

Article



Taylor Bird Swarm Algorithm Based on Deep Belief Network for Heart Disease Diagnosis

Afnan M. Alhassan ^{1,2,*} and Wan Mohd Nazmee Wan Zainon ¹

- ¹ School of Computer Science, Universiti Sains Malaysia, George Town 11800, Malaysia; nazmee@usm.my
- ² College of Computing and Information Technology, Shaqra University, Shaqra 11961, Saudi Arabia
- * Correspondence: afnan@student.usm.my

Received: 27 July 2020; Accepted: 18 September 2020; Published: 22 September 2020



Abstract: Contemporary medicine depends on a huge amount of information contained in medical databases. Thus, the extraction of valuable knowledge, and making scientific decisions for the treatment of disease, has progressively become necessary to attain effective diagnosis. The obtainability of a large amount of medical data leads to the requirement of effective data analysis tools for extracting constructive knowledge. This paper proposes a novel method for heart disease diagnosis. 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 significant features to classify medical data. Incorporating sparse FCM for the feature selection process provides more benefits for interpreting the models, as this sparse technique provides important features for detection, 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 detection. Here, 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: deep belief network; heart disease diagnosis; sparse FCM; bird swarm algorithm

1. Introduction

Contemporary medicine depends on a large amount of information accumulated in medical datasets. The extraction of such constructive knowledge can help when making scientific decisions to diagnose disease. Medical data can enhance the management of hospital information and endorse the growth of telemedicine. Medical data primarily focuses on patient care first, and research resources second. The main rationalization to collect medical data is to promote patient health conditions [1]. The accessibility of numerous medical data causes redundancy, which requires effectual and significant techniques for processing data to extract beneficial knowledge. However, the diagnostics of various diseases indicate significant issues in data analysis [2]. Quantifiable diagnosis is performed by adoctor's guidance rather than patterns of the medical dataset; thus, there is the possibility of incorrect diagnosis [3]. Cloud-based services can assist with managing medical data, including compliance management, policy integration, access controls, and identity management [4].

Now a day, heart disease is a foremost source of death. We are moving towards a new industrial revolution; thus, lifestyle changes should take place to prevent risk factors of heart disease, such as obesity, diabetes, hypertension, and smoking [5]. The treatment of disease is a complex mission in medical field. The discovery of heart disease, with different risk factors, is considered a multi-layered issue [6]. Thus, patient medical data are collected to simplify the diagnosis process. Offering a valuable service (at less cost) is a major limitation in the healthcare industry. In [7], valuable quality service refers

to the precise diagnosis (and effective treatment) in patients. Poor clinical decisions cause disasters, which may affect the health of patients. Automated approaches, such as the machine-learning approach [8,9] and data mining [10] approach, assist with attaining clinical tests, or diagnoses, at reduced risks [11,12]. The classification and pattern recognition by machine learning algorithms are 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 clinical decision support systems (CDSS) [13,14].

Several methods are devised for automatic heart disease detection to evaluate the efficiency of the decision tree and Naive Bayes [15]. Moreover, optimization with the genetic algorithm is employed for minimizing the number of attributes without forfeiting accuracy and efficiency to diagnose heart disease [16]. Data mining methods for heart disease diagnosis include the bagging algorithm, neural network, support vector machine, and automatically defined groups [17]. In [18], the study acquired 493 samples from a cerebrovascular disease prevention program, and utilized three classification techniques (the Bayesian classifier, decision tree, and backpropagation neural network) for constructing classification models. In [19], a method is devised for diagnosing coronary artery disease. The method utilized 303 samples by adapting the feature creation technique. In [20], a methodology is devised for automatically detecting the efficiency of features to reveal heart rate signals. In [21], a hybrid algorithm is devised with K-Nearest Neighbour (KNN), and the genetic algorithm for effectual classification. The method utilized a genetic search as a decency measure for ranking attributes. Then, the classification algorithm was devised on evaluated attributes for heart disease diagnosis. The extraction of valuable information from huge data is a time-consuming task [22]. The size of the medical dataset is increasing in a rapid manner and advanced techniques of data mining help physicians make effective decisions. However, the issues of heart disease data involve feature selection, in which the imbalance of samples and the lack of magnitude of features are just some of the issues [23]. Although there are methods for heart disease detection with real-world medical data, these methods are devised to improve accuracy and time for computation in disease detection [24]. In [25], a hybrid model with the cuckoo search (CS)—and a rough set—is adapted for diagnosing heart disease. The drawback is that a rough set produces an unnecessary number of rules. To solve these challenges in heart disease diagnoses; a novel method, named the Taylor-based bird swarm algorithm-deep belief network (Taylor-BSA–DBN), is proposed for medical data classification.

