in two groups i.e. group based on Individual and group based on population. Simulated Annealing and Hill Climbing are some of the example of Individual based algorithm. Local optima is the basic problem with these algorithms. Brain Storm Optimization, Particle Swarm Optimization and Moth Flame Optimization (MFO) belongs to the category of population based and have the high capability to avoid problem of local optima [21]. Evolution, Physic and Swarm are three categories which comes under population based feature selection algorithms. With respect to the nature, evolutionary algorithm simulates its evolutionary process. One instance of the evolutionary computation is the Bio Geography Based Optimization (BBO) algorithm [6]. Gravitational Search Algorithm is inspired from physical phenomena in nature [7]. Social behavior of animal species like ant and birds, impacted swarm based algorithms. Some of its examples are Bees Algorithm [5], Grey Wolf Optimizer [21] and Moth Flame Optimization [21].

Two conflicting milestones i.e. Discovery and Exploitation should be balanced to give optimum result in optimization algorithm. Population based algorithms have the potential to

balance discovery and exploitation in order to first find a reasonable approximation of the global optimum then boost its accuracy[2].

Optimization algorithms are sectioned into a filter based and wrapper based feature selection, based on its dependency criteria. In filter based perspective, data dependent criteria is used for getting possible solutions in the feature space whereas classification dependent criteria is used in wrapper based perspective for getting solutions in the feature space[8].

MFO is the one of the approach belongs to category of population and wrapper based feature selection. MFO algorithm finds very competitive outcomes compared with other well-known meta-heuristic algorithms such Particle Swarm Optimization and Biogeography based algorithm. MFO algorithm diversification is very high consequently it requires local optima to be avoided. Its Diversification and intensification balance is very efficient in seeking the right solution to address real issues [21].

A structure can be designed to establish the connectivity between optimization algorithm and the problem to get optimum solution [4]. Many researchers already worked with classifier like Neural Network, Fuzzy Logic, Hybrid Neurofuzzy model, support vector machine, principal component analysis, Artificial Immune System and Genetic Algorithm. Rüdiger W. Brause (2001) presented disease diagnosis model by the use of growing neural networks and rule based networks, he detailed neural network applicability to classify medical problems and showed its merits and demerits with regards to the medical background. He concluded that human diagnostic ability are worse than the neural network diagnostic systems [19] Lotfi Jadeh worked on fuzzy set theory and in 1974 presented Fuzzy logic and its application to approximate reasoning and stated that fuzzy logic is capable for making rule based intelligent system[12].During the year 1992, P. K. Simpson proposed two separate neural network training approaches: one for the problem of clustering and another for the problem of classification, Fuzzy min max neural networks were presented for their experiments [18]. Work on "Least square support vector machine with fuzzy weighing pre-processing" presented by E. Comak and achieved accuracy was 94.29% for the liver disorders problem [3]. It is noted that SVM does not accomplish the classification task in the non-linear case but in linear case SVM classifies the data. To support nonlinearity data needs to be converted in linear feature space by using different compatible kernels. SVM is a statistical learning theory and offline mechanism to develop medical decision support system. Lukka et al. suggested a 66.50 percent classification accuracy based on fuzzy robust PCA algorithms and similarity classifier for liver disorders [13]. Here, data has different point scales. So firstly there is need to normalize the data then it is to be applied to PCA. PCA is used for dimensionality reduction and it is not influenced by variables which are in high magnitude. Seral Ozens et al. experimented combination of genetic algorithm and artificial immune system for heart and liver disorders" [17]. The key downside here is that there is no classification bias for many AIS classifier. The pure distance criterion required to determine the degree of affinity. So AIS is associated with genetic algorithms in this approach to handle above mentioned

problem. It is point to be noted that GA uses probabilistic rules to guide searching. Manjeevan seera et al. (2014) [15]", used "Hybrid Intelligent system for medical data classification (Fuzzy Min Max-CART-RF)". They experimented FMM, FMM-CART and FMM-CART-RF for getting classification accuracies from the labeled dataset of liver disorders and pima Indian diabetes. In the classification stage here, test samples fall between various groups in the overlapping/ and/or containment regions. The authors used the contraction strategy to solve this problem. In contraction approach, principle of minimal disturbance is used to avoid the problem of overlapped and/or containment regions. Here some amount of misclassification arises because the use of principle of minimal disturbance. Lukka (2011) [14], gave the model "Feature Selection using fuzzy entropy measures with similarity classifier" for evaluating classification accuracy for Pima Indian Diabetes. Here feature selection is applied on high dimensional medical dataset. Proposed algorithm is used to select the appropriate and optimal features but forgotten to increase the accuracy constraint of the classifier. The feature selection algorithm which supports both binary dataset and the multi class dataset sometimes produces high accuracy on the binary dataset but gives low accuracy when it is used in the multi class data set. Orkcu & Bal (2011) [16], used Binary Coded GA, BP and Real Coded GA for the disease dataset of Pima Indian diabetes. Highest 77.10 % accuracy was achieved in the consideration of Real/Actual Coded Genetic Algorithm. The actual Coded and Binary Coded Genetic Algorithm operated explicitly with character strings representing the parameter collection and not the parameters themselves. It also uses probabilistic rules not deterministic rules to guide their search.

