

Modified Monarch Butterfly Based Feature Selection for Multi Medical Data Classification Using Deep Neural Network

N. Balakumar, B. Prabadevi

Abstract: *In healthcare informatics, the individual's disease prediction and its classification are essential. With the utilization of data mining systems, we can analyze the disease at a beginning stage and enhances the patient's survival rate. But still, it has some issues like removing missing values and feature selection from the medical datasets. To overcome that, optimal features are selected from the datasets by the use of an innovative optimization algorithm. In the proposed work, multi-datasets (Liver, Lung, Heart, and Thyroid) are considered for the disease prediction analysis. Initially, the missing values from the input datasets are removed during the preprocessing stage. Next, to that, the optimal features are chosen by the algorithm called Modified Monarch Butterfly Optimization (MMBO). Based on the selected features, the data are classified into two norms: healthy and non-healthy by the proposed classifier i.e. Deep Neural Network (DNN). The performances of the proposed algorithm and classifier are tested on the multi-datasets in terms of sensitivity, specificity, and accuracy. The results demonstrate the MMBO-DNN algorithm achieves high accuracy of and less execution time compared to existing algorithms.*

Index Terms: *Medical Data Classification, Missing values, Optimal Feature Selection, MMBO, and DNN Classifier.*

I. INTRODUCTION

As of late, data classification is seen as a most quickly making possibility approach in the field of data mining, with various applications in every aspect of learning [1]. In the medical region, scientists have been investigating various data classification methods to enhance the accuracy of medical diagnosis [2]. For the most part, a medical database contains a huge gathering of medical data, which incorporates some parameters, every one of the parameters characterizes a specific kind of disease [3, 4]. Medical data classification contains extraordinary significance in the field of medicine in view of the gigantic improvement of robotization strategies [5]. By applying data classification algorithms on medical data, it is conceivable to help diagnosis of diseases and particularly in its beginning times [6]. When it gets effectively distinguished in beginning periods itself, the recovery of patient's life is high.

Normally, the patient's data consists of different features and its diagnostics process should be entered with the most extreme care to give services [7]. As the data stored in medical databases contain missing values and redundant data, mining of the medical data becomes cumbersome [8]. As it can influence the outcomes of mining, it is fundamental to have great data preparation and data reduction before applying data mining algorithms [9]. Forecast of disease turns out to be speedy and simpler if data is exact and reliable and free from noise [10, 12]. Feature selection is a proficient data preprocessing method in data digging for diminishing dimensionality of data. In medical diagnosis, it is essential to distinguish the most significant risk factors related to disease [11]. The advantage of feature identification from the datasets is it provides quick and better results by removing unwanted attributes [13].

One of the prediction techniques in data mining is classification which first uses training data to build up a model after that the final model is applied on testing data to get results of prediction [14]. Different classification algorithms have been connected on disease datasets for the diagnosis of various diseases from the medical datasets and the outcomes have been observed to be exceptionally encouraging [15, 16]. There is a most extreme need to build up a novel classification method which can speed up and rearrange the procedure of diagnosis of disease [17]. In this way, feature selection decreases the classifier capacities to the measurements that are regarded most applicable for exact classification [18]. Henceforth, the research utilizes distinctive data mining systems to classify medical data and furthermore to explore the effectiveness of utilizing diverse input features and classifiers for evaluating the medical datasets.

This paper is organized as follows: A survey on various feature selection approaches and classification techniques in medical data classification are presented in section 2. Section 3 describes the purpose of medical data classification; Section 4 gives a detailed description of the methodology along with the proposed techniques. Section 5 explains the classification results of the proposed work and the conclusion part is presented in Section 6 along with perspectives for this work.

Revised Manuscript Received on December 22, 2018

N. Balakumar, Department of Computer Applications, Pioneer College of Arts and Science, Coimbatore, India. Email: msg2balakumar@gmail.com ,

B. Prabadevi, Department of Computer Applications, Pioneer College of Arts and Science, Coimbatore, India. prabadevi.b@gmail.com



II. EXISTING LITERATURE-A SURVEY

In 2018 Jain, D., and Singh, V. [19] presented a study on the use of feature selection as well as classification procedures for the diagnosis and prediction of chronic diseases. Dimensionality decrease helps in enhancing the execution of machine learning algorithm. This work shows an overview of various feature selection techniques and their inalienable advantages and disadvantages. At that point, they investigated conventionally and parallel classification systems for disease forecast.

Shuo Yang *et al.* [20] had exhibited an enhanced classification approach for the prediction of diseases based on the established Iterative Dichotomiser (Id3) algorithm. The enhanced Id3 algorithm defeats multi-esteem predisposition issue while choosing test/split qualities, comprehends the issue of numeric trait discretization and stores the classifier as guidelines by utilizing a heuristic methodology for simple comprehension and memory reserve funds. Analysis of results demonstrated that the enhanced Id3 calculation is better than the present four classification algorithms in terms of accuracy, stability, and minor error rate.

A Computer Aided Diagnosis (CAD) scheme for the location of breast disease has been created by Ahmed Abdel-Zaheret *al.* [21] utilizing Deep Belief Network (DBN) unsupervised way taken after by back propagation supervised path. The development was back-propagation Neural Network (NN) with Liebenberg Marquardt learning capacity while weights were instated from the DBN-NN. Our system was tried on the Wisconsin Breast Cancer Dataset (WBCD). The classifier complex gives a precision of 99.68% demonstrating promising outcomes over beforehand distributed examinations.