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

The major portion of the paper focuses on:

 Proposed Taylor-BSA–DBN for heart disease diagnosis:Taylor-BSA–DBN(a classifier) is 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 Taylor-BSA–DBN is adapted for heart disease diagnosis.

Other sections of the paper are arranged as follows: Section 2 elaborates the descriptions of conventional heart disease detection strategies utilized in the literature, as well as challenges faced, which are considered as the inspiration for developing the proposed technique. The proposed method for heart disease diagnosis using modified DBN is portrayed in Section 3. The outcomes of the proposed strategy with other methods are depicted in Section 4; Section 5 presents the conclusion.

2. Motivations

This section illustrates eight strategies employed for heart disease diagnosis, along with its challenges.

Literature Survey

Reddy, G.T. et al. [22] devised an adaptive genetic algorithm with fuzzy logic (AGAFL) model for predicting heart disease, which assists clinicians in 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 are optimized by adapting an adaptive genetic algorithm. Initially, the significant features that affect heart disease are chosen using the rough set theory. Then, the second step predicts heart disease with the AGAFL classifier. The method is effective in handling noisy data and works effectively with large attributes. Nourmohammadi-Khiarak et al. [23] devised a method for selecting features and reducing the number of features.

Here, the imperialist competitive algorithm was devised to choose important features from heart disease. This algorithm offers an optimal response in selecting features. Moreover, the k-nearest neighbor 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. [26] devised a model using Cleveland heart samples for heart disease diagnosis. The method employed cluster-based Decision Tree 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. [27] 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, reduce computational overhead, and enhance prediction performance. The Radial Basis Function-Transductive Support Vector Machines (RBF-TSVM) classifier was employed for heart disease prediction. Abdel-Basset, M. et al. [28] devised a model using Internet of Things (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 (i.e., for observing patients suffering from heart failure). Moreover, the model provided an accurate solution that decreases the rate of death and the cost of treatment. Nilashi, M. et al. [24] 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 (FSVM), 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. Shah, S.M.S. et al. [29] 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: 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. [25] devised a hybrid method for diagnosing heart disease. The method combined the 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 feasibility, but failed to induce an intuitionistic fuzzy rough set and CS for diagnosing heart disease.

3. Proposed Taylor-BSA–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. 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 tend 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 it has some powerful methods. Thus, the proposed Taylor-BSA–DBN is devised to process medical data for attaining effective heart disease diagnosis. Figure 1 portrays the schematic view of the proposed Taylor-BSA–DBN for heart disease diagnosis. The complete process of the proposed model is pre-processing feature selection, and detection. 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 are 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 detection 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 the proposed Taylor-based bird swarm algorithm (Taylor-BSA)–deep belief network (DBN) for heart disease diagnosis.

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

$$A = \{A_{G,H}\}; (1 \le G \le B); (1 \le H \le C)$$
(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 data are replaced with a log function, wherein the base of the log is set by the analyst (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 are fed to the feature selection module, considering the Sparse FCM algorithm [30], which is the modification of the standard FCM. The benefit of using Sparse FCM is to provide high dimensional data clustering. The pre-processed data contain different types of attributes, each indicating individual value. In the medical data classification strategy, the sparse FCM is applied for determining the features from the 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, poses the ability to handle disease diagnosis without delay, and is easier with optimization techniques.

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. The Taylor series [31] 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 aone-step process, and it can deal with higher-order terms. The 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. Moreover, it is a simple method to solve complex functions. BSA [32] is duly based on the social behaviors of birds that follow some idealistic rules. BSA is more accurate than other standard optimizations with highly efficient, accurate, and robust performances. 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. Moreover, 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, including: $F_{i,j}$; $(1 \le i \le j)$, 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 behavior of birds is expressed as *Ft*.