Zhiying Xu et al. (2020) presented a soft computational model for diagnosing skin cancer. Here convolution neural network is optimized by the satin bowerbird optimization (SBO) algorithm hence overall classification accuracy enhanced when optimal features are classified by Support vector machine. As performance metric was concerned they achieved 95% accuracy [23]. Navid Razmjooy et al discussed implantation capabilities based model for the automatic skin cancer detection [26]. As per the study of authors Ana Carolina Borges Monteiro et al., Software Defined Radio embedded with Wireless Body Area Network system proved beneficial for doctors to keep an eye on patient. As this technology can measures level of blood pressure, temperature etc. This type of technological advancement proved very important in healthcare field [24]. Benefits of Health 4.0 cyber physical system (HCPS), elaborated by Ana Carolina Borges Monteiro et al. in healthcare system. The HCPS observes parameters of patient's condition using biosensors [25]. Study of related articles helped to motivate for the research in the medical diagnosis field. Advancement in medical diagnosis with optimized performance needed to equip with latest technology to maintain smart healthcare system.

As main findings are concerned, FMMNN suffers from problem hyperboxes overlapping and containment. For solving this, contraction approach is used. It is observed that some misclassification still exists here if too many features are handled while classification process [1]. MFO generate

selected optimal features to be supplied to FMMNN. Hence overlapping and containment cases between hyper boxes in FMMNN reduced. Consequently chances of misclassification minimized. This is the main motivation of using combination of FMMNN and MFO in our model.

III. Moth Flame Optimization (MFO)

Moths navigate in a fixed angle targeting the moon. In night they are unable to distinguish between moon and artificial light or flame. For movement they follow transverse orientation navigation strategy also maintains a similar angle with flame and the moon [21]. Compared to the moon, flame remains highly closed to the moths so maintaining a similar angle triggers spiral navigation route for moths [10] hence convergence towards flames occurred.

A. MFO Algorithm

It is assumed that moths are the matrix of possible solutions and their layout in the space are the problem variables. The best obtained layout (optimal position) of moths is kept in matrix of flames [21]. Following paragraphs shows the overview of steps of MFO algorithm.

Step 1: Required main Parameters: count of moths (search agents), count of flames, count of variables, count of maximum iterations, lower bound (lb (i)) and upper bound (ub (i)) of variables.

Step 2: Representation of Population: MFO algorithm can be modelled by the position of moths and flames. The MP matrix represents the position matrix of moths.

$$MP = \begin{pmatrix} MP_{1,1} & MP_{1,2} & .... MP_{1,d} \\ MP_{2,1} & MP_{2,2} & .... MP_{2,d} \\ MP_{n,1} & MP_{n,2} & .... MP_{n,d} \end{pmatrix} \qquad (1)$$

here n is the count of moths and d is the count of variables (dimensions). $MP_{i,j}$ Shows the value of jth position variable of the ith moth, is decided by Eq. (2)

$$MP_{i,j} = (ub(i)-lb(i) * rand(\ ) + lb(i)) \qquad (2)$$

here rand ( ) is the random number generator in the interval [0, 1] having random distribution. FP represents the position matrix of flames as follows:

$$FP = \begin{pmatrix} FP_{1,1} & FP_{1,2} & .... FP_{1,d} \\ FP_{2,1} & FP_{2,2} & .... FP_{2,d} \\ FP_{n,1} & FP_{n,2} & .... FP_{n,d} \end{pmatrix} \qquad (3)$$

$FP_{i,j}$ Shows the value of jth position variable of the ith flame.

In MFO, moths are the search agents that move around the search space and flames are the best position of moths.

