applied sciences

MDPI

Article

Taylor bird swarm algorithm based Deep belief network for Heart Disease Diagnosis using medical data classification

Afnan M. Alhassan*[1,2], Wan Mohd Nazmee Wan Zainon[1]

[1]Department of Computer Science, Universiti Sains Malaysia, George Town 11800, Malaysia

[2]College of Computing and Information Technology, Shaqra University, Shaqra, Saudi Arabia

*Corresponding author: Afnan M. Alhassan (afnan@student.usm.my)

Abstract: The contemporary medicines produce a huge amount of information contained in the medical database. Thus, the extraction of valuable knowledge and making a scientific decision for the treatment of disease has progressively become necessary for attaining effective diagnosis. The obtainability of large medical data leads to the requirement of effective data analysis tools for extracting constructive knowledge. This paper proposes a novel method for medical data classification. Here, the pre-processing of medical data is done using log-transformation that converts the data to its uniform value range. Then, the feature selection process is performed using sparse Fuzzy-c-means (FCM) for selecting the significant features for classifying the medical data. The incorporation of sparse FCM for the feature selection process provides more benefits for interpreting the models, as this sparse technique provides important features for the classification and can be utilized for handling high dimensional data. Then, the selected features are given to the Deep belief network (DBN), which is trained using the proposed Taylor-based bird swarm algorithm (Taylor-BSA) for the classification. Here, the proposed Taylor-BSA is designed by combining the Taylor series and Bird swarm algorithm (BSA). The proposed Taylor-BSA-DBN outperformed other methods with maximal accuracy of 93.4%, maximal sensitivity of 95%, and maximal specificity of 90.3% respectively.

Keywords: Medical data classification, DBN, Heart disease diagnosis, Sparse FCM.

1. Introduction

The current medicine produces much information accumulated in the medical dataset. The extraction of constructive knowledge helps to make scientific decisions to diagnose disease. The medical can enhance the management of hospital information

and endorse the growth of telemedicine. Medical is chiefly focussed on patient care and only secondarily for research resources. The main rationalization to collect medical data is to promote each patient's health conditions [1]. The accessibility of numerous medical data causes reflection which poses effectual and significant techniques for processing the data to extract beneficial knowledge. However, the diagnostics of various diseases indicate significant issues in data analysis [2]. The quantifiable diagnosis is performed by the doctor's guidance rather than patterns of the medical dataset and thus there is a possibility of incorrect diagnosis [3]. The cloud assists to increase the high-quality association of managing medical data, which includes compliance management, policy integration, access controls, and identity management [23].

Heart disease is a foremost source of death worldwide. As we are moving towards a new industrial revolution, thus the changes in lifestyle arise obesity, diabetes, hypertension, and smoking which in turn affect heart disease [4]. The treatment of disease is a complex and complex mission in medicine. The discovery of heart disease with different factors is considered a multi-layered issue [5]. Thus, the patient's medical data is employed for collecting to simplify the diagnosis process and use knowledge for devising the disease. Offering valuable services with fewer costs is a major limitation by healthcare. In [6], valuable quality service specifies precise diagnosis in patients and offer effective treatment. The poor clinical decisions cause disasters which may affect the health of patients. Automated approaches like the machine-learning approach and data mining approaches assist to attain clinical tests or diagnosis at reduced risk [7] [8]. The classification and pattern recognition by the machine learning algorithms are superior to the traditional statistical approaches. Also, it is widely included in prognostic and diagnosis monitoring. The machine learning approach supports decision-making, which increases the safety of the patients and avoids medical errors so that it can be used in the clinical decision support systems (CDSS) [31] [32].

Several methods ~~are~~ devised for automatic heart disease detection to evaluate the efficiency of ~~the~~ decision tree and Naïve Baye Moreover, optimization with ~~the~~ genetic ~~algorithm~~ is employed for minimizing the ~~count of attributes~~ without forfeiting accuracy and efficiency to diagnose heart disease [10]. ~~The~~ data mining methods for heart disease diagnosis ~~includes~~ bagging ~~algorithm~~, neural network, support vector machine, and automatically defined groups [9].In [11] employed 493 samples from cerebrovascular disease prevention and utilized three classification techniques namely Bayesian classifier, decision tree, and backpropagation neural network for constructing classification models. In [12], a method is devised for diagnosing coronary artery disease. The method utilized 303 samples by adapting the feature creation technique. In [13], a methodology is devised for automatically detecting the efficiency of features to reveal heart rate signals. In [14], a hybrid

algorithm is devised with KNN and genetic algorithm for effectual classification. The method utilized a genetic search as a ~~decency~~ measure for ranking attributes. Then the classification algorithm is devised on evaluated attributes for heart disease diagnosis. The extraction of valuable information from huge data is a time-consuming task [15]. The size of the medical dataset is increasing in a rapid manner and the advanced techniques of data mining help physicians for making effective decisions. However, the issues of heart disease data involve feature selection in which the imbalance of samples and lack of magnitude of features are some of the issues [16]. Even though there are methods for heart disease prediction with real-world medical data, but these methods are devised for improving accuracy and time for computation in disease prediction [20]. In [22], a hybrid model with CS and a rough set is adapted for diagnosing heart disease. The drawback is a rough set produces an unnecessary number of rules. These challenges in the heart disease diagnosis are considered as motivation and a novel method named, TaylorBSA-DBN is proposed for medical data classification.

The purpose of the research is to present a medical data classification strategy, for which the proposed TaylorBSA-DBN is employed. The major contribution of the research is the classification of medical data using selected features. Here, the feature selection is performed using Sparse FCM for selecting imperative features. In addition, DBN is employed for classifying the medical data using the features. Here, the DBN is trained by proposed TaylorBSA in such a way that the model parameters are learned optimally. The proposed TaylorBSA is developed through the inheritance of the high global convergence property of BSA in the Taylor series. Hence, the proposed TaylorBSA-DBN renders effective accuracy, sensitivity, and specificity while facilitating medical data classification.

The major contribution of the paper is:

- **Proposed TaylorBSA-DBN for medical data classification:** A classifier, TaylorBSA-DBNis proposed by modifying the training algorithm of the DBN with the Taylor-BSA algorithm, which is newly derived by combining the Taylor series and BSA algorithm, for the optimal tuning of weights and biases. The proposed TaylorBSA-DBN is adapted for medical data classification.

Other sections of the paper are arranged as: Section 2 elaborates description of the conventional medical data classification strategies utilized in literature and challenges faced, which are considered as the inspiration for developing the proposed technique. The proposed method for medical data classification using modified DBN is portrayed in Section 3. The outcomes of the proposed strategy with other methods are depicted in Section 4 and Section 5 present conclusion.

2. Motivations

This section illustrates eight classification strategies employed for medical data classification along with its challenges.

2.1. Literature survey

The eight conventional strategies using medical data for heart disease diagnosis is illustrated below: Reddy, G.T *et al.* [15] devised an adaptive genetic algorithm with fuzzy logic (AGAFL) model for predicting ~~the~~ heart disease which assists clinicians for treating heart disease at earlier phases. The model comprises rough sets with a fuzzy rule-based classification module and heart disease feature selection module. The obtained rules from fuzzy classifiers were optimized by adapting an adaptive genetic algorithm. Initially, the significant features that affect heart disease are chosen using rough set theory. Then, the second step predicted heart disease with ~~the~~ AGAFL classifier. The method is effective in handling noisy data and works effectively with large attributes. Nourmohammadi-Khiarak *et al.* [16] devised a method for selecting features and reduce the number of features.