Step 2. Determination of objective function: the selection of the best position of the 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 behavior, or else the vigilance behavior commences. On the other hand, the swarm splits as scroungers and producers, which is modeled as flight behaviors. Finally, the feasibility of the solutions is verified and the best solution is retrieved.

Step 4. Foraging behavior of birds: the individual bird searches for the food based on its own experience, and the behavior of the swarm, which is given below. The standard equation of the foraging behavior of birds [32] is given by,

$$F_{i,j}^{h+1} = F_{i,j}^{h} - F_{i,j}^{h} Rand(0,1)[Z+T] + Rand(0,1)[P_{i,j}Z+Y_{j}T]$$
(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, $P_{i,j}$ 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 the Taylor series [31], the update equation is expressed as,

$$F_{i,j}^{h+1} = 0.5F_{i,j}^{h} + 1.3591F_{i,j}^{h-1} - 1.359F_{i,j}^{h-2} + 0.6795F_{i,j}^{h-3} - 0.2259F_{i,j}^{h-4} + 0.0555F_{i,j}^{h-5} - 0.0104F_{i,j}^{h-6} + 1.38e^{-3}F_{i,j}^{h-7} - 9.92e^{-5}F_{i,j}^{h-8}$$
(4)

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

Substituting Equation (5) in Equation (3),

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

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

$$(7)$$

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

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

Step 5. Vigilance Behavior of Birds: the birds move towards the center, during which, the birds compete with each other; the vigilance behavior of birds is modeled as,

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

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

$$V_2 = w_2 \times \exp\left[\left(\frac{RQ(U)_i - RQ(U)_T}{\left|RQ(U)_T - RQ(U)_i\right| + \psi}\right) \frac{v \times RQ(U)_T}{\sum RQ + \psi}\right]$$
(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 Behavior: this behavior is of the birds' 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. Some birds in the group act as producers and others as scroungers. The behavior is modeled as,

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

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

$$\tag{14}$$

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 better 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 Algorithm 1.

Algorithm 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 behavior of birds Determine the fitness of the solutions While $h < h_{max}$ For k = 1 : bIf Rand(0,1) < probForaging behavior using Equation (3) Else Vigilance behavior using Equation (12) End if End for Else Split the swarm as scroungers and producers For k = 1 : bIf *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 + 1End while Optimal solution is obtained End

3.3.2. Architecture of Deep Belief Network

The DBN [33] is a subset of Deep Neural Network (DNN) and comprises different layers of Multilayer Perceptrons (MLPs) and Restricted Boltzmann Machines (RBMs). RBMs comprise of visible

and hidden units that are associated with weights. The basic structural design of the DBN is illustrated in Figure 2.



Figure 2. Architectural view diagram of DBN classifier.

Training of Deep Belief Network

This section elaborates on the training process of the proposed Taylor-BSA–DBN classifier. A 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. The probability of reconstruction of hidden neurons is obtained by resampling the hidden states.
 - 6. The negative gradient is computed.
 - 7. Weights are updated by subtracting the negative gradient from the positive gradient.
 - 8. Weights are updated for the next iteration, using the steepest or gradient descent algorithm.
 - 9. Energy is calculated 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 a backpropagation approach by feeding the training data, which are 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 initialize the weights.

- 2. Read the input sample from the result of the preceding layer.
- 3. Obtain the average error, 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. Obtain the new weights from the hidden and the visible layers by applying gradient descent.
- 6. Identify the new weights using the updated 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 using Java libraries via Java Archive (JAR) files, utilizing a PC, Windows 10 OS, 2GB RAM, and an Intel i3 core processor. The simulation setup of the proposed system is depicted in Table 1.

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

Table 1.	Simulation	setup.
----------	------------	--------

4.2. Dataset Description

The experimentation is done using Cleveland, Hungarian, and Switzerland datasets taken from healthcare data based on University of California Irvine (UCI) machine learning repository [34], 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 number of instances and 75 attributes, ofwhich, 13 attributes are employed for experimentation. Furthermore, the dataset is characterized as multivariate with integer and real attributes. The attributes (features), such asresting 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 fluoroscopy (ca), and 3 = normal; 6 = fixed defect; 7 = reversible defect (thal).