Step 3: Persistence of fitness function: Fitness function evaluates each month. FOM matrix stores fitness value of moths processed by fitness function (see section 3.2). Similarly the matrix FOF is used to store the fitness values of flames.

$$FOM = \begin{pmatrix} fom_1 \\ fom_2 \\ fom_n \end{pmatrix} \qquad (4)$$

$$FOF = \begin{pmatrix} fof_1 \\ fof_2 \\ fof_n \end{pmatrix} \qquad (5)$$

In MFO, Flames are treated as flags dropped by moths while searching the search space. In the event of seeking a better solution, each moth looks for a flame and updates its position .Flame matrix updates their fitness values as per the fitness values of moths.

Step 4: General texture of MFO and iteration process: - The general texture of MFO algorithm is as follows

$$MFO = (K,R,T) \qquad (6)$$

where, K is the function that generate a random population of moths and corresponding fitness, R is the main function that makes the moths move around search space, T is a termination function. Function R uses logarithmic spiral, it is main update mechanism for moths to simulate orientation behaviour of moth. Eq. (7) update location value of moth with respect to the flame.

$$MP_i = S(MP_i, FP_j) = D_i . e^{bt} . \cos(2 \Pi t) + F_j \qquad (7)$$

where, the spiral function is S, MP_i is the ith moth, and FP_j is the jth flame, D_i defines the distance between the ith moth and the jth flame for defining the shape of logarithmic spiral. b is a constant and t is a random number in the range [-1, 1]. D_i is measured using following Eq. (8)

$$D_i = |FP_j - MP_i| \qquad (8)$$

Spiral function $S(MP_i, FP_j)$ decides how moths update their position around flame. Eq. (7) allows a moth to fly around a flame only. Hence the exploration/discovery and exploitation of the search area are therefore ensured. In order to emphasize exploitation further it is believed that t is a random number in range [r, 1] where r is the constant of convergence. Its values linearly decreased from -1 to -2 over the course of iteration. But exploitation of the best solutions may degrade due to update of moth's position with regard to n various positions in the search space. Because of this reason, MFO can fall in local minima. To solve this problem Eq. (9) is used to reduce the quantity of flames over the course of iterations. Hence exploitation degradation achieved and best promising solution obtained.

$$\text{Flame count} = \text{round } (NF - c \times \frac{NF - 1}{T}) \qquad (9)$$

where, NF is the maximum count of flames, c is the current iteration, and T is the maximum count of iterations.

Step 5: Obtaining optimal solution: - General steps of R function are as follows:

    Update flame count using Eq. (9)

    FOM = FitnessFunction (MP);

    If iteration == 1

        FP = sort (MP);

        FOF = sort (FOM);

    else

        FP = sort(MPt-1, MPt);

        FOF = sort(MPt-1, MPt);

    end

    for i = 1: n

        for j = 1: d

            Update r and t

            Calculate D using Eq. (8) with

            respect to the corresponding moth.

            Update MP (i,j) using Eq. (7) with

            respect to the corresponding moth.

        end

    end

R function is executed until T function i.e. T: MP → {true, false} returns true. After termination of the R function the best moth is returned as the best obtained approximation of the optimum.

B. Objective Function (Fitness Function) used in MFO

The aim of the optimization has to improve the classification accuracy of the classifier by selecting minimum set of the attributes. This improves the accuracy of the system with improved speed of classification which is achieved by the reduction of attributes using MFO method. Let $[X_1 \ldots \ldots X_n]$ are n attributes of data sample (disease data sample). If k attributes selected out of n then optimization function have k variables $[X_1 \ldots \ldots X_k]$ Range of each variable is as follows.

$1 <= g_k <= n$ where constraint is $g_i \neq g_j$ for all value of i and j.

In this way we select the reduced attribute set as $A_R = [X_{g_1} \ldots \ldots X_{g_k}]$ where A_R is set of reduced attributes. Following equation represent the main motive of the objective function. It is formed as to reduce classification error by FMMNN classifier for A_R.

minimized classification error $= \text{FMMNN}(A_R, T)$ (10)

where T is the target output.

FMMNN classifier consist three layers. In the first (Input) layer, positions of moths are supplied as the parameters. In the middle layer, formations of hyperboxes (logically separated regions as per its class) take place. Its output layer gives the actual classified output. Then classification error is calculated based on the comparison between target output and actual output. Section 4 gives the more details about FMMNN. In MFO, values of optimal (efficient and reduced) set of attributes (where minimum classification error is obtained) considered as fitness value of moths.

IV. FUZZY MIN MAX NEURAL NETWORK (FMMNN)

F_A, F_B and F_c are the three consecutive layers of FMMNN Fig. 1, where F_A represents input layer, F_B acts as hidden or hyper box layer and F_c is the representative output layer which contains class nodes. In hyper box layer, a hyperbox is regarded as a fuzzy set. The minimum and maximum points of the hyper boxes are denoted by matrices V and W respectively, matrix U gives relation between hyperbox layer and output layer.