Here, the imperialist competitive algorithm was devised for choosing important features from heart disease. This algorithm offers an optimal response in selecting features. Moreover, the K-nearest neighbour algorithm was utilized for classification. The method showed that the accuracy of feature selection was enhanced. However, the method failed to utilize incomplete or missed data. Magesh, G. and Swarnalatha, P [17] devised a model using Cleveland's heart samples for heart disease diagnosis. The method employed cluster-based DT learning (CDTL) for diagnosing heart disease. Here, the original set was partitioned using target label distribution. From elevated distribution samples, the possible class was derived. For each class set, the features were detected using entropy for diagnosing heart disease.Thiyagaraj, M. and Suseendran, G [18] developed Particle Swarm Optimization and Rough Sets with Transductive Support Vector Machines (PSO and RS with TSVM) for heart disease diagnosis. This method improved data integrity to minimize data redundancy. The normalization of data was carried out using Zero-Score (Z-Score). Then, the PSO was employed for selecting the optimal subset of attributes and reduce computational overhead and enhances prediction performance. At last, the Radial Basis Function-Transductive Support Vector Machines (RBF-TSVM) classifier was employed for heart disease prediction. Abdel-Basset, M *et al.* [19] devised a model using IoT for determining and monitoring heart patients. The goal of the healthcare model was to obtain improved precision for diagnosis. The neutrosophic multi-criteria decision making (NMCDM) technique was employed for aiding patients for observing patients suffering from heart failure. Moreover, the model provided an accurate solution that can benefit over declining mortality and the cost of treatment. Nilashi, M *et al.* [20] devised a predictive technique for heart disease diagnosis with machine learning models. Here, the method adapted unsupervised and supervised learning for

diagnosing heart disease. In addition, the method employed Self-Organizing Map, Fuzzy Support Vector Machine (Fuzzy SVM), and Principal Component Analysis (PCA) for missing value assertion. Moreover, incremental PCA and FSVM are devised for incremental learning of data to minimize the time taken for computation in disease prediction. IShah, S.M.S *et al.* [21] devised an automatic diagnostic technique for diagnosing heart disease. The method evaluated the pertinent feature subset by employing the benefits of feature selection and extraction models. For accomplishing the feature selection, two algorithms, namely accuracy based feature selection algorithm (AFSA) and Mean Fisher based feature selection algorithm (MFFSA) for heart disease diagnosis. However, the method failed to employ PCA for dimension reduction. Acharjya, D.P [22] devised a hybrid method for diagnosing heart disease. The method combined cuckoo search (CS) and rough set to infer decision rules. Moreover, the CS was employed for discovering essential features. In addition, three major features were evaluated with rough set rules. The method improved the feasibility but failed to induce an intuitionistic fuzzy rough set and CS for diagnosing heart disease.

3. Proposed Taylor-BSA-based DBN for medical data classification

The accessibility of a large amount of medical data led to the requirement of strong data analysis tools for extracting valuable knowledge. The researchers are adapting data mining and statistical tools for improving the analysis of data on huge datasets. The diagnosis of a disease is the foremost application in which data mining tools are offering triumphant results. Medical data tends to be rich in information, but poor in knowledge. Thus, there is a deficiency of effectual analysis tools for discovering hidden relation and trends from medical data generated from clinical records. The processing of medical data brings a manifestation if has some powerful methods. Thus, the proposed Taylor-BSA-based DBN is devised to process the medical data for attaining effective medical data classification. Figure 1 portrays the schematic view of the proposed Taylor-BSA-based DBN for medical data classification. The complete process of the proposed model is pre-processing feature selection, and classification. At first, the medical data is fed as an input to the pre-processing phase, wherein log transformation is applied to pre-process the data. Log transformation is applied for minimizing skew and to normalize the data. Once the pre-processed data is obtained, then it is further subjected to the feature selection phase. In the feature selection phase, the imperative features are selected with Sparse FCM. After obtaining imperative features, the classification is performed with DBN wherein the training of DBN is carried out using Taylor-BSA. The proposed Taylor-BSA is devised by combining the Taylor series and BSA. The output produced from the classifier is the classified medical data.

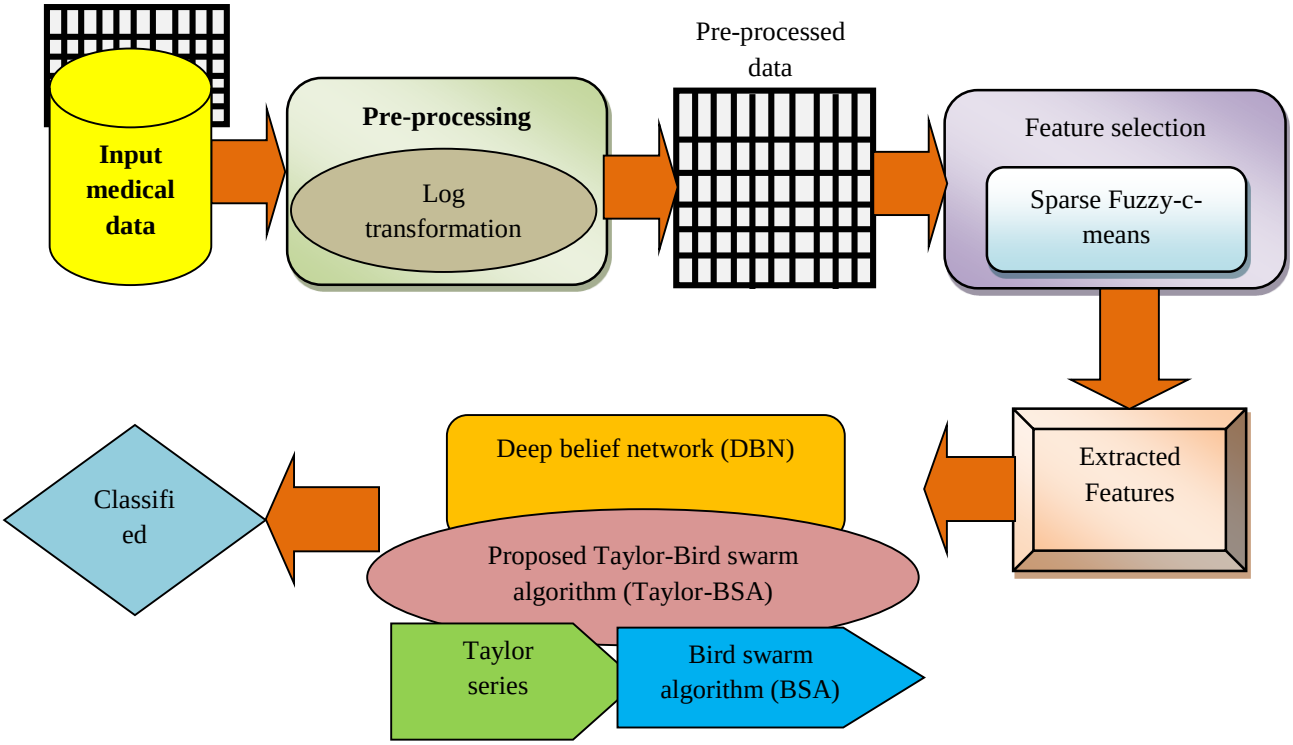

Figure 1. Schematic view of proposed Taylor-BSA based DBN for classifying medical data

Consider an input medical data be given as A, with various attributes and is expressed as

$$A = \{A_{G,H}\}; (1 \le G \le B); (1 \le H \le C) \quad (1)$$

where $A_{G,H}$ denotes H^{th} attribute in G^{th} data, B specifies a total number of data, and C specifies total attributes in each data. The dimension of the database is represented as $[B \times C]$.