In 2018 Ghaddar, B., and Naoum-Sawaya, J. [22] proposed a classifier called support vector machine (SVM) for solving the problem of attributes or feature selection. by using this SVM approach, presented a generic approach for cancer classification based on gene expression to the diagnose the tumor disease. the results demonstrated the SVM approach accomplished low error rates compared to other algorithms.

In 2011 Lavanya, D. and, Rani, [23] examined the execution of Decision tree classifier-CART with and without feature selection as far as accuracy, time to fabricate a model and size of the tree on different Breast Cancer Datasets. The outcomes demonstrated that a specific feature selection utilizing CART has improved the classification precision of a specific dataset.

[24] Shen, L., *et al.* 2016 proposed a novel approach called support vector machine with fruit fly optimization algorithm (SVM_FOA) for diagnosing the diseases from the medical datasets. The author considered four medical datasets for analyzing the individual's criteria. The performance of the SVM_FOA was analyzed on the four datasets in terms of accuracy, sensitivity, specificity and computation time.

III. PURPOSE OF MEDICAL DATA CLASSIFICATION

Classification of the continuous data with a large number of records and features has become the most challenging data mining tasks of human activity. In the last 20 years, it is being applied in the field of medical data

classification for early diagnosis of individual's diseases. For that, a specific classifier is used with an optimization algorithm which improves the data classification accuracy [29-31].

IV. METHODOLOGY FOR DATA CLASSIFICATION

The methodology aims to improve the medical data classification accuracy and to provide the preventive measures to the patients at an early stage [32]. The study comprises four medical datasets namely, liver, lung, heart, and thyroid. Each dataset has a number of attributes, from that some specific attributes (it specifies the type of disease) will be used to classify the data [33, 34]. The stages involved in medical data classification are

- Removing missing values
- Optimal feature (attribute) selection
- Data classification

Initially, the missing values and redundant data are removed during the preprocessing stage this will lead to high-quality results as well as reduced costs for data mining. In the next stage, the optimal features are selected from the datasets using an optimization algorithm called MMBO. Based on the selected optimal features, the data are classified corresponding to the specific diseases. For the classification purpose, DNN classifier is used to identify the disease from the datasets as in two categories (healthy and non-healthy). The schematic diagram of the proposed medical data classification is shown in figure 1.

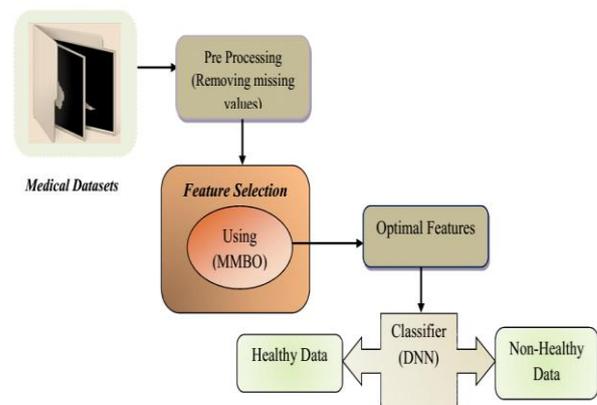


Fig 1: Schematic Diagram of the proposed Medical Data Classification

A. Removing Missing Values from Medical Datasets

The preprocessing step is an important one in data classification; it includes the handling of missing values. In the study, the original dataset has missing values as well as redundant features. The missing values of the medical datasets can be handled in two ways; either by removing the instances or replacing them with mean, average, maximum or minimum of the specified data [25].

B.Optimal Feature Selection using MMBO Algorithm

Processing an extensive number of features in databases is challenging assignment for implementation

procedure. In the data mining process, feature selection is one of the fundamental tasks and it aims to reduce the data’s dimensionality and their noise in datasets. The intention of feature selection is to remove redundant features from the datasets and to keep the relevant data. Feature selection gives a better performance in both the learning and classification process if the possibility of over-fitting increases with the number of features. Here, an MMBO algorithm is presented for selecting the optimal features.

Monarch Butterfly Optimization: Monarch Butterfly Optimization (MBO) is a metaheuristic algorithm which is inspired by the migration behavior of MBO [26]. The migration behavior of MBO is explained as follows:

Rule 1: All the monarch butterflies are only located in Land 1 or Land 2.

Rule 2: By using migration operator in Land 1 or in Land 2 monarch butterfly, each child monarch butterfly individual is generated

Rule 3: In MBO algorithm, if the newly generated MBO has better fitness function as compared to its parent, then it will be replaced by its parent. As a result, the population range is remaining unchanged.

Rule 4: By comparing the generated with parent one, the butterfly with better fitness one is assigned for the next generation and they cannot be altered by any other operators. This will never depreciate with the increment of generations.

Modified Monarch Butterfly: The general monarch butterfly is modified by the random value selection i.e. the value is selected based on the velocity update of Particle Swarm Optimization (PSO) algorithm in the equation (3) and those steps are explained as below.

Initialization: Monarch Butterfly optimizer is initialized with n number of population in land 1 and Land 2; here, each monarch butterflies position representing a given feature set combination and evaluate the fitness according to the position of each MB.

Updating the positions of MMBO

Updating the position of MBO is performed by two operations namely, migration operator and butterfly adjusting operator.

Migration operator: In consideration, the monarch butterflies in Land 1 and Land 2 are named as Subpopulation 1 and Subpopulation 2 respectively. In the study, the features are initialized as the subpopulation 1 and subpopulation 2. This migration process can be explained as follows

$$f_{i,t}^{G+1} = f_{r1,t}^G \dots\dots\dots (1)$$

Two criteria in the migration process:(i) When $r \leq p$ the element t in the newly generated monarch butterfly is generated by (2). Here, r can be computed as

$$r = rand * period \dots\dots\dots (2)$$

In equation (2), the random value is chosen based on the velocity update of Particle Swarm Optimization (PSO) algorithm (modified MBO). Its updation process is explained as in equation (3).