4.3. Evaluation Metrics

The performance of the proposed Taylor-BSA–DBN is employed for analyzing the methods, including 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,

$$Accuracy = \frac{T^p + T^n}{T^p + T^n + F^p + F^n}$$
(15)

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,

$$Sensitivity = \frac{T^p}{T^p + F^n}$$
(16)

4.3.3. Specificity

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

$$Specificity = \frac{T^n}{T^n + F^p}$$
(17)

4.4. Comparative Methods

The methods employed for the analysis include the Support Vector Machine (SVM) [35], Naive Bayes (NB) [36], DBN [33], and the proposed Taylor-BSA–DBN.

4.5. Comparative Analysis

The analysis of the 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:

Analysis Considering Cleveland Database

Table 2 elaborates the analysis of methods using the Cleveland database, considering training data with accuracy, sensitivity, and specificity parameters. The maximum accuracy, sensitivity, and specificity is considered as the best performance. Here, the proposed system offers better performances than the existing methods, such as SVM, NB, and DBN, respectively.

Methods	SVM	SVM NB DBN		Proposed Taylor-BSA-DBN				
Training Percentage	0,111			r				
	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				
		Specifi	city					
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				

Table 2. Analysis of methods with cluster size = 5 using the Cleveland database. Abbreviations: SVM, Support Vector Machine; NB, Naive Bayes; DBN, deep belief network.

Analysis Considering Hungarian Database

Table 3 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.

Methods	SVM	NB	DBN	Proposed Taylor-BSA-DBN			
Training Percentage	5 1 11			110p0000 10000 2011 2 211			
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			
		Specific	city				
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			

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

Analysis Considering Switzerland Database

Table 4 elaborates the analysis of methods using the Switzerland database considering training data with accuracy, sensitivity, and specificity parameters. The better performances of the proposed system, with values, are 0.8462, 0.8571, and 0.8333 for performance metrics, such as accuracy, sensitivity, and specificity.

Methods	SVM	NB	DBN	Proposed Taylor-BSA-DBN			
Training Percentage	5 1 11						
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			
		Specific	city				
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			

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

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:

Analysis Considering Cleveland Database

Table 5 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 performances. Here, the proposed system offers better performance than the existing methods, such as SVM, NB, and DBN, respectively.

Analysis Considering Hungarian Database

Table 6 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.

Methods	SVM	NB	DBN	Proposed Taylor-BSA-DBN				
Training Percentage	5 1 11	ND						
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				
		Specific	city					
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				

Table 5. Analysis of methods with cluster size = 9 using the Cleveland database.

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

Methods	SVM	NB	DBN	Proposed Taylor-BSA-DBN	
Training Percentage	5 • 101				
		Accura	асу		
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	
		Specifi	city		
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	

Analysis Considering Switzerland Database

Table 7 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.

Methods	SVM	NB	DBN	Proposed Taylor-BSA-DBN				
Training Percentage	5 • 101							
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				
		Specifi	city					
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				

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

4.5.3. Analysis Based on Receiver Operating Characteristic (ROC) Curve

Table 8 depicts the comparative analysis based on ROC curve, using Cleveland, Hungarian, and Switzerland databases. In the 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 the 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, and DBN is 0.9105, 0.9443, 0.9569, and 0.9794, respectively.

4.5.4. Analysis Based on k-Fold

Table 9 depicts the comparative analysis based on k-fold using the Cleveland, Hungarian, and Switzerland databases, for cluster size = 5. The Hungarian datasets offer the maximum accuracy of 0.9021, when k-fold = 8. By considering k-fold = 7, the specificity offered by the Cleveland datasets for the methods, such as SVM, NB, DBN, and the proposed Taylor-BSA–DBN, is 0.8032, 0.8189, 0.8256, and 0.8321, respectively. The proposed Taylor-BSA–DBN offers maximum accuracy, sensitivity, and specificity, when considering k-fold = 8.

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.

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					
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					
		Sw	vitzerland						
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					

Table 8. Analysis based on ROC.