3.1 Pre-processing

The importance of pre-processing is to facilitate smoother processing of the input data. Additionally, the pre-processing is carried out for eliminating the noise and artefacts contained in the data. In this method, the pre-processing is carried out by using log transformation, in which each data is replaced with a log function, wherein the base of the log is set by the analyst that maybe 2, or 10. The process is used to compress the massive data. In addition, the log transformation has extensively adapted the method to solve skewed data and assist data normalization. The log transformation is formulated as,

$$D = \log_{10}(A) \tag{2}$$

The dimension of pre-processed dataset A becomes $[B \times C]$.

3.2 Selection of features with Sparse FCM clustering

The pre-processed data is fed to the feature selection module considering the Sparse FCM algorithm [24], which is the modification of standard FCM and the benefit of using Sparse FCM is to provide high dimensional data clustering. The pre-processed data contains different types of attributes each indicating individual value. In the medical data classification strategy, the sparse FCM is applied for determining the features from each data. The sparse FCM clustering algorithm clusters nodes to attain communication between nodes through the cluster head and facilitate effective detection of the attacker node. Generally, in sparse FCM, dimensional reduction is effective, and poses the ability to handle disease diagnosis without delay and is easier with optimization technique.

3.3. Classification of medical data with proposed Taylor-BSA-based DBN

In this section, medical data classification using the proposed Taylor-BSA method is presented and the classification is progressed using the feature vector.

3.3.1 Proposed Taylor-BSA algorithm

The proposed Taylor-BSA is the combination of the Taylor series and BSA. Taylor series [25] explains the functions of complex variables, and it is the expansion of a function into an infinite sum of terms. It not only serves as a powerful tool, but also helps in evaluating integrals and infinite sums. Moreover, the Taylor series is the one-step process, and it can deal with higher-order terms. Taylor series seems to be advantageous for derivations and can be used to get theoretical error bounds. Above all, the Taylor series ensures the accuracy of classification. Also, it is a simple method

to solve complex functions. BSA [26] is duly based on the social behaviours of birds that follow some idealistic rules. BSA is highly accurate than other standard optimizations with highly efficient, accurate, and robust performance. In addition, there is a perfect balance between exploration and exploitation in BSA. The DBN has recently become a popular approach in machine learning for its promised advantages such as fast inference and the ability to encode richer and higher order network structures. DBN is used to extract better feature representations, and several related tasks are solved simultaneously by using shared representations. Also, it has the advantages of a multi-layer structure, and pre-training with the fine-tuning learning method. The Algorithmic steps of the proposed Taylor-BSA are described below:

Step 1) Initialization: The first step is the initialization of population and other algorithmic parameters, which includes: $F_{i,j}$; $(1 \le i \le j)$, where, the population size is denoted as j, $h_{\max}$ represent maximal iteration, $prob$ indicate the probability of foraging food, and the frequency of flight behaviour of birds is expressed as Ft.

Step 2) Determination of objective function: The selection of the best position of bird is termed as a minimization issue. The minimal value of error defines the optimal solution.

Step 3) Position update of the birds: For updating the positions, birds have three phases, which are decided using probability. Whenever the random number $Rand(0,1) < prob$, then the update is based on foraging behaviour or else, the vigilance behaviour commences. On the other hand, swarm splits as scroungers and producers, which is modelled as flight behaviours. Finally, the feasibility of the solutions is verified and the best solution is retrieved.

Step 4) Foraging behaviour of birds: The individual bird searches the food based on their own experience, and the behaviour of swarm, which is given below. The standard equation of the foraging behaviour of birds [26] is given by,

$$F_{i,j}^{h+1} = F_{i,j}^{h} - F_{i,j}^{h}\, Rand\,(0,1)[Z+T] + Rand\,(0,1)[\mathrm{P}_{i,j} Z + Y_j T] \tag{3}$$

where, $F_{i,j}^{h+1}$ and $F_{i,j}^{h}$ denotes the location of i^{th} bird in j^{th} dimension at $(h+1)$ and h, $\mathrm{P}_{i,j}$ refers to the previous best position of the i^{th} bird, $Rand\,(0,1)$ is independent uniformly distributed numbers, Y_j indicates the best previous location shared by the birds swarm, Z denotes the cognitive accelerated coefficients, and T denotes the social accelerated coefficients. Here, Z and T are positive numbers.

According to Taylor series [25], the update equation is expressed as,

$$\begin{aligned} F_{i,j}^{h+1} &= 0.5F_{i,j}^{h} + 1.3591F_{i,j}^{h-1} - 1.359F_{i,j}^{h-2} + 0.6795F_{i,j}^{h-3} \\ &- 0.2259F_{i,j}^{h-4} + 0.0555F_{i,j}^{h-5} - 0.0104F_{i,j}^{h-6} + 1.38e^{-3}F_{i,j}^{h-7} - 9.92e^{-5}F_{i,j}^{h-8} \end{aligned} \tag{4}$$

$$F_{i,j}^{h} = \frac{1}{0.5}\begin{bmatrix} F_{i,j}^{h+1} - 1.3591F_{i,j}^{h-1} + 1.359F_{i,j}^{h-2} - 0.6795F_{i,j}^{h-3} + 0.2259F_{i,j}^{h-4} \\ - 0.0555F_{I,J}^{h-5} + 0.0104F_{i,j}^{h-6} - 1.38e^{-3}F_{i,j}^{h-7} + 9.92e^{-5}F_{i,j}^{h-8} \end{bmatrix} \tag{5}$$

Substituting equation (5) in equation (3),

$$F_{i,j}^{h+1} = F_{i,j}^{h} - \left[\begin{array}{l} 2F_{i,j}^{h+1} - 2.7182F_{i,j}^{h-1} + 2.718F_{i,j}^{h-2} - 1.359F_{i,j}^{h-3} \\ -0.4518F_{i,j}^{h-4} - 0.111F_{i,j}^{h-5} + 0.0208F_{i,j}^{h-6} - 0.00276F_{i,j}^{h-7} + 0.0001984F_{i,j}^{h-8} \end{array}\right] \quad (6)$$
$$Rand(0,1)[Z+T] + Rand(0,1)[P_{i,j}Z + Y_jT]$$

$$F_{i,j}^{h+1} + 2F_{i,j}^{h+1} = F_{i,j}^{h} + \left[\begin{array}{l} 2.7182F_{i,j}^{h-1} - 2.718F_{i,j}^{h-2} + 1.359F_{i,j}^{h-3} \\ +0.4518F_{i,j}^{h-4} + 0.111F_{i,j}^{h-5} - 0.0208F_{i,j}^{h-6} + 0.00276F_{i,j}^{h-7} - 0.0001984F_{i,j}^{h-8} \end{array}\right] \quad (7)$$
$$Rand(0,1)[Z+T] + Rand(0,1)[P_{i,j}Z + Y_jT]$$

$$3F_{i,j}^{h+1} = F_{i,j}^{h} + \left[\begin{array}{l} 2.7182F_{i,j}^{h-1} - 2.718F_{i,j}^{h-2} + 1.359F_{i,j}^{h-3} \\ +0.4518F_{i,j}^{h-4} + 0.111F_{i,j}^{h-5} - 0.0208F_{i,j}^{h-6} + 0.00276F_{i,j}^{h-7} - 0.0001984F_{i,j}^{h-8} \end{array}\right] \quad (8)$$
$$Rand(0,1)[Z+T] + Rand(0,1)[P_{i,j}Z + Y_jT]$$

$$F_{i,j}^{h+1} = \frac{1}{3}\left[\begin{array}{l} F_{i,j}^{h} + \left[\begin{array}{l} 2.7182F_{i,j}^{h-1} - 2.718F_{i,j}^{h-2} + 1.359F_{i,j}^{h-3} \\ +0.4518F_{i,j}^{h-4} + 0.111F_{i,j}^{h-5} - 0.0208F_{i,j}^{h-6} + 0.00276F_{i,j}^{h-7} - 0.0001984F_{i,j}^{h-8} \end{array}\right] \\ Rand(0,1)[Z+T] + Rand(0,1)[P_{i,j}Z + Y_jT] \end{array}\right] \quad (9)$$