Velocity updation of PSO: Based on the behavior of swarm i.e. PSO algorithm, the velocity vector for a particle is updated according to Gbest and Best value. The formulation for updating the velocity of the particles in the PSO is given as:

$$v_i(t+1) = v_i(t) + b_1 rand(Pbest(t) - r_i(t)) + b_2 rand(Gbest - r_i(t)) \dots\dots\dots (3)$$

Where, V_i is the particle velocity, r_i is the current particle, rand is a random number between (0, 1), b_1, b_2 are the learning factor, usually $b_1 = b_2 = 2$.

(ii) If $r > p$ then the element t in the newly generated monarch butterfly is generated by

$$f_{i,t}^{G+1} = f_{r2,t}^G \dots\dots\dots (4)$$

Parameter Description: From equation (1), $f_{i,t}^{G+1}$

symbolizes the t^{th} element of f_i at generation $G+1$ that introduces the position of the monarch butterfly i . $f_{r1,t}^G$

Indicates the t^{th} element of f_{r1} that is the newly generated position of the monarch butterfly $r1$. G is the current generation process. The term $r1$ is randomly chosen monarch butterfly from subpopulation 1. From equation (2), the period indicates the migration period, rand is a randomly generated number drawn from the uniform distribution. By altering the value of p, the MMBO algorithm can balance the direction of migration operator. The value of p is set to 5/12 as per migration period.

Butterfly Adjusting Operator: The position of MMBO is also updated by another operator i.e. butterfly adjusting operator. Considering all the MMBO elements (medical data attributes) are in j. If ($rand \leq p$), the position can be updated as

$$= \begin{cases} f_{j,t}^{G+1} = f_{best,t}^G, & \text{if } rand \leq p \\ f_{j,t}^{G+1} = f_{r3,t}^G & \text{if } rand > p \end{cases} \dots\dots\dots (5)$$

Under this condition, if $rand > BAR$ (Butterfly Adjusting Rate), it can be further updated as follows:

$$f_{j,t}^{G+1} = f_{j,t}^{G+1} + \alpha \times (dx_t - 0.5) \dots\dots\dots (6)$$

Where BAR indicates butterfly aligning rate. The term α indicates the weighting factor that is afforded as $\alpha = WS_{max} / G^2$ where WS_{max} max walk step that a monarch butterfly individual can move in one step at the current generation G. The parameter dx is the walk step of



the monarch butterfly j that can be computed by performing Levy flight.

$$dx = Levy(f_j^g) \dots\dots\dots (7)$$

With the use of this MMBO algorithm, the optimal features are selected from the medical dataset which reduces the complexity of the classification process. The healthy and non-healthy data from the considered medical datasets are

identified by the presented classification algorithm. The diagrammatic representation of MMBO procedure is depicted in figure 2.

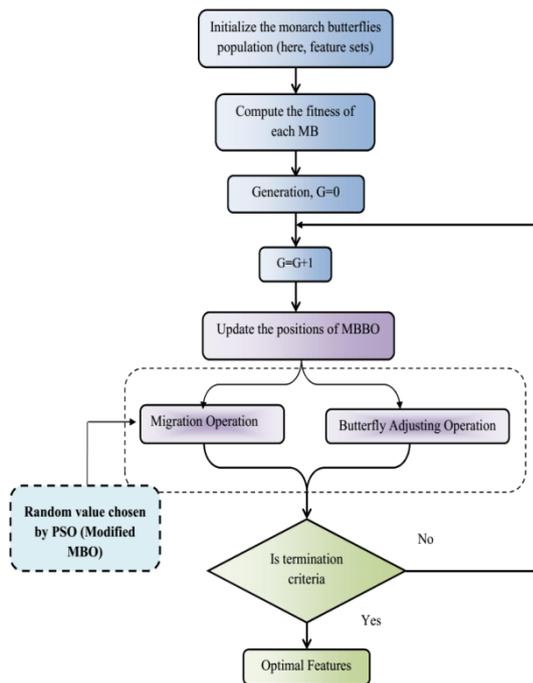


Fig 2: Steps involved in MMBO Algorithm

C. Data Classification

In the field of classification study, Deep Neural Network (DNN) model is used to classify the data as two categories: healthy and non-healthy. The optimal feature set is fed into the classification technique; the attained data with optimal features improve the accuracy of the classification task in comparison of applying the classification task on the original datasets.

Deep Neural Network (DNN)

An artificial neural network model with the multiple layers of the hidden units and outputs is termed as DNNs. In its parameter learning, DNN consists of pre-training (using generative deep belief network or DBN) and fine-tuning stages [27].

(i) Pre-training stage

In the training stage, we use a Deep Belief Network (DBN) that is a deep architecture and typical feedforward network in which the input flows from the input layer to the output layer through a number of hidden layers which are in excess of two layers. The DBN model permits the network to produce visible activations on the basis of its hidden units' states, which characterizes the network belief. Here, we implemented the Restricted Boltzmann Machine (RBM) to work out the above problem.

Restricted Boltzmann Machine: An RBM is an exclusive type of Markov random field that has one layer of (typically Bernoulli) stochastic hidden units and one layer of (typically Bernoulli or Gaussian) stochastic visible or observable units.