A. Transition from Input Layer to Hyperbox Layer

FMMNN uses membership function as given in eq. (11) to calculate input features belongingness for the hyper boxes

$$b_j = (a_j, V_j, W_j) = \frac{1}{n} \sum_{i=1}^{n} (1 - f(a_{h_i} - W_{j_i}, \gamma) - f(V_{j_i} - X_{h_i}, \gamma)) \ (11)$$

$f_{x,y} = 1$ if x*y>1

$f_{x,y} = x*y$ if 0<=x*y<=1

$f_{x,y} = 0$ if x*y<0

here a_h = Input features, n = dimension of feature space, b_j = j^{th} hyper box, V_j = set of min points, W_j = set of max points, f(x, y) is a two parameter function where γ is sensitivity(fuzziness control) parameter. The connections between input layer and hyperbox layer maintained by v and w matrix.

In hyper box layer, if training performed 1st time creates a hyper box using $V_j = W_j = A_h$ else expand existing hyperbox b_j using new training data if the criteria given in Eq. (12) satisfied:

$$\theta \geq \sum_{i=1}^{n} (\max(w_{j_i}, a_{h_i}) - \min(v_{j_i}, a_{h_i})) \tag{12}$$
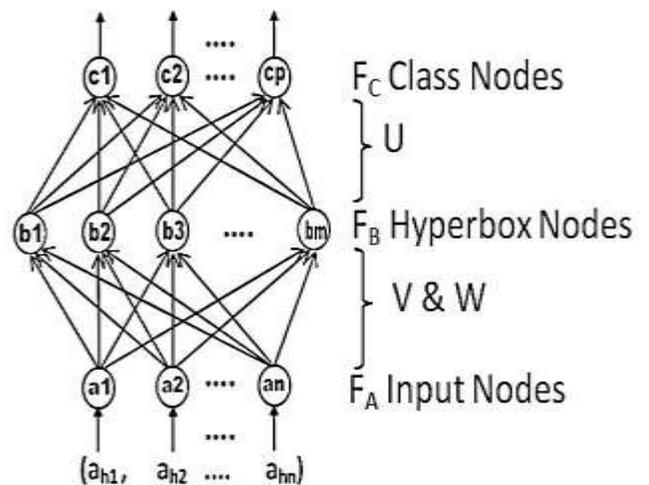
here θ = Expansion coefficient

Fig. 1. Layers of Fuzzy Min Max Neural Network [1].

When a training data is given to the network then it try to accommodate in one of the existing hyperbox of that class provided that the hyperbox size is not exceeding the specified maximum limit. To know the overlapping or containment between hyperboxes following tests are to be done. Assume that hyperbox b_j and b_k are compared with each other then

a) Isolation test: If $\{W_{j_i} < V_{k_i}\}$ or $\{W_{k_i} < V_{j_i}\}$ true for any value of i, it shows that two hyperboxes (b_k, b_j) are isolated.

b) Containment test: If $\{V_{j_i} < V_{k_i} < W_{k_i} < W_{j_i}\}$ or $\{V_{k_i} < V_{j_i} < W_{j_i} < W_{k_i}\}$ true then b_j is contained by b_k or b_k is contained by b_j.

c) Overlap test: if test (a) and test (b) are not satisfied then overlapping between b_k and b_j is obvious.

We used contraction approach to remove hyperboxes overlapping and containment problem. In this, if the hyperboxes from different classes overlap, a hyperbox contraction process is initiated to eliminate the overlaping regions. Note that overlapping regions caused by hyperboxes from the same class are allowed [1].

B. Transition from Hyperbox Layer to Output Layer

Connection between hyperbox node and class nodes in is represented by u matrix.

$$u_{ij} = 1 \quad if \{b_j \varepsilon c_i\} \text{ otherwise } u_{ij} = 0$$

By knowing the hyperboxes belongingness to the class, network easily classify the linearly non separable data. c_i decides the membership of ith class and it is given by

$$c_i = \max_{j=1}^{m} (b_j u_{ij}) \tag{13}$$

V. METHODOLOGY OF MFO-FMMNN

Our proposed hybrid model consist two parts. First part handled by Moth Flame Optimization phase (MFO) while the second part took over by the classification phase. Methodology structure of hybrid model is shown in Fig. 2. Methodology steps are described below:

1) Input: All features of medical (disease) data samples are supplied to MFO-FMMNN.