Step 5) Vigilance Behaviour of Birds: The birds move towards the centre during which the birds compete with each other and the vigilance behaviour of birds is modelled as,

$$F_{i,j}^{h+1} = F_{i,j}^{h} + V_1\left(\mu_j - F_{i,j}^{h}\right) \times Rand(0,1) + V_2\left[U_{oj} - F_{i,j}^{h}\right] \times Rand(-1,1) \quad (10)$$

$$V_1 = w_1 \times \exp\left(\frac{-RQ(U)_i}{\sum RQ + \psi} \times v\right) \quad (11)$$

$$V_2 = w_2 \times \exp\left[\left(\frac{RQ(U)_i - RQ(U)_T}{\left|RQ(U)_T - RQ(U)_i\right| + \psi}\right)\frac{v \times RQ(U)_T}{\sum RQ + \psi}\right] \quad (12)$$

where, V represents the number of birds, w_1 and w_2 are the positive constants lying in the range of $[0,2]$, $RQ(U)_i$ denotes the optimal fitness value of i^{th} bird, and $\sum RQ$ corresponds to the addition of the best fitness values of the swarm. ψ be the constant that keeps optimization away from zero-division error. T signifies the positive integer.

Step 6) Flight behaviour: This behaviour of the bird's progress when the birds fly to another site in case of any threatening events and foraging mechanisms. When the birds reach a new site, they search for food. Few birds in the group try acting as producers and few as scroungers. The behaviour is modelled as,

$$F_{i,j}^{h+1} = F_{i,j}^{h} + Rand\,r(0,1) \times F_{i,j}^{h} \quad (13)$$

$$F_{i,j}^{h+1} = F_{i,j}^{h} + \left(F_{\gamma,j}^{h} - F_{i,j}^{h}\right) \times Fl \times Rand(0,1) \quad (14)$$

where, $Random(0,1)$ refer to the Gaussian distributed random number with zero-mean and standard deviation.

Step 7) Determination of best solution: The best solution is evaluated based on error function. If the newly computed solution is best than the previous one, then it is updated by the new solution.

Step 8) Terminate: The optimal solutions are derived in an iterative manner until the maximum number of iterations is reached. The pseudo-code of the proposed Taylor-BSA algorithm is illustrated in table 1.

Table 1. Pseudocode for the proposed Taylor-BSA algorithm

Input: Bird swarm population $W_{k,l}$ **;** $(1 \le k \le b)$
Output: Best solution
Procedure:
Begin
Population initiation: $F_{i,j}$**;** $(1 \le i \le p)$
Read the parameters: b **-population size;** $h_{\max}$ **maximal iteration,** $prob$ **-probability of foraging food,** Fl **-frequency of flight behaviour of birds**
Determine the fitness of the solutions
While $h < h_{\max}$
For $k = 1:b$
If $Rand(0,1) < prob$
Foraging behaviour using equation (3)
Else
Vigilance behaviour using equation (12)
End if
End for
Else
Split the swarm as scroungers and producers
For $k = 1:b$
If k **is a producer**
Update using equation (13)
Else
Update using equation (14)
End if
End for
Check the feasibility of the solutions
Return the best solution
$h = h + 1$

End while
Optimal solution is obtained
End

3.3.2 Architecture of DBN

The DBN [27] is a subset of Deep Neural Network (DNN) and comprises different layers of Multilayer Perceptrons (MLPs) and Restricted Boltzmann Machines (RBMs). RBMs comprises visible and hidden units that are associated with weights. The basic structural design of the DBN is illustrated using figure 2.

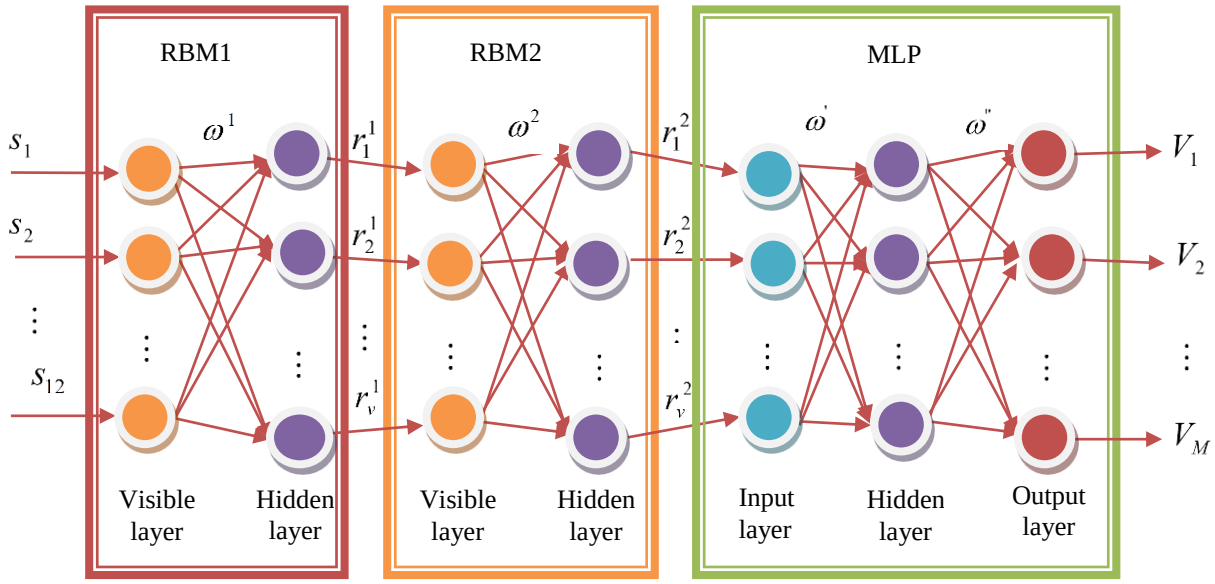

Figure 2. Architectural view diagram of DBN classifier

a) Training of Deep Belief Network

This section elaborates on the training process of the proposed Taylor-BSA-based DBN classifier. RBM has unsupervised learning based on the gradient descent method, whereas MLP performs a supervised learning method using the standard backpropagation algorithm. Therefore, the training of DBN is based on a gradient descent- backpropagation algorithm. Here, the most appropriate weights are chosen

optimally for the update. The training procedure of the proposed DBN classifier is described below,

I. Training of RBM layers

A training sample N is given as the input to the first layer of RBM. It computes the probability distribution of the data and encodes it into the weight parameters. The steps involved in the training process of RBM are illustrated below,

1. The input training sample is read and the weight vector is produced randomly.
2. The probability function of each hidden neuron in the first RBM is calculated.
3. The positive gradient is computed using a visible vector and the probability of the hidden layer.
4. The probability of each visible neuron is obtained by reconstructing the visible layer from the hidden layer.
5. Then, the probability of reconstruction of hidden neurons is obtained by resampling the hidden states.
6. Compute the negative gradient.
7. Weights are updated by subtracting the negative gradient from the positive gradient.
8. Update the weights for the next iteration using the steepest or gradient descent algorithm.
9. Calculate the energy for a joint configuration of the neurons in the visible and the hidden layers.