Initialization: Primarily we initialize the visible units' v to the training vector.

$$F(v, h) = -\sum_{i=1}^I \sum_{j=1}^J Opt_{-} F_{ij} v_i h_j - \sum_{i=1}^I \alpha_i v_i - \sum_{j=1}^J \beta_j h_j \dots\dots\dots (8)$$

Where $Opt_{-} F_{ij}$ (optimal features selected from MMB algorithm) represents the symmetric interaction term between the visible unit v_i and the hidden unit h_j , α, β is the bias term, I, J is the number of visible and hidden units. Between hidden units in an RBM, there are no direct influences; it is enormously easy to get an impartial sample of $(v_i, h_j)_{data}$

$$\rho(h_j = 1|v) = \delta\left(\sum_{i=1}^I Opt_{-} F_{ij} v_i + \alpha_j\right) \dots\dots\dots (9)$$

Where $\delta(x)$ is the logistic sigmoid function $\frac{1}{(1 + \exp(x))}$,

v_i, h_j is the unbiased sample.

Updating process: We update the hidden and visible units are parallel in the provided visible and hidden units. For executing the stochastic steepest ascent in the log probability of the training data, this shows the way to a much-uncomplicated learning rule as

$$\Delta Opt_{-} F_{ij} = \theta(v_i, h_j)_{data} - (v_i, h_j)_{reconstruction} \dots\dots\dots (10)$$

During the training process of every iteration, dissimilar RBM is stacked on top of it and it forms a multilayer model. The values attained for the units in RBM layers are apportioned by means of network current weight and biases. The output layer of the already-trained layers is engaged as input to the novel RBM. The accomplished deep network weights are engaged in priming a fine-tuning phase. The diagrammatic representation of DNN classifier is shown in figure 3

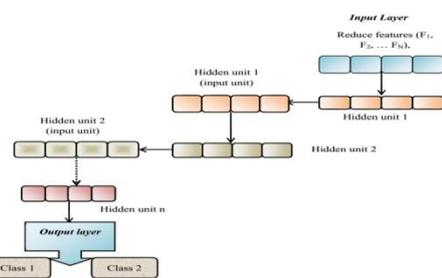


Fig 3: Diagrammatic representation of DNN

(ii) Fine Tuning Phase

The working principle of this phase is based on the ordinary backpropagation algorithm. To categorize the data from the original datasets, an output layer is intended in the top of the DNN. Also, there is N number of input neurons (depending upon the features), and three hidden layers are utilized in our DNN. The optimized weight is intended through the training stage with the help of the training data set, where back propagation starts the weights attained in the pre-training phase.

D. Optimal Solution

With the use of MMBO algorithm, optimal features are selected and then DNN classifier is used to classify the data as healthy and non-healthy from the collected datasets. The more accurate medical data classification is achieved for the four datasets (Liver, lung, heart, and thyroid). Finally, the maximum accuracy is attained for the proposed MMBO-DNN algorithm.

V. RESULT AND DISCUSSION

The proposed methodology is implemented using the working platform of MATLAB 2016a with the system configuration, i5 processors with 4GB RAM. In the result analysis section, the performances of the proposed MMB with DNN algorithm is discussed and compared with existing techniques. The performance measures such as sensitivity, specificity, and accuracy are assessed and compared with the actual result and predicted the result.

A. Description of Datasets

In the proposed medical data classification process, four medical datasets are collected from the UCI machine learning repository that is: liver, lung, heart, and thyroid disease datasets.

Liver disease dataset: Here, the considered dataset is Indian Liver Patient Dataset. It is a multivariate dataset and consists of 416 patient and 167 non-patient records. It has 583 instances and 10 attributes, the characteristics of the attribute is an integer and real [28].

Lung disease dataset: Here, the considered dataset is Lung Cancer Dataset. It is a multivariate dataset and consists of 32 instances and 56 attributes, the characteristics of the attribute is in an integer. The pathological of lung cancer is described in 3 classes.

Heart disease dataset: This dataset contains heart disease databases which refer 76 attributes and 14 subsets. The values are integer based from 0 to 4. From those experiments with this database differentiating the value as presence and 0 is referred to as an absence.

Thyroid disease dataset: This database was left at the UCI by Ross Quinlan. It has 3772 training instances and 3428 testing instances. It has 15 categorical and 6 real attributes. The problem is to determine whether a patient referred to the clinic is hypothyroid. Therefore three classes are built: normal (not hypothyroid), hyperfunction and subnormal functioning.

Table 1: Database Details

Database name		No. of classes	No. of instances	No. of attributes	Attribute characteristics	Missing values
Indian Liver Patient Dataset	Multivariate	2	583	10	Integer, Real	No
Lung Cancer	Multivariate	3	32	56	Integer	Yes

Dataset						
Heart Disease Dataset	Multivariate	2	303	76	Categorical, Integer, Real	Yes
Thyroid Disease Dataset	Multivariate, Domain-Theory	3	72000	21	Categorical, Real	No

B. Performance Measures Evaluation

Accuracy: Accuracy is a measure which decides the probability that how many results are precisely classified.

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \dots\dots\dots (11)$$

Sensitivity: The extent of real positives which are correctly distinguished is the measure of the sensitivity. It identifies with the capacity of the test to distinguish positive outcomes.

$$Sensitivity = \frac{TP}{TP + FN} \dots\dots\dots (12)$$

Specificity: The extent of negatives which are correctly distinguished is the measure of the specificity. It identifies with the capacity of the test to distinguish negative outcomes.

$$Specificity = \frac{TN}{FP + TN} \dots\dots\dots (13)$$

Where, the expansion of TP, TN, FP, and FN are True positive, True Negative, False positive, and False Negative respectively.