2) Settings of parameters: Following parameters need to be set in this step. (i) count of moths(search agents) (ii) count

of flames (iii) count of variables (iv) count of maximum iteration (v) vector of Lower bound vector (vi) vector of upper bound. (3) Setting of population of moths and count of flames:

3) Origin of moth and flame count: Set the first random population of moth, obtained from the Eq. (2). Count of flames is calculated using Eq. (9).

4) Inception of optimization:

a) MFO started with iteration value 1.here iteration is the count of loops used to test algorithm.

b) The fitness values of all the moths calculated using objective function (see Section IIIB, the values which produced minimum error in its extent when medical features processed by FMMNN as fitness f unction is concern)

c) Sorting of the first population of moths takes place based on their fitness values.

d) The population with the best moths fitness values will be selected as the flames hence the fitness values of flames and moths equals

e) Updating of moths position takes place with respect to its corresponding flame, calculated according to Eq. (7).

5) Culmination of optimization:

a) When the first iteration terminates we get the best set of moth and flames and the best fittest values of moths and flames.

b) Start to generate the offspring generation of moths (other subset of features) then update the position and fitness values of moths and flames at each subsequent iterations according to (Eq. (1) to (5) and (7))

c) In each iteration the sequence of flames are Modified based on the best solutions, then moths update their position with respect to the updated flame.

d) When iteration value reaches its maximum limit, we get the new best set of moths and flames and also their corresponding fitness value. Now optimization process stops and the best moths (optimal set of features) obtained.

6) Classification: Using optimal set of features, FMMNN classifier calculates classification accuracy.

a) In MFO phase, FMMNN runs iteratively with regards to obtain fitness values of moths as per iterations.

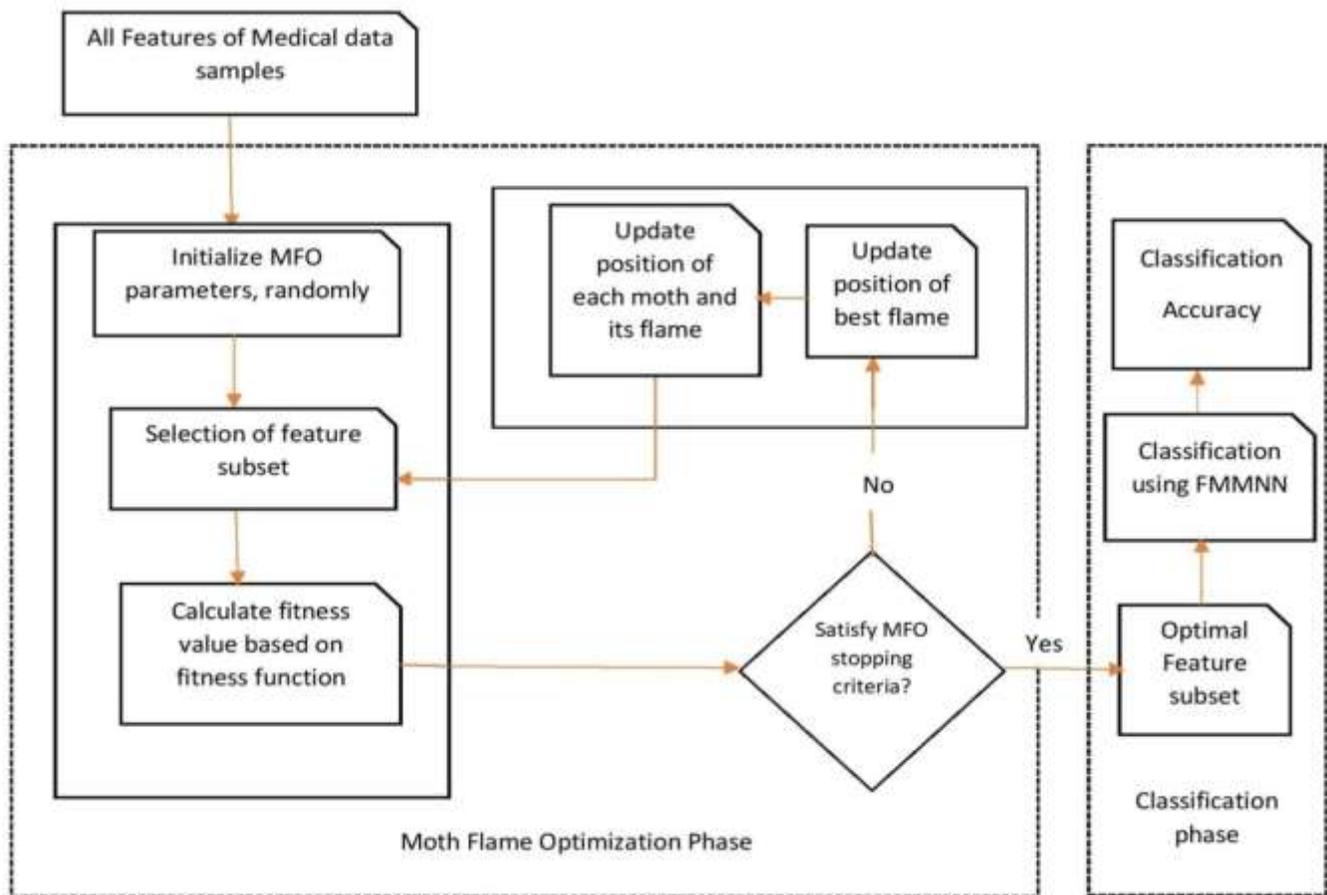*b)* In classification phase, FMMNN runs to obtain classification accuracy.