Motrice	Methods	SVM	NB	DBN	Proposed Taylor-BSA-DBN			
wienies -	k-Fold	. 3 v ivi	ND	DDN	Tioposed Tuylor Dorr DDrv			
Cleveland								
	5	0.7021	0.7088	0.7126	0.7239			
Accuracy	6	0.7122	0.7189	0.7245	0.7365			
Accuracy	7	0.7345	0.7543	0.7634	0.7843			
	8	0.7528	0.7843	0.7965	0.8132			
	5	0.7567	0.7678	0.7898	0.7956			
Sensitivity	6	0.7834	0.8045	0.8156	0.8232			
benbruity	7	0.8032	0.8189	0.8256	0.8321			
	8	0.8145	0.8229	0.8365	0.8448			
	5	0.7586	0.7656	0.7699	0.7865			
Specificity	6	0.7854	0.7745	0.7965	0.8043			
opeemeny	7	0.7940	0.8088	0.8124	0.8227			
	8	0.8021	0.8178	0.8249	0.8339			
			Hung	arian				
	5	0.7960	0.8758	0.8791	0.8822			
Accuracy	6	0.8071	0.8413	0.8838	0.8854			
Accuracy	7	0.7985	0.8030	0.8324	0.8917			
	8	0.7982	0.8626	0.8948	0.9021			
	5	0.7959	0.8027	0.8197	0.8231			
Sensitivity	6	0.7857	0.7891	0.7925	0.8367			
Sensitivity	7	0.7891	0.7925	0.7959	0.8393			
	8	0.7857	0.7925	0.8061	0.8458			
	5	0.7494	0.7513	0.7645	0.7656			
Specificity	6	0.7146	0.7343	0.7645	0.7760			
opeenieny	7	0.7433	0.7535	0.7645	0.7719			
	8	0.7530	0.7645	0.7697	0.7873			
			Switze	erland				
	5	0.7528	0.7789	0.7799	0.7896			
1 001180 011	6	0.7658	0.7828	0.7925	0.8012			
Accuracy	7	0.7712	0.7958	0.8028	0.8156			
	8	0.7828	0.8128	0.8159	0.8259			
	5	0.7428	0.7689	0.7847	0.7956			
Soncitivity	6	0.7625	0.7758	0.7858	0.8028			
Schonivity	7	0.7748	0.7896	0.7952	0.8125			
	8	0.7828	0.7986	0.8078	0.8225			
	5	0.7658	0.7589	0.7750	0.7896			
Specificity	6	0.7758	0.7832	0.7962	0.8020			
Specificity	7	0.7841	0.7911	0.8025	0.8196			
	8	0.7958	0.8002	0.8178	0.8219			

Table 9. Analysis based on k-fold.

Cluster Size	Database	Metrics	SVM	NB	DBN	Proposed Taylor-BSA-DBN
		Accuracy	0.754	0.765	0.774	0.871
	Cleveland	Sensitivity	0.758	0.767	0.771	0.879
		Specificity	0.750	0.758	0.781	0.862
Cluster size - 5		Accuracy	0.696	0.750	0.765	0.913
Cluster size = 5	Hungarian	Sensitivity	0.750	0.800	0.813	0.933
		Specificity	0.611	0.667	0.684	0.875
		Accuracy	0.677	0.704	0.714	0.846
	Switzerland	Sensitivity	0.688	0.692	0.714	0.857
		Specificity	0.667	0.714	0.714	0.833
		Accuracy	0.776	0.830	0.897	0.934
	Cleveland	Sensitivity	0.823	0.867	0.913	0.950
		Specificity	0.694	0.765	0.875	0.903
Cluster size = 0		Accuracy	0.746	0.791	0.871	0.902
Cluster size = 9	Hungarian	Sensitivity	0.742	0.800	0.879	0.909
		Specificity	0.750	0.781	0.862	0.893
		Accuracy	0.679	0.755	0.820	0.840
	Switzerland	Sensitivity	0.679	0.750	0.830	0.846
		Specificity	0.679	0.760	0.820	0.833

Table 10. Comparative analysis.

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 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. Com	putational	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.

Dataset	Methods	Accuracy	y 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		

Table 12. Statistical Analysis.

5. Conclusions

Contemporary medicine depends on a huge amount of information contained in medical databases. 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 heart disease diagnosis 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 detection, and can be utilized for handling high dimensional data. The obtained selected features are fed to DBN, which is trained by the 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 heart diseases, such ascongenital heart disease, coronary artery disease, and arrhythmia.