Table 2: Performance Analysis of Proposed and Existing Algorithms

Database	Performance Metrics	MMBO-DNN	MBO-DNN	GA-DNN	PSO-DNN	FOA-SVM
Lung Cancer Database	Accuracy	97.59	88.52	84.22	71.25	93.52
	Sensitivity	95.58	86.55	75.22	78.52	93.2
	Specificity	94.8	82.22	86.45	86.22	90
Liver Database	Accuracy	97.48	94.52	83.56	86.44	82.22
	Sensitivity	90	82.2	92.22	88.52	92
	Specificity	96.45	90	88.2	86.22	92
Heart Disease Database	Accuracy	95.45	90	86.22	86.22	92.1
	Sensitivity	97.48	86.22	88	89.45	93.2
	Specificity	95	88.52	86.45	79.22	91.22
Thyroid Database	Accuracy	96	83.33	86.22	69.45	93.33
	Sensitivity	92.22	88.5	85.22	75.52	92.2
	Specificity	93	88.2	78.22	83.21	90



Table 2 describes the performance measures such as accuracy, sensitivity, and specificity of proposed MMBO-DNN and then it is compared with existing algorithms like MBO-DNN, GA-DNN, PSO-DNN and FOA-SVM. For lung cancer dataset, the attained accuracy is high i.e. 97.59 in MMBO with DNN classifier, because the existing monarch butterfly is modified by the velocity update of PSO. Similarly, the three measures for Liver, Heart and Thyroid datasets are analyzed and compared with the other algorithms.

For a different number of features in each dataset, the accuracy of different algorithms like proposed MMBO, MBO, GA and FOA are analyzed in figure 4. By the utilization of these algorithms, the accuracy is calculated and it is rated based on the features. Figure (a) explains the Lung cancer dataset, its accuracy is in the range of 90% and 93%. Figure (b) describes the liver database; its accuracy is analyzed for the optimal feature sets 1 to 7. Figure (c) explains the Heart disease database; its accuracy is rated for the feature sets 25 to 75. Figure (d) illustrates the Thyroid disease dataset; its accuracy is in the range of 90% and 95% for the feature set 3 to 21. On comparing the four algorithms, the proposed model attains high accuracy for all the four datasets (Lung cancer, liver, Heart disease, and thyroid disease database).

Accuracy Analysis

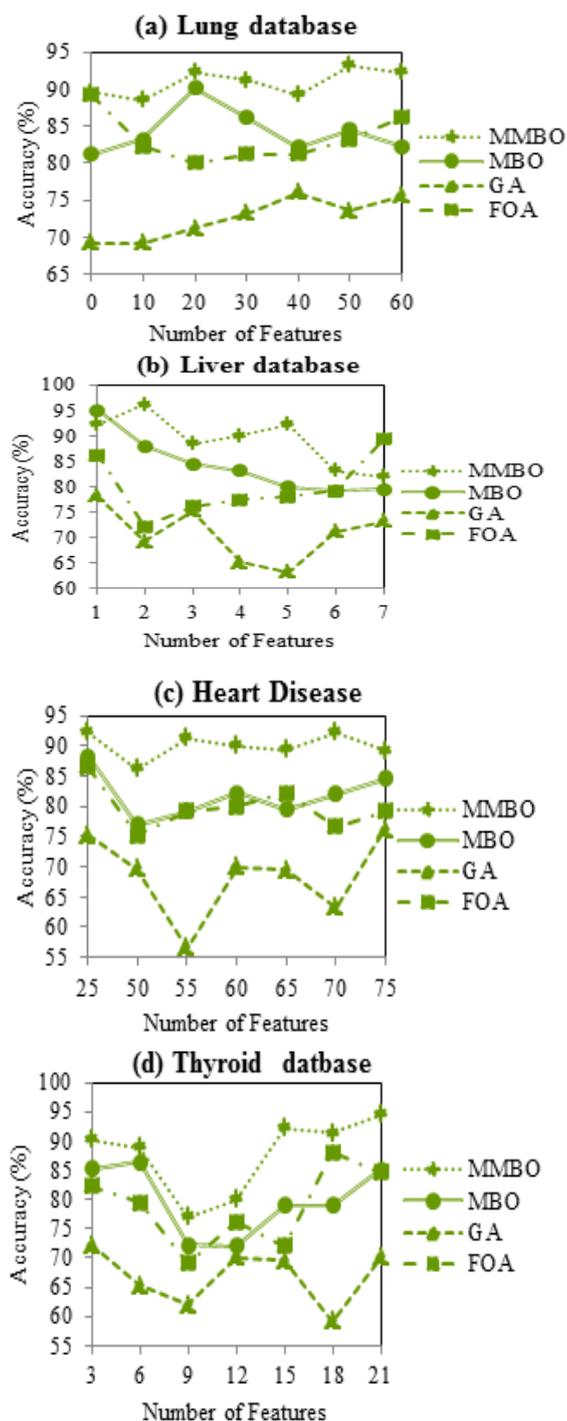


Fig 4: Comparative analysis of Algorithms (a) Lung Cancer Database, (b) Liver Database, (c) Heart Disease Database and (d) Thyroid database