II. Training of MLP

The training procedure in MLP is based on backpropagation approach by feeding the training data, which is the hidden output of the second RBM layer through the network. Analyzing the data, the network is adjusted iteratively until the optimal weights are chosen. Moreover, Taylor-BSA is employed to compute the optimal weights, which are determined using the error function. The training procedure is summarized below,

1. Randomly initialized the weights.
2. Read the input sample from the result of the preceding layer.
3. Then, the average error is obtained based on the difference between the obtained output and the desired output.
4. Calculate the weight updates in the hidden and the visible layers.
5. The new weights are obtained from the hidden and the visible layers by applying gradient descent.
6. The new weights are identified using the update equation of Taylor-BSA.
7. Estimate the error function using gradient descent and Taylor-BSA.
8. Choose the minimum error and repeat the steps.

4. Results and Discussion

This section elaborates on the assessment of the proposed strategy with classical strategies for medical data classification using accuracy, sensitivity, and specificity. The analysis is done by varying training data. In addition, the effectiveness of the proposed Taylor-BSA-DBN is analyzed.

4.1 Experimental setup

The implementation of the proposed strategy is carried out in JAVA with jar files libraries using PC having Windows 10 OS, 2GB RAM, and Intel i3 core processor. The simulation setup of the proposed system is depicted in Table 2.

Table 2. Simulation setup

Parameter	Value
Number of input layers	2
Number of hidden layers	2
Number of output layers	1
Cluster size	5 to 9
Number of selected features in Cleveland dataset	123
Number of selected features in Hungarian dataset	139
Number of selected features in Switzerland dataset	139
Learning rate	0.1

4.2. Dataset description

The experimentation is done using Cleveland, Hungarian, and Switzerland dataset taken from healthcare data based on UCI machine learning Repository [28] and which is commonly used for both detection and classification. The Cleveland database is taken from the Cleveland Clinical Foundation contributed by David W. Aha. The Hungarian dataset is obtained from the Hungarian Institute of Cardiology. The Switzerland dataset is obtained from the University Hospital, Basel, Switzerland. The dataset comprises ~~of~~ 303 ~~numbers of~~ instances, and 75 attributes, from which 13 attributes are employed for experimentation, and the dataset is characterized as multivariate with integer and real attributes. The attributes (features), like resting blood pressure (trestbps), maximum heart rate achieved (thalach), the slope of the peak exercise ST segment (slope), Age (age), Sex (sex), fasting blood sugar (fbs), ST depression induced by exercise relative to rest (oldpeak),Chest pain (cp), serum cholesterol (chol), exercise-induced angina (exang), resting electrocardiographic

results (restecg), number of major vessels (0-3) colored by flourosopy (ca), and 3 = normal; 6 = fixed defect; 7 = reversible defect (thal).

4.3. Evaluation Metrics

The performance of proposed Taylor-BSA-DBN is employed for analyzing the methods includes the accuracy, sensitivity and specificity.

4.3.1. Accuracy: The accuracy is described as the degree of closeness of an estimated value with respect to its original value in optimal medical data classification, and it is represented as,

$$Accuracy = \frac{T^p + T^n}{T^p + T^n + F^p + F^n} \tag{55}$$

where, T^p represent true positive, F^p indicate false positive, T^n indicate true negative and F^n represents false negative, respectively.

4.3.2. Sensitivity: This measure is described as the ratio of positives that are correctly identified by the classifier and it is represented as,

$$Sensitivity = \frac{T^p}{T^p + F^n} \tag{56}$$

4.3.3. Specificity: This measure is defined as the ratio of negatives that are correctly identified by the classifier and is formulated as.

$$Specificity = \frac{T^n}{T^n + F^p} \tag{57}$$

4.4 ~~Comparative~~ Methods:

The methods employed for the analysis include: Support vector machine (SVM) [29], Naive Bayes (NB) [30], DBN [27] and proposed Taylor-BSA-DBN.

4.5. Comparative analysis

The analysis of proposed Taylor-BSA-DBN with the conventional methods with accuracy, sensitivity, and specificity parameters is evaluated. The analysis is performed by varying the training data using Cleveland, Hungarian and Switzerland databases.

4.5.1. Analysis with cluster size=5

The analysis of methods considering cluster size=5 using Cleveland, Hungarian and Switzerland databases are specified below:

a) Analysis considering Cleveland database

Table 3 elaborates the analysis of methods using the Cleveland database considering training data with accuracy, sensitivity, and specificity parameters. The maximum accuracy, sensitivity, and specificity are considered as the best performance. Here, the proposed system offers better performance than the existing methods, such as SVM, NB, and DBN, respectively.

Table 3. Analysis of methods with cluster size=5 using Cleveland database

Methods / Training Percentage	SVM	NB	DBN	Proposed Taylor-BSA-DBN
Accuracy				
50	0.7590	0.7603	0.7874	0.8625
60	0.7143	0.7682	0.7851	0.8632
70	0.7460	0.7627	0.8122	0.8531
80	0.7236	0.7619	0.7869	0.8644
90	0.7538	0.7647	0.7742	0.8710
Sensitivity				
50	0.7535	0.7613	0.7908	0.8693
60	0.7120	0.7611	0.7886	0.8699
70	0.7473	0.7558	0.8172	0.8602
80	0.7167	0.7656	0.7903	0.8710
90	0.7576	0.7667	0.7714	0.8788
Specificity				
50	0.7566	0.7667	0.7838	0.8551
60	0.7165	0.7750	0.7815	0.8559
70	0.7447	0.7692	0.8068	0.8452
80	0.7302	0.7581	0.7833	0.8571
90	0.7500	0.7576	0.7813	0.8621

b) Analysis considering Hungarian database

Table 4 elaborates the analysis of methods using the Hungarian database considering training data with accuracy, sensitivity, and specificity parameters. The proposed system offers the best performance when considering 90% of training data.

Table 4. Analysis of methods with cluster size=5 using Hungarian database

Methods / Training Percentage	SVM	NB	DBN	Proposed Taylor-BSA-DBN
Accuracy				
50	0.7810	0.8043	0.8428	0.9200
60	0.7906	0.7976	0.8595	0.8907
70	0.7674	0.8143	0.8182	0.8551
80	0.7576	0.8095	0.8710	0.8710
90	0.6957	0.7500	0.7647	0.9130
Sensitivity				
50	0.8160	0.8456	0.8776	0.9388
60	0.8300	0.8347	0.8908	0.9160
70	0.8052	0.8539	0.8571	0.8876
80	0.8065	0.8400	0.8795	0.9000
90	0.7500	0.8000	0.8125	0.9333
Specificity				
50	0.7294	0.7326	0.7805	0.8846
60	0.7143	0.7500	0.8030	0.8438
70	0.7115	0.7451	0.7500	0.7959
80	0.6757	0.7647	0.8182	0.8182
90	0.6111	0.6667	0.6842	0.8750

c) Analysis considering Switzerland database

Table 5 elaborates the analysis of methods using the Switzerland database considering training data with accuracy, sensitivity, and specificity parameters. The better performance of the proposed system with values is 0.8462, 0.8571, and 0.8333, for the performance metrics, such as accuracy, sensitivity, and specificity.

Table 5. Analysis of methods with cluster size=5 using Switzerland database

Methods / Training Percentage	SVM	NB	DBN	Proposed Taylor-BSA-DBN
Accuracy				
50	0.7619	0.7710	0.7895	0.8644
60	0.7009	0.7374	0.7800	0.8557
70	0.7073	0.7568	0.7895	0.8904
80	0.7551	0.7647	0.7778	0.8400
90	0.6774	0.7037	0.7143	0.8462

Sensitivity				
50	0.7656	0.7742	0.7818	0.8710
60	0.6981	0.7292	0.7843	0.8627
70	0.7073	0.7500	0.7948	0.8974
80	0.7500	0.7692	0.7857	0.8462
90	0.6875	0.6923	0.7143	0.8571
Specificity				
50	0.7581	0.7667	0.7966	0.8571
60	0.7037	0.7451	0.7755	0.8478
70	0.7073	0.7632	0.7838	0.8823
80	0.7600	0.7600	0.7692	0.8333
90	0.6667	0.7143	0.7143	0.8333

4.5.2. Analysis with cluster size=9

The analysis of methods considering cluster size=9 using Cleveland, Hungarian, and Switzerland databases are specified below:

a) Analysis considering Cleveland database

Table 6 depicts the analysis of methods using the Cleveland database considering training data with accuracy, sensitivity, and specificity parameters. The maximum accuracy, sensitivity, and specificity are considered as the best performance. Here, the proposed system offers better performance than the existing methods, such as SVM, NB, and DBN, respectively.