Author Contributions: Conceptualization, A.M.A. methodology, A.M.A.; software, A.M.A.; validation, A.M.A.; resources, A.M.A.; data curation, A.M.A.; writing—original draft preparation, A.M.A. and W.M.N.W.Z.;

writing—review and editing, A.M.A. and W.M.N.W.Z.; visualization, A.M.A. and W.M.N.W.Z.; supervision, W.M.N.W.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Abdel-Basset, M.; Gamal, A.; Manogaran, G.; Long, H.V. A novel group decision making model based on neutrosophic sets for heart disease diagnosis. *Multimed. Tools Appl.* **2019**, *79*, 9977–10002. [CrossRef]
- 2. Acharjya, D.P. A Hybrid Scheme for Heart Disease Diagnosis Using Rough Set and Cuckoo Search Technique. *J. Med. Syst.* **2020**, *44*, 27.
- Ahn, G.J.; Hu, H.; Lee, J.; Meng, Y. Representing and Reasoning about Web Access Control Policies. In Proceedings of the IEEE 34th Annual Computer Software and Applications Conference, Seoul, Korea, 19–23 July 2010; pp. 137–146.
- Alizadehsani, R.; Habibi, J.; Hosseini, M.J.; Mashayekhi, H.; Boghrati, R.; Ghandeharioun, A.; Bahadorian, B.; Sani, Z.A. A data mining approach for diagnosis of coronary artery disease. *Comput. Methods Programs Biomed.* 2013, 11, 52–61. [CrossRef] [PubMed]
- 5. Alzahani, S.M.; Althopity, A.; Alghamdi, A.; Alshehri, B.; Aljuaid, S. An overview of data mining techniques applied for heart disease diagnosis and prediction. *Lect. Notes Inf. Theory* **2014**, *2*, 310–315. [CrossRef]
- Babič, F.; Olejár, J.; Vantová, Z.; Paralič, J. Predictive and descriptive analysis for heart disease diagnosis. In Proceedings of the Federated Conference on Computer Science and Information Systems (FedCSIS), Prague, Czech Republic, 3–6 September 2017; pp. 155–163.
- 7. Chang, X.; Wang, Q.; Liu, Y.; Wang, Y. Sparse Regularization in Fuzzy c-Means for High-Dimensional Data Clustering. *IEEE Trans. Cybern.* **2017**, *47*, 2616–2627. [CrossRef]
- 8. Fatima, M.; Pasha, M. Survey of Machine Learning Algorithms for Disease Diagnostic. *J. Intell. Learn. Syst. Appl.* **2017**, *9*, 1–16. [CrossRef]
- Ghumbre, S.; Patil, C.; Ghatol, A. Heart disease diagnosis using support vector machine. In Proceedings of the International Conference on Computer Science and Information Technology (ICCSIT), Mumbai, India, 10–12 June 2011.
- Ghumbre, S.U.; Ghatol, A.A. Heart Disease Diagnosis Using Machine Learning Algorithm. In Proceedings of the International Conference on Information Systems Design and Intelligent Applications, Visakhapatnam, India, 5–7 January 2012; Volume 132, pp. 217–225.
- 11. Giri, D.; Acharya, U.R. Automated diagnosis of Coronary Artery Disease affected patients usingLDA, PCA, ICA and Discrete WaveletTransform. *Knowl. Based Syst.* **2013**, *37*, 274–282. [CrossRef]
- 12. Heart Disease Data Set. Available online: http://archive.ics.uci.edu/ml/datasets/Heart+Disease (accessed on 22 April 2020).
- 13. Jabbar, M.A.; Deekshatulu, B.; LandChandra, P. Classification of Heart Disease Using K- Nearest Neighbor and Genetic Algorithm. *Procedia Technol.* **2013**, *10*, 85–94. [CrossRef]
- 14. Jabbar, M.A.; Deekshatulu, B.L.; Chandra, P. Heart disease classification using nearest neighbor classifier with feature subset selection. *An. Ser. Inform.* **2013**, *11*, 47–54.
- 15. Kukar, M.; Kononenko, I.; Groselj, C.; Kralj, K.; Fettich, J. Analysing and Improving the Diagnosis of Ischaemic Heart Disease with Machine Learning. *Artif. Intell. Med.* **1999**, *16*, 25–50. [CrossRef]
- 16. Magesh, G.; Swarnalatha, P. Optimal feature selection through a cluster-based DT learning (CDTL) in heart disease prediction. *Evol. Intell.* **2020**, 1–11. [CrossRef]
- 17. Mangai, S.A.; Sankar, B.R.; Alagarsamy, K. Taylor Series Prediction of Time Series Data with Error Propagated by Artificial Neural Network. *Int. J. Comput. Appl.* **2014**, *89*, 41–47.
- 18. Mannepalli, K.; Sastry, P.N.; Suman, M. A novel Adaptive Fractional Deep Belief Networks for speaker emotion recognition. *Alex. Eng. J.* **2017**, *56*, 485–497. [CrossRef]
- 19. Medhekar, D.S.; Bote, M.P.; Deshmukh, S.D. Heart disease prediction system using naive Bayes. *Int. J. Enhanc. Res. Sci. Technol. Eng.* **2013**, *2*, 1–5.
- 20. Meng, X.; Gao, X.Z.; Lu, L.; Liu, Y.; Zhang, H. A new bio-inspired optimisation algorithm: Bird Swarm Algorithm. *J. Exp. Theor. Artif. Intell.* **2016**, *28*, 673–687. [CrossRef]