Comparative Analysis

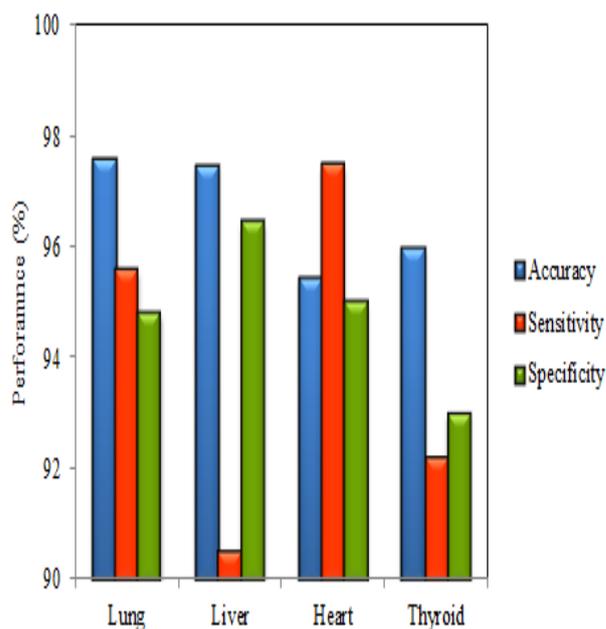


Fig 5: Comparative Analysis of Four Datasets

The comparative analysis of performance measures such as accuracy, sensitivity, and specificity for the proposed MMBO_DNN is analyzed in figure 5. On considering lung datasets, the accuracy is 97.48%, sensitivity is 95.76% and specificity is 95%. Similarly, for liver datasets, the accuracy is 97.48%, sensitivity is 90% and specificity is 96.45%. For Heart database, the accuracy is 95.45%, sensitivity is 97.48% and specificity is 95%. Finally, thyroid datasets its accuracy is 96%, sensitivity is 92.22% and specificity is 93%.

Table 3: Training and Testing Data

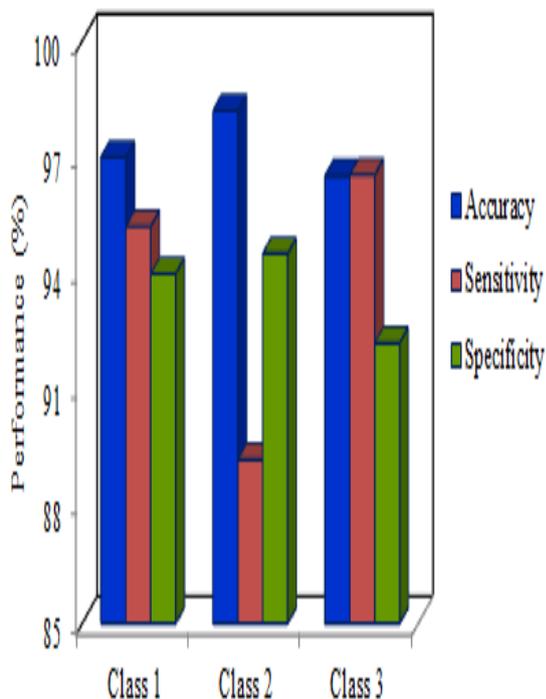
Database	Training Data	Testing Data	Accuracy	Sensitivity	Specificity	Class
Lung cancer database	20	8	95.56	94.52	94	3
Liver Database	225	52	96.45	92.52	94.85	2
Heart Disease Database	228	63	97	94.85	96	2
Thyroid Database	4287	278	95	94.75	93	3

Table 3 demonstrates a different database for a different disease. Here we calculate the training data, testing data, accuracy, sensitivity, specificity, and class for each database.

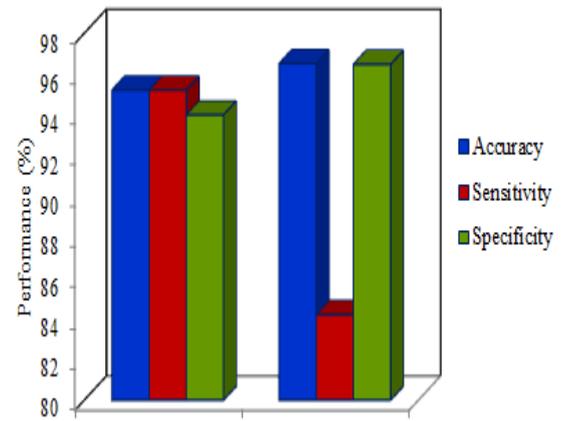
For lung cancer database training data is 20%, we have achieved the testing data as 8%, accuracy 95.56 %, sensitivity is 94.52 %, specificity is 94 % and class is 3.

Moreover, for another different disease database we have obtained the various results.

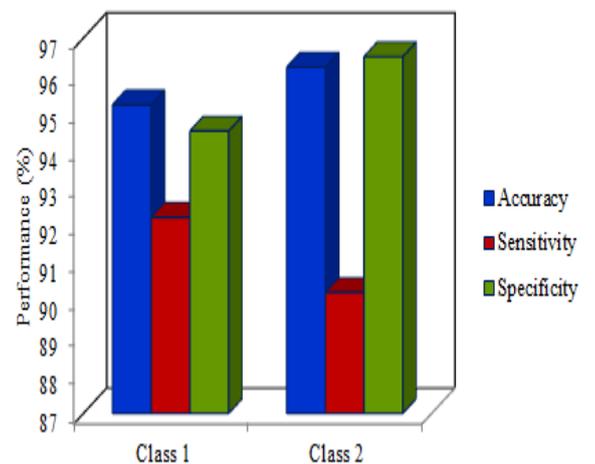
Performance Analysis of Four Datasets



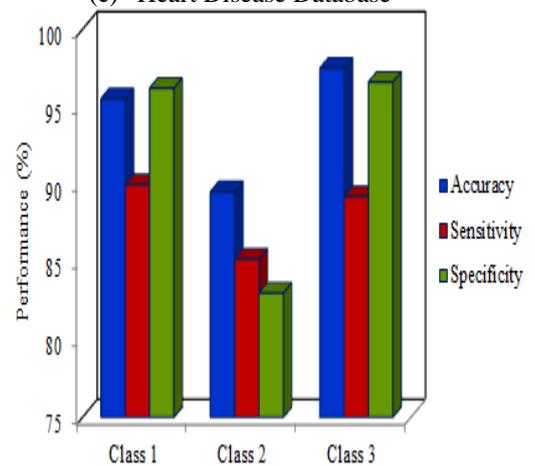
(a) Lung Database



(b) Liver Database



(c) Heart Disease Database



(d) Thyroid Database
Fig 6: Database analysis

In the figure (a) and (b) explains to analyze the patient as healthy or non-healthy from a given database. Figure (a) considered the lung database and divide the class as class 1, class 2 and class 3. The x-axis represents the class and y-axis represents the performances such as accuracy, sensitivity, and specificity. We have obtained the 97%