Table 6. Analysis of methods with cluster size=9 using Cleveland database

Methods / **Training Percentage**	**SVM**	**NB**	**DBN**	**Proposed Taylor-BSA-DBN**
Accuracy				
50	0.7590	0.7603	0.7993	0.8690
60	0.7166	0.7430	0.7851	0.8632
70	0.7354	0.7363	0.8123	0.8857
80	0.7419	0.7607	0.7619	0.8475
90	0.7460	0.7910	0.8710	0.9016
Sensitivity				
50	0.7535	0.7613	0.8039	0.8758
60	0.7107	0.7440	0.7886	0.8699
70	0.7303	0.7368	0.8172	0.8925

80	0.7419	0.7544	0.7656	0.8548
90	0.7419	0.8000	0.8788	0.9091
Specificity				
50	0.7566	0.7667	0.7945	0.8613
60	0.7222	0.7419	0.7815	0.8559
70	0.7340	0.7419	0.8068	0.8780
80	0.7419	0.7581	0.7667	0.8393
90	0.7500	0.7813	0.8621	0.8929

b) Analysis considering Hungarian database

Table 7 shows the analysis of methods using the Hungarian database considering training data with accuracy, sensitivity, and specificity parameters. The proposed system offers the best performance when considering 90% of training data.

Table 7. Analysis of methods with cluster size=9 using Hungarian database

Methods / **Training Percentage**	**SVM**	**NB**	**DBN**	**Proposed Taylor-BSA-DBN**
Accuracy				
50	0.7500	0.7957	0.8515	0.9200
60	0.7513	0.7870	0.8075	0.8907
70	0.7777	0.8273	0.8500	0.9118
80	0.7755	0.8298	0.8974	0.9341
90	0.7674	0.8000	0.8298	0.8696
Sensitivity				
50	0.7907	0.8389	0.8844	0.9388
60	0.8017	0.8218	0.8487	0.9160
70	0.8242	0.8652	0.8732	0.9326
80	0.8226	0.8667	0.9130	0.9500
90	0.8077	0.8438	0.8667	0.9000
Specificity				
50	0.6897	0.7209	0.7927	0.8846
60	0.6667	0.7353	0.7353	0.8438
70	0.6981	0.7600	0.8163	0.8723
80	0.6944	0.7647	0.8750	0.9032
90	0.7059	0.7222	0.7647	0.8125

c) Analysis considering Switzerland database

Table 8 depicts the analysis of methods using the Switzerland database considering training data with accuracy, sensitivity, and specificity parameters. The better performance of the proposed system with values is 0.7778, 0.7857, and 0.7692, for the performance metrics, such as accuracy, sensitivity, and specificity.

Table 8. Analysis of methods with cluster size=9 using Switzerland database

Methods / Training Percentage	SVM	NB	DBN	Proposed Taylor-BSA-DBN
Accuracy				
50	0.7460	0.7479	0.8017	0.8644
60	0.7170	0.7624	0.7684	0.8947
70	0.7368	0.7500	0.7662	0.8904
80	0.6786	0.7551	0.8200	0.8400
90	0.7333	0.7600	0.7679	0.7778
Sensitivity				
50	0.7414	0.7500	0.8065	0.8710
60	0.7170	0.7609	0.7647	0.9020
70	0.7297	0.7561	0.7692	0.8974
80	0.6786	0.7500	0.8300	0.8462
90	0.7300	0.7500	0.7667	0.7857
Specificity				
50	0.7419	0.7541	0.7966	0.8571
60	0.7170	0.7600	0.7755	0.8864
70	0.7436	0.7436	0.7632	0.8824
80	0.6786	0.7600	0.8200	0.8333
90	0.7143	0.7300	0.7556	0.7692

4.5.3. Analysis based on ROC

Table 9 depicts the comparative analysis based on ROC using Cleveland, Hungarian, and Switzerland database. In Cleveland dataset, when the false positive rate (FPR) is 5, the corresponding true positive rate (TPR) of the methods, such as SVM, NB, DBN, and Proposed Taylor-BSA-DBN is 0.8857, 0.9119, 0.9535, and 0.9684, respectively. By considering the Hungarian dataset, when the FPR is 4, the corresponding TPR of the proposed method is a maximum of 0.9348. For the same FPR, the TPR of the methods, such as SVM, NB, and DBN is 0.9030, 0.9130, and 0.9233, respectively. By considering the Switzerland dataset, when the FPR is 6, the TPR of the methods, such as SVM, NB,

DBN, and Proposed Taylor-BSA-DBN is 0.9105, 0.9443, 0.9569, and 0.9794, respectively.

Table 9. Analysis based on ROC

Methods	SVM	NB	DBN	Proposed Taylor-BSA-DBN
FPR	TPR			
Cleveland				
1	0	0	0	0
2	0.7913	0.7949	0.8429	0.8761
3	0.7961	0.8330	0.8523	0.8798
4	0.8462	0.8753	0.9149	0.9284
5	0.8857	0.9119	0.9535	0.9684
6	0.9153	0.9569	0.9788	0.9847
7	0.9710	0.9783	0.9895	0.9975
8	0.9952	0.9989	1	1
9	1	1	1	1
10	1	1	1	1
Hungarian				
1	0	0	0	0
2	0.8233	0.8410	0.8553	0.8941
3	0.8286	0.8647	0.8734	0.8953
4	0.9030	0.9130	0.9233	0.9348
5	0.9246	0.9417	0.9596	0.9789
6	0.9521	0.9697	0.9803	0.9999
7	0.9793	0.9800	0.9946	1
8	0.9981	0.9985	1	1
9	1	1	1	1
10	1	1	1	1
Switzerland				
1	0	0	0	0
2	0.7593	0.7682	0.8024	0.8258
3	0.7620	0.7923	0.8399	0.8682
4	0.8452	0.8735	0.8781	0.9101
5	0.8725	0.9184	0.9194	0.9564
6	0.9105	0.9443	0.9569	0.9794
7	0.9701	0.9722	0.9865	0.9924
8	0.9946	0.9953	0.9994	1
9	1	1	1	1

10	1	1	1	1

4.6 Comparative Discussion

Table 10 portrays the analysis of methods using accuracy, sensitivity, and specificity parameter with varying training data. The analysis is done with Cleveland, Switzerland, and Hungarian databases. Using cluster size=5, and considering the Cleveland database, the proposed Taylor-BSA-DBN showed maximal accuracy of 0.871, which is 13.43%, 12.17%, and 11.14%, better than the existing methods, such as SVM, NB, and DBN, respectively. In the existing methods, the DBN offers maximum sensitivity of 0.771, but the proposed method is 12.29% better than the existing DBN. The proposed method has a maximum specificity of 0.862. The percentage of improvement of the proposed method with the existing methods, such as SVM, NB, and DBN is 12.99%, 12.06%, and 9.40%, respectively. Considering the Hungarian database, the proposed Taylor-BSA-DBN showed maximal accuracy of 0.913, maximal sensitivity of 0.933, and maximal specificity of 0.875. Considering the Switzerland database, the proposed Taylor-BSA-DBN showed maximal accuracy of 0.846, which is 19.98%, 16.78%, and 15.60% better than the existing methods, such as SVM, NB, and DBN, respectively. Similarly, the proposed system has a maximum sensitivity of 0.857. The percentage of improvement of the proposed system sensitivity, with the existing methods, such as SVM, NB, and DBN is 19.72%, 19.25%, and 16.69%, respectively. Likewise, the proposed Taylor-BSA-DBN showed maximal specificity of 0.833.