Fig. 2. Methodology Structure of MFO-FMMNN.

VI. EXPERIMENTS AND RESULT ANALYSIS

Our experiments and results are based on data samples of Liver Disorders and Pima Indian Diabetes. Following paragraphs shows the deliberation.

A. Experiment on Liver Disorders Dataset

Publicly available data sets of Liver Disorders taken from the University of California at Irvine (UCI), machine learning data repository (This repository consist genuine database and can be considered for research [11]). It is used for evaluation of medical decision support system. In the dataset following feature values of the patient are considered i.e. Age, Gender , Bilirubin Total, Bilirubin Direct, Alkaline Phosphotase, Alamine Aminotransferase, Aspartate Aminotransferase, Protiens Total, Albumin, Albumin and Globulin Ratio(Lichman et al. , 2013). A series of systematic evaluations was conducted for 1 to 10 features for 50 Maximum count of iterations. Fig. 3 shows result of percentage of classification accuracy evaluated by MFO-FMMNN. When 5 features are considered, achieved accuracy 97.74 %. This result is compared with the result calculated by FMMNN classifier then it found that MFO-FMMNN proves better than FMMNN as per accuracy result shown in Table I.
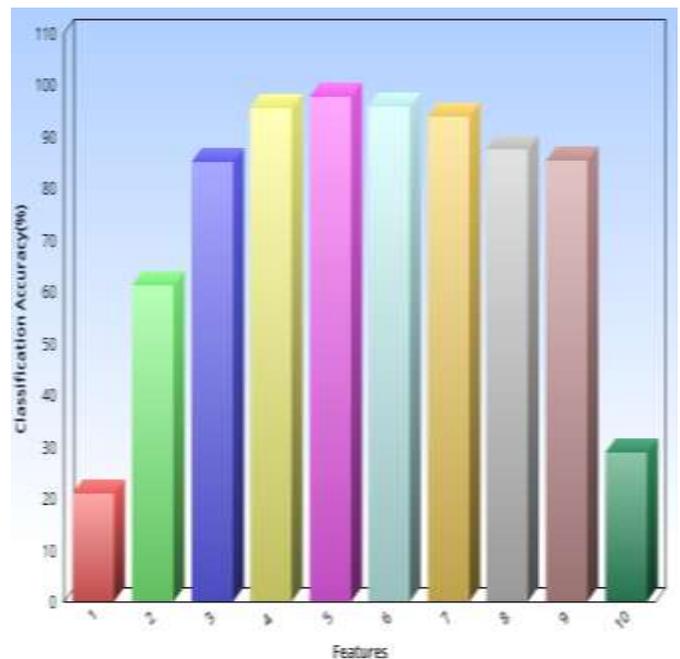

Fig. 3. Features vs. Classification Accuracy Calculated by MFO-FMMNN for Liver Disorders Dataset.