accuracy in class 1, for sensitivity 95% and specificity as 94%. Figure (b) explains the liver database for class 1 and class 2. 96% accuracy is obtained for class 1 and same as sensitivity and 95% for specificity. Furthermore, we have obtained the different values for different classes. Figure 6 (c) explains heart disease database for class 1 and class 2. We have to predict the result as healthy or non-healthy according to their classes. We have obtained the 95% accuracy for class 1, 93 % for sensitivity and 94% for specificity. In class 2, we have obtained the better accuracy and specificity and sensitivity gives the lower performances. Fig (b) explains the thyroid database for class 1, class 2 and class 3. For class 1, we have obtained the accuracy and specificity value of 95 %. Moreover, for the other two classes, we have obtained the different performance values.

VI. CONCLUSION

The paper focused on medical data classification in the data mining field. The major challenge of the medical data classification is finding the best subset of features from the number of attributes present in the considered dataset. Here, four medical datasets were considered for medical data classification to found out the healthy and non-healthy data. At, to begin with, the data are sustained to the pre-processing where the missing values are evacuated and then the optimal features were selected by the proposed MMBO with DNN algorithm. From the analysis, it can be seen that MMBO with DNN is the most appropriate method for data classification; it selects optimal features and it only requires a moderate computational cost for solving medical data classification problems. The performance metrics such as sensitivity, specificity and the accuracy of the MMBO-DNN achieves the maximum value i.e. 95.58, 94.8, and 97.59 for the lung cancer datasets. Results concluded that the proposed MMBO-DNN achieved better classification results compared to other techniques. In the future, we extend our approach with new measures for medical data values and removing the missing values by new imputation approach, classification, and prediction. Future contributions will concentrate on developing novel and hybrid optimization algorithms for feature selection and medical data classification.

REFERENCES

- ManjeevanSeera, and CheePeng Lim, "A hybrid intelligent system for medical data classification", Journal of Expert Systems with Applications, Vol. 41, no. 5, pp. 2239-2249, 2014.
- Rajeswari, K., Vaithiyanathan, V. and Pedde, S.V., 2013. Feature selection for classification in medical data mining. International Journal of Emerging Trends and Technology in Computer Science (IJETCS), 2(2), pp.492-7.
- HariBabu, and E. S. Gopi, "Medical data classifications using a genetic algorithm based generalized kernel linear discriminant analysis," Journal of Procedia Computer Science, Vol. 57, pp. 868-875, 2017.
- Senthil Kumar and H. Hannah Inbarani, "Optimistic multi-granulation rough set based classification for medical diagnosis", Journal of Procedia Computer Science, Vol. 47, pp. 374-382, 2015.
- Murad Al-Rajab, Joan Lu, and QiangXu, "Examining applying high-performance genetic data feature selection and classification algorithms for colon cancer diagnosis", Journal of Computer Methods and Programs in Biomedicine, Vol. 146, pp. 11-24, 2017.
- Marian Gorzalczy and FilipRudziński, "Interpretable and accurate medical data classification—a multi-objective genetic-fuzzy optimization approach", Journal of Expert Systems with Applications, Vol. 71, pp. 26-39, 2017.
- Yen-Chen Chen, Yo-Cheng Chang, Wan-Chi Ke, and Hung-Wen Chiu, "Cancer adjuvant chemotherapy strategic classification by artificial neural network with gene expression data: An example for non-small cell lung cancer", Journal of biomedical informatics, Vol. 56, pp. 1-7, 2015.
- Mohapatra, SreejitChakravarty, and P. K. Dash, "Microarray medical data classification using kernel ridge regression and modified cat swarm optimization based gene selection system", Journal of Swarm and Evolutionary Computation, Vol. 28, pp. 144-160, 2016.
- Hanaa Salem, GamalAttiya, and Nawal El-Fishawy, "Classification of human cancer diseases by gene expression profiles", Journal of Applied Soft Computing, Vol. 50, pp.124-134, 2017.
- Ramos-González, J., López-Sánchez, D., Castellanos-Garzón, J. A., de Paz, J. F., &Corchado, J. M. (2017). A CBR framework with gradient boosting based feature selection for lung cancer subtype classification. Computers in Biology and Medicine, 86, 98–106.
- Dennis and S. Muthukrishnan, "AGFS: Adaptive Genetic Fuzzy System for medical data classification." Journal of Applied Soft Computing, Vol. 25, pp. 242-252, 2014.
- Cigdem, O., and Demirel, H. (2018). Performance analysis of different classification algorithms using different feature selection methods on Parkinson's disease detection. Journal of Neuroscience Methods.
- ManiruzzamanNishith Kumar, MdMenhazulAbedin, MdShaykhul Islam, Harman S. Suri, Ayman S. El-Baz, and Jasjit S. Suri. "Comparative approaches for classification of diabetes mellitus data: Machine learning paradigm", Journal of Computer Methods and Programs in Biomedicine, Vol. 152, pp. 23-34, 2017.
- SreejitChakravartyMohapatra and Pradipta K. Dash. "An improved cuckoo search based extreme learning machine for medical data classification", Journal of Swarm and Evolutionary Computation, Vol. 24, pp. 25-49, 2015.
- SurangaKasthurirathne, Brian E. Dixon, Judy Gichoya, HuipingXu, Yuni Xia, Burke Mamlin, and Shaun J. Grannis, "Toward better public health reporting using existing off the shelf approaches: The value of medical dictionaries in automated cancer detection using plaintext medical data", Journal of Biomedical Informatics, Vol. 69, pp. 160-176, 2017.
- Chip Lynch, BehnazAbdollahi, Joshua D. Fuqua, R. Alexandra, James A. Bartholomai, Rayeane N. Balgeman, Victor H. van Berkel, and Hermann B. Frieboes, "Prediction of lung cancer patient survival via supervised machine learning classification techniques", Journal of Medical Informatics, Vol. 108, pp. 1-8, 2017.
- Harb, H.M. and Desuky, A.S., 2014. Feature selection on the classification of medical datasets based on particle swarm optimization. International Journal of Computer Applications, 104(5).
- Tomcikova, D., Felsoci, M., Spinar, J., Miklik, R., Mikusova, T., Vitovec, J., ...Parenica, J. (2013). Risk of in-hospital mortality identified according to the typology of patients with acute heart failure: Classification tree analysis on data from the Acute Heart Failure Database–Main registry. Journal of Critical Care, 28(3), 250–258.
- Jain, D., & Singh, V. (2018). Feature selection and classification systems for chronic disease prediction: A review. Egyptian Informatics Journal.
- Shuo Yang, Jing-ZhiGuo, and Jun-Wei Jin, "An improved Id3 algorithm for medical data classification", Journal of Computers & Electrical Engineering, 2017.
- Ahmed Abdel-Zaher and Ayman M. Eldeib, "Breast cancer classification using deep belief networks", Journal of Expert Systems with Applications, Vol. 46, pp.139-144, 2016.
- Ghaddar, B., &Naoum-Sawaya, J. (2018). High dimensional data classification and feature selection using support vector machines. European Journal of Operational Research, 265(3), 993–1004.
- Lavanya, D. and Rani, D.K.U., 2011. Analysis of feature selection with classification: Breast cancer datasets. Indian Journal of Computer Science and Engineering (IJCSSE), 2(5), pp.756-763.
- Shen, L., Chen, H., Yu, Z., Kang, W., Zhang, B., Li, H., ... Liu, D. (2016). Evolving support vector machines using fruit fly optimization for medical data classification. Knowledge-Based Systems, 96, 61–75.