- 21. Mohan, S.; Thirumalai, C.; Srivastava, G. Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques. *IEEE Access* 2019, 7, 81542–81554. [CrossRef]
- 22. Nilashi, M.; Ahmadi, H.; Manaf, A.A.; Rashid, T.A.; Samad, S.; Shahmoradi, L.; Aljojo, N.; Akbari, E. Coronary Heart Disease Diagnosis Through Self-Organizing Map and Fuzzy Support Vector Machine with Incremental Updates. *Int. J. Fuzzy Syst.* **2020**, *23*, 1376–1388. [CrossRef]
- 23. Nourmohammadi-Khiarak, J.; Feizi-Derakhshi, M.R.; Behrouzi, K.; Mazaheri, S.; Zamani-Harghalani, Y.; Tayebi, R.M. New hybrid method for heart disease diagnosis utilizing optimization algorithm in feature selection. *Health Technol.* **2019**, *10*, 667–678. [CrossRef]
- 24. Oyyathevan, S.; Askarunisa, A. An expert system for heart disease prediction using data mining technique: Neural network. *Int. J. Eng. Res. Sports Sci.* **2014**, *1*, 1–6.
- Palaniappan, S.; Awang, R. Intelligent heart disease prediction system using data mining techniques. In Proceedings of the International Conference on Computer Systems and Applications, Doha, Qatar, 31 March–4 April 2008.
- 26. Palaniappan, S.; Awang, R. Intelligent Heart Disease Prediction System Using Data Mining Techniques. *Int. J. Comput. Sci. Netw. Secur.* 2008, *8*, 343–350.
- 27. Patil, S.B.; Kumaraswamy, Y.S. Extraction of significant patterns from heart disease warehouses for heart attack prediction. *Int. J. Comput. Sci. Netw. Secur.* **2009**, *9*, 228–235.
- 28. Pattekari, S.A.; Parveen, A. Prediction system for Heart Disease using Naive Bayes. *Int. J. Adv. Comput. Math. Sci.* 2012, *3*, 290–294.
- 29. Ranganatha, S.; Raj, H.P.; Anusha, C.; Vinay, S.K. Medical data mining and analysis for heart disease dataset using classification techniques. In Proceedings of the National Conference on Challenges in Research & Technology in the Coming Decades (CRT), Ujire, India, 27–28 September 2013.
- 30. Reddy, G.T.; Reddy, M.P.K.; Lakshmanna, K.; Rajput, D.S.; Kaluri, R.; Srivastava, G. Hybrid genetic algorithm and a fuzzy logic classifier for heart disease diagnosis. *Evol. Intell.* **2019**, *13*, 185–196. [CrossRef]
- 31. Safdar, S.; Zafar, S.; Zafar, N.; Khan, N. Machine learning based decision support systems (DSS) for heart disease diagnosis: A review. *Artif. Intell. Rev.* **2018**, *50*, 597–623. [CrossRef]
- Shah, S.M.S.; Shah, F.A.; Hussain, S.A.; Batool, S. Support Vector Machines-based Heart Disease Diagnosis