TABLE I.   COMPARISON OF CLASSIFICATION ACCURACY BETWEEN FMMNN AND MFO-FMMNN FOR LIVER DISORDERS DATASET

SN	Model/methods used	Accuracy achieved
1	FMMNN	68.80
2	MFO+FMMNN	97.74

B. Experiment on Pima Indian Diabetes Dataset

This dataset is taken from the National Institute of Diabetes and Digestive and Kidney Diseases. The objective of the dataset is to classify whether or not a patient suffered from diabetes [22]. Patients data considered here are of females at least 21 years old of Pima Indian heritage. Total eight features are considered in each data sample i.e. patient age, plasma glucose concentration a 2 h in an oral glucose tolerance test, blood pressure (diastolic), triceps skin fold thickness, 2-h serum insulin, index of body mass, diabetes pedigree function of diabetes, times of number of pregnant. When PID dataset is processed by MFO-FMMNN for 70 Maximum numbers of iterations then calculated classification accuracy is 86.95 % (see Fig. 4). Our output is compared with the classification result calculated by previous author's proposed work. Results are summarized in Table II, which represents the comparison of classification accuracy between FMMNN and MFO-FMMNN for Pima Indian diabetes dataset. Accuracy achieved by MFO-FMMNN super shaded FMNNN.
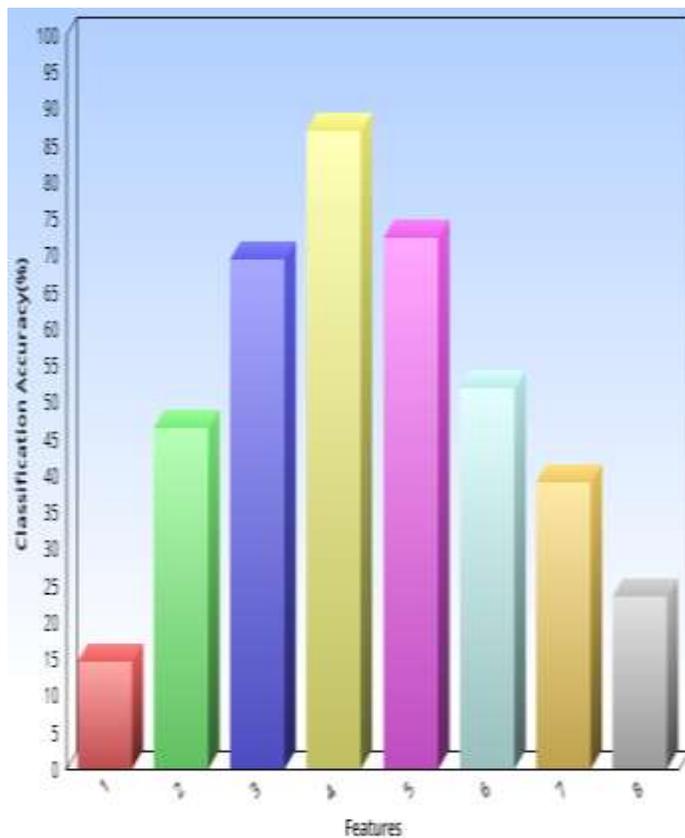

Fig. 4.   Features vs. Classification Accuracy (%) Calculated by MFO-FMMNN for Pima Indian Diabetes Data Sample.

TABLE II.   COMPARISON OF CLASSIFICATION ACCURACY BETWEEN FMMNN AND MFO-FMMNN FOR PIMA INDIAN DIABETES DATASET

SN	Model/methods used	Accuracy achieved
1	Fuzzy Min Max Neural Network	70.40
2	MFO+FMMNN	86.95

VII. CONCLUSION

MFO-FMMNN can solve difficult problems in the real world with restricted search spaces. MFO balances discovery and exploitation of search spaces hence able to avoid local minima problem. This is the main reason behind getting optimal features set. FMMNN used in this paper fulfilled two purposes i.e. calculating fitness value of features (Moths) in MFO phase and classify the data in single pass classification phase. It has the property of incremental learning such that newly introduced data can be classified easily. Presented hybrid model is able to calculate optimal accuracy by considering different sets of features of supplied dataset hence best accuracy from best set of features guaranteed. The limitation of this model is that MFO method is not capable to generate solutions for problems having more than one objective. Decision trees can be used to provide explanatory rules of working of this hybrid model. In future this model can also be implemented and tested for image basis medical datasets if wavelet transforms accompanied with it.

REFERENCES

[1] A.V. Nandekar and P.K. Biswas, "A Fuzzy Min-Max Neural Network Classifier With compensatory Neuron Architecture," IEEE TRANSACTIONS on neural networks, vol. 18, no. 1, January 2007.

[2] A.E. Eiben, C. Schippers, On evolutionary exploration and exploitation, Fundam. Inform. 35(1998) 35–50.

[3] E. Comak et al (2007). A new medical decision making system: least square support vector machine (LSSVM) with fuzzy weighing preprocessing, Expert system with applications 32(2), 409-414.

[4] D.H. Wolpert, W.G. Macready, No free lunch theorems for optimization, IEEE Trans. Evol.Comput. 1 (1997) 67–82.