25. Sim, J., Lee, J. S., & Kwon, O. (2015). Missing Values and Optimal Selection of an Imputation Method and Classification Algorithm to Improve the Accuracy of Ubiquitous Computing Applications. *Mathematical Problems in Engineering*, 2015, 1–14.
26. Gai-Ge Wang, Suash Deb, Zhihua Cui, 2015., Monarch butterfly optimization, *Journal of Neural Comput& Applications*.
27. Mohsen, H., El-Dahshan, E.-S. A., El-Horbaty, E.-S. M., & Salem, A.-B. M. (2018). Classification using deep learning neural networks for brain tumors. *Future Computing and Informatics Journal*, 3(1), 68–71.
28. Gulia, A., Vohra, R. and Rani, P., 2014. Liver patient classification using intelligent techniques. *International Journal of Computer Science and Information Technologies*, 5(4), pp.5110-5115.
29. Lakshmanaprabu, S. K., Shankar, K., Ilayaraja, M., Nasir, A. W., Vijayakumar, V., & Chilamkurti, N. (2019). Random forest for big data classification in the internet of things using optimal features. *International Journal of Machine Learning and Cybernetics*, 1-10. <https://doi.org/10.1007/s13042-018-00916-z>
30. Lakshmanaprabu, S. K., Mohanty, S. N., Shankar, K., Arunkumar, N., & Ramirez, G. (2019). Optimal deep learning model for classification of lung cancer on CT images. *Future Generation Computer Systems*, 92, 374-382.
31. Karthikeyan, K., Sunder, R., Shankar, K., Lakshmanaprabu, S. K., Vijayakumar, V., Elhoseny, M., & Manogaran, G. (2018). Energy consumption analysis of Virtual Machine migration in cloud using hybrid swarm optimization (ABC–BA). *The Journal of Supercomputing*, 1-17.
32. Shankar, K., Elhoseny, M., Lakshmanaprabu, S. K., Ilayaraja, M., Vidhyavathi, R. M., Elsoud, M. A., & Alkhambashi, M. (2018). Optimal feature level fusion based ANFIS classifier for brain MRI image classification. *Concurrency Computat Pract Exper*, e4887.
33. Lakshmanaprabu, S. K., Shankar, K., Khanna, A., Gupta, D., Rodrigues, J. J., Pinheiro, P. R., & De Albuquerque, V. H. C. (2018). Effective Features to Classify Big Data Using Social Internet of Things. *IEEE Access*, 6, 24196-24204.
34. Muslihudin, M., Wanti, R., Hardono, N., Shankar, K., Ilayaraja, M., Maseleno, A., ... & Mukodimah, S. (2018). Prediction of Layer Chicken Disease using Fuzzy Analytical Hierarcy Process. *International Journal of Engineering & Technology*, 7(2.26), 90-94.