Using cluster size=9, and considering the Cleveland database, the proposed Taylor-BSA-DBN showed maximal accuracy of 0.934, which is 16.92%, 11.13%, and 3.96%, better than the existing methods, such as SVM, NB, and DBN, respectively. In the existing methods, the DBN offers maximum sensitivity of 0.913, but the proposed method is 3.89% better than the existing DBN. The proposed method has a maximum specificity of 0.903. The percentage of improvement of the proposed method with the existing methods, such as SVM, NB, and DBN is 23.15%, 15.28%, and 3.10%, respectively. Considering the Hungarian database, the proposed Taylor-BSA-DBN showed maximal accuracy of 0.902, maximal sensitivity of 0.909, and maximal specificity of 0.893. Considering the Switzerland database, the proposed Taylor-BSA-DBN showed maximal accuracy of 0.840, which is 19.17%, 10.12%, and 2.38%, better than the existing methods, such as SVM, NB, and DBN, respectively. Similarly, the proposed system has a maximum sensitivity of 0.846. The percentage of improvement of the proposed system sensitivity with the existing methods, such as SVM, NB, and DBN is 19.74%, 11.35%, and 1.89%, respectively. Likewise, the proposed Taylor-BSA-DBN showed maximal specificity of 0.833.

Table 10. Comparative analysis

Cluster size	Database	Metrics	SVM	NB	DBN	Proposed Taylor-BSA-DBN
Cluster size=5	Cleveland	Accuracy	0.754	0.765	0.774	**0.871**
		Sensitivity	0.758	0.767	0.771	**0.879**
		Specificity	0.750	0.758	0.781	**0.862**
	Hungarian	Accuracy	0.696	0.750	0.765	**0.913**
		Sensitivity	0.750	0.800	0.813	**0.933**
		Specificity	0.611	0.667	0.684	**0.875**
	Switzerland	Accuracy	0.677	0.704	0.714	**0.846**
		Sensitivity	0.688	0.692	0.714	**0.857**
		Specificity	0.667	0.714	0.714	**0.833**
Cluster size=9	Cleveland	Accuracy	0.776	0.830	0.897	**0.934**
		Sensitivity	0.823	0.867	0.913	**0.950**
		Specificity	0.694	0.765	0.875	**0.903**
	Hungarian	Accuracy	0.746	0.791	0.871	**0.902**
		Sensitivity	0.742	0.800	0.879	**0.909**
		Specificity	0.750	0.781	0.862	**0.893**
	Switzerland	Accuracy	0.679	0.755	0.820	**0.840**
		Sensitivity	0.679	0.750	0.830	**0.846**
		Specificity	0.679	0.760	0.820	**0.833**

Table 11 shows the computational time of the proposed system and the existing methods, such as SVM, NB, and DBN, in which the proposed Taylor-BSA-DBN has a minimum computation time of 6.31 sec.

Table 11. Computational Time

Methods	SVM	NB	DBN	Proposed Taylor-BSA-DBN
Time (Sec)	10.08	8.79	7.56	**6.31**

Table 12 shows the statistical analysis of the proposed work and the existing methods based on mean and variance.

Table 12. Statistical Analysis

Dataset	Methods	Accuracy	Mean	Variance	Sensitivity	Mean	Variance	Specificity	Mean	Variance
Cluster size = 5										
Cleveland	SVM	0.754	0.752	0.002	0.758	0.754	0.004	0.750	0.748	0.002
	NB	0.765	0.761	0.004	0.767	0.765	0.002	0.758	0.754	0.004
	DBN	0.774	0.771	0.003	0.771	0.768	0.003	0.781	0.779	0.002
	Proposed Method	0.871	0.869	0.002	0.879	0.878	0.001	0.862	0.860	0.002
Hungarian	SVM	0.696	0.693	0.003	0.750	0.748	0.002	0.611	0.608	0.003
	NB	0.750	0.746	0.004	0.800	0.799	0.001	0.667	0.665	0.002
	DBN	0.765	0.763	0.002	0.813	0.810	0.003	0.684	0.682	0.002
	Proposed Method	0.913	0.911	0.002	0.933	0.932	0.001	0.875	0.873	0.002
Switzerland	SVM	0.677	0.675	0.002	0.688	0.684	0.004	0.667	0.665	0.002
	NB	0.704	0.702	0.003	0.692	0.690	0.002	0.714	0.711	0.003
	DBN	0.714	0.711	0.003	0.714	0.713	0.001	0.714	0.712	0.002
	Proposed Method	0.846	0.844	0.002	0.857	0.855	0.002	0.833	0.831	0.002
Cluster size = 9										
Cleveland	SVM	0.776	0.773	0.003	0.823	0.822	0.001	0.694	0.691	0.003
	NB	0.830	0.826	0.004	0.867	0.865	0.002	0.765	0.761	0.004
	DBN	0.897	0.895	0.002	0.913	0.911	0.002	0.875	0.873	0.002
	Proposed Method	0.934	0.932	0.002	0.950	0.948	0.002	0.903	0.901	0.002
Hungarian	SVM	0.746	0.743	0.003	0.742	0.740	0.002	0.750	0.748	0.002
	NB	0.791	0.790	0.001	0.800	0.797	0.003	0.781	0.780	0.001
	DBN	0.871	0.868	0.003	0.879	0.878	0.001	0.862	0.860	0.002
	Proposed Method	0.902	0.900	0.002	0.909	0.907	0.002	0.893	0.891	0.002
Switzerland	SVM	0.679	0.677	0.002	0.679	0.675	0.004	0.679	0.677	0.002
	NB	0.755	0.752	0.003	0.750	0.748	0.002	0.760	0.758	0.002
	DBN	0.820	0.818	0.002	0.830	0.827	0.003	0.820	0.819	0.001
	Proposed Method	0.840	0.838	0.002	0.846	0.844	0.002	0.833	0.832	0.001

5. Conclusion

The contemporary medicines produce a huge amount of information contained in the medical database. The obtainability of large medical data leads to the requirement of effective data analysis tools for extracting constructive knowledge. This paper proposes a novel fully automated DBN for medical data classification using medical data. The proposed Taylor-BSA is employed to train DBN. The proposed Taylor-BSA is designed by combining the Taylor series and BSA algorithm, which can be utilized for finding the optimal weights for establishing effective medical data classification. Here, the sparse-FCM is employed for selecting significant features. The incorporation

of sparse FCM for the feature selection process provides more benefits for interpreting the models, as this sparse technique provides important features for the classification and can be utilized for handling high dimensional data. The obtained selected features are fed to DBN which is trained by proposed Taylor-BSA. The proposed Taylor-BSA is designed by integrating the Taylor series and BSA in order to generate optimal weights for classification. The proposed Taylor-BSA-DBN outperformed other methods with maximal accuracy of 93.4%, maximal sensitivity of 95%, and maximal specificity of 90.3%, respectively. The proposed method does not classify the type of heart disease. In the future, other medical data classification datasets will be employed for computing efficiency of the proposed method. In addition, the proposed system will be further improved to classify the heart diseases, like Congenital heart disease, Coronary artery disease, and Arrhythmia.