[5] D. Pham, A. Ghanbarzadeh, E. Koc, S. Otri, S. Rahim, M. Zaidi,The bees algorithm-a novel tool for complex optimisation problems, IPROMS 2006, 2006, pp. 454–459.

[6] D. Simon, Biogeography-based optimization, IEEE Trans. Evol. Comput. 12 (2008) 702–713. (2014) 46–61.

[7] E. Rashedi, H. Nezamabadi-Pour, S. Saryazdi, GSA: a gravitational search algorithm, Inf. Sci. 179 (2009) 2232–2248.

[8] I. Guyon, A. Elisseeff, "An introduction to variable and attribute selection", Machine Learning Research,Vol.3, pp. 1157-1182, 2003.

[9] Hossam M. Zawba et. al, "Feature Selection approach based on moth flame optimization algorithm" 978-15090-0623-6/16/$31.00©2016 IEEE

[10] K.D. Frank, C. Rich, T. Longcore, "Effects of artificial night lighting on moths", Journal of Ecological consequences of artificial night lighting, Hindawi Publishing Corporation, Vol. 2015, No. 1,pp. 305-344, 2006.

[11] Lichman, M. (2013). UCI Machine Learning Repository Irvine, CA: University of California, School of Information and Computer Science

[12] Lotfi jadeh(1974). "Fuzzy logic and its application to approximate reasoning". In: Information Processing 74, Proc. IFIP Congr. 1974 (3), pp. 591–594.

[13] Luukka at al.(2009). Classification Based on Fuzzy Robust PCA algorithms and similarity Classifier. Expert Systems with Applications, Elsevier, 5 March 2008.

[14] Luuka P.(2011) Feature selection using fuzzy entropy measure with similarity classifier. Expert system with application.38 (4), 4600-4607.

[15] Manjeevan Seera ,CheePeng Lim. "A hybrid intelligent system for medical data classification", ELSEVIER, Expert System with application 41(2014) 2239-2249.

[16] Orkcu H. H. & Bal H.(2011) Comparing performance of backpropagation and genetic algorithms in data classification. Expert system with applications 38(4),3703-3709.

[17] Ozens S & Gunes (2009). Attribute weighing via genetic algorithm for attribute weighted artificial immune system (AWAIS) and its application to heart disease and liver disorders problems. Expert system with applications, 36(1), 386,392.

[18] P. K. Simpson, "Fuzzy min-max neural networks –Part 1: Classification", IEEE Trans. Neural Networks, Vol.3, No.5, pp. 776-786, 1992.

[19] R. W. Brause, "Medical Analysis and Diagnosis by Neural Networks", J.W. Goethe - University,Computer Science Dept., Frankfurt a. M.,Germany,2001.

[20] S. Mirjalili," Moth - Flame Optimization Algorithm: A Novel Nature inspired Heuristic Paradigm", Knowledge-Based Systems, Vol. 89, pp.228-249, 2015.

[21] S. Mirjalili, S.M. Mirjalili, A. Lewis, Grey wolf optimizer, Adv. Eng. Softw. 69.

[22] Smith, J.W., Everhart, J.E., Dickson, W.C., Knowler, W.C., & Johannes, R.S. (1988). Using the ADAP learning algorithm to forecast the onset of diabetes mellitus. In Proceedings of the Symposium on Computer Applications and Medical Care (pp. 261--265). IEEE Computer Society Press.

[23] Zhiying Xu*, Fatima Rashid Sheykhahmad, Noradin Ghadimi*, Navid Razmjooy (2020). Computer-aided diagnosis of skin cancer based on soft computing techniques. Open Medicine, 15(1), 860- 871.

[24] Ana Carolina Borges Monteiro , Vania V. Estrela , Abdeldjalil Khelassi , Yuzo Iano, Navid Razmjooy , Diego T. M. Rocha , Delcimar Martins (2019). Why software-defined radio (SDR) matters in healthcare? Med. Technol. J. 3 (3), 421–429.

[25] Ana Carolina Borges Monteiro , Vania V. Estrela , Abdeldjalil Khelassi , Yuzo Iano, Navid Razmjooy. (2019) Health 4.0: Applications management technologies and review type of article. Med. Technol. J., 2(2), 262-276

[26] Navin Razmjooy, Mohsen Ashourian, Maryam Karimifard, Vania V. Estrela, Hermes J. Loschi, Douglas do Nascimento, Renaldo P. Franca, Mikhail Vishnevski(2020), Computer-added Diagnosis of skin cancer: a review.Current medical Imaging, 16(7), 781-793.