References

1. Ranganatha, S., Raj, H.P., Anusha, C. and Vinay, S.K., "Medical data mining and analysis for heart disease dataset using classification techniques," In proceedings of National Conference on Challenges in Research & Technology in the Coming Decades (CRT), Ujire, India, 2013.
2. Babič, F., Olejár, J., Vantová, Z. and Paralič, J., "Predictive and descriptive analysis for heart disease diagnosis," In proceedings of Federated Conference on Computer Science and Information Systems (FedCSIS), pp. 155-163, 2017.
3. Jabbar, M.A., Deekshatulu, B.L. and Chandra, P., "Heart disease classification using nearest neighbor classifier with feature subset selection," Anale. Seria Informatica, vol.11, pp.47-54, 2013.
4. Alzahani, S.M., Althopity, A., Alghamdi, A., Alshehri, B. and Aljuaid, S, "An overview of data mining techniques applied for heart disease diagnosis and prediction," Lecture Notes on Information Theory, Taif University, Taif, Saudi Arabia, vol.2, no.4, pp. 310-315, 2014.
5. S. B. Patil and Y. S. Kumaraswamy, "Extraction of significant patterns from heart disease warehouses for heart attack prediction," International Journal of Computer Science and Networks Security, vol. 9, pp. 228-235, 2009.
6. S. Oyyathevan and A. Askarunisa, "An expert system for heart disease prediction using data mining technique: Neural network," International Journal of Engineering Research and Sports Science, vol. 1, pp. 1-6, 2014.
7. S. Palaniappan and R. Awang, "Intelligent heart disease prediction system using data mining techniques," In proceedings of International Conference on Computer Systems and Applications, Doha, Qatar, 2008.
8. G. Subbalakshmi, "Decision support in heart disease prediction system using naive bayes," Indian Journal of Computer Science and Engineering, vol. 2, no. 2, pp. 170-174, 2011

9. Shouman, M., Turner, T. and Stocker, R., "Using data mining techniques in heart disease diagnosis and treatment," In proceedings of IEEE Japan-Egypt Conference on Electronics, Communications and Computers, Alexandria, Egypt, pp. 173-177, 2012.
10. Ghumbre, S.U. and Ghatol, A.A., "Heart Disease Diagnosis Using Machine Learning Algorithm," In Proceedings of the International Conference on Information Systems Design and Intelligent Applications, Springer, Visakhapatnam, India, vol. 132, pp. 217-225, 2012.
11. Yeh, D.-Y., Cheng, C.-H., and Chen, Y.-W."A predictive model for cerebrovascular disease using data mining," Expert Systems with Applications, vol. 38, no. 7, pp. 8970-8977, 2011.
12. Alizadehsani, R., Habibi, J., Hosseini, M. J., Mashayekhi, H., Boghrati, R., Ghandeharioun, "A data mining approach for diagnosis of coronary artery disease,"Computer Methods and Programs in Biomedicine, vol. 11, no. 1, pp. 52-61, 2013.
13. Giri, D.and Acharya, U. R,. "Automated diagnosis of Coronary Artery Disease affected patients usingLDA, PCA, ICA and Discrete WaveletTransform," Knowledge-Based Systems, vol. 37, pp. 274-282, 2013.
14. Jabbar,M.A,Deekshatulu,B.LandChandra,Priti ,"Classification of Heart Disease Using K- Nearest Neighbor and Genetic Algorithm,"Procedia Technology, vol. 10, pp. 85– 94, 2013.
15. Reddy, G.T., Reddy, M.P.K., Lakshmanna, K., Rajput, D.S., Kaluri, R. and Srivastava, G., "Hybrid genetic algorithm and a fuzzy logic classifier for heart disease diagnosis," Evolutionary Intelligence, vol. 13, pp.185-196, 2019.
16. Nourmohammadi-Khiarak, J., Feizi-Derakhshi, M.R., Behrouzi, K., Mazaheri, S., Zamani-Harghalani, Y. and Tayebi, R.M., "New hybrid method for heart disease diagnosis utilizing optimization algorithm in feature selection," Health and Technology, vol. 10, pp. 667-678, 2019.
17. Magesh, G. and Swarnalatha, P., "Optimal feature selection through a cluster-based DT learning (CDTL) in heart disease prediction," Evolutionary Intelligence, pp.1-11, 2020.
18. Thiyagaraj, M. and Suseendran, G., "Enhanced Prediction of Heart Disease Using Particle Swarm Optimization and Rough Sets with Transductive Support Vector Machines Classifier," In Data Management, Analytics and Innovation, Springer, vol. 2, pp. 141-152, 2020.
19. Abdel-Basset, M., Gamal, A., Manogaran, G. and Long, H.V., "A novel group decision making model based on neutrosophic sets for heart disease diagnosis," Multimedia Tools and Applications, vol. 79, pp. 9977–10002, 2019.
20. Nilashi, M., Ahmadi, H., Manaf, A.A., Rashid, T.A., Samad, S., Shahmoradi, L., Aljojo, N. and Akbari, E, "Coronary Heart Disease Diagnosis Through Self-

Organizing Map and Fuzzy Support Vector Machine with Incremental Updates," International Journal of Fuzzy Systems, vol. 23, pp. 1376–1388, 2020.

21. Shah, S.M.S., Shah, F.A., Hussain, S.A. and Batool, S., "Support Vector Machines-based Heart Disease Diagnosis using Feature Subset, Wrapping Selection and Extraction Methods," Computers & Electrical Engineering, vol. 84, pp. 106628, 2020.
22. Acharjya, D.P., "A Hybrid Scheme for Heart Disease Diagnosis Using Rough Set and Cuckoo Search Technique," Journal of Medical Systems, vol. 44, no. 27, 2020.
23. G.-J. Ahn, H. Hu, J. Lee, and Y. Meng, "Representing and Reasoning about Web Access Control Policies," In proceediong of IEEE 34th Annual Computer Software and Applications Conference, Seoul, South Korea, pp. 137–146, 2010.
24. Chang X, Wang Q, Liu Y and Wang Y, "Sparse Regularization in Fuzzy c-Means for High-Dimensional Data Clustering," IEEE Transactions on Cybernetics, vol. 47, no. 9, pp. 2616 – 2627, 2017.
25. Mangai SA, Sankar BR and Alagarsamy K, "Taylor Series Prediction of Time Series Data with Error Propagated by Artificial Neural Network," International Journal of Computer Applications, vol.89, no.1, 2014.
26. Meng X, Gao XZ, Lu L, Liu Y and Zhang H, "A new bio-inspired optimisation algorithm: Bird Swarm Algorithm," Journal of Experimental & Theoretical Artificial Intelligence, vol. 28, no. 4, pp. 673-687, 2016.
27. Mannepalli K, Sastry PN, and Suman M, "A novel Adaptive Fractional Deep Belief Networks for speaker emotion recognition," Alexandria Engineering Journal, vol. 56, no. 4, 2016.
28. Heart Disease Data Set, "http://archive.ics.uci.edu/ml/datasets/Heart+Disease" Accessed on April 2020.
29. Ghumbre, S., Patil, C. and Ghatol, A., "Heart disease diagnosis using support vector machine," In proceedings of International conference on computer science and information technology (ICCSIT), Mumbai, India, 2011.
30. Medhekar, D.S., Bote, M.P. and Deshmukh, S.D., "Heart disease prediction system using naive Bayes," International Journal of Enhanced Research in Science Technology & Engineering (IJERSTE), vol.2, no.3, 2013.
31. Fatima, M., and Pasha, M., "Survey of Machine Learning Algorithms for Disease Diagnostic", Journal of Intelligent Learning Systems and Applications, vol. 9, no. 1, 2017.
32. Safdar, S., Zafar, S., Zafar, N., and Khan, N., "Machine learning based decision support systems (DSS) for heart disease diagnosis: a review', Artificial Intelligence Review, 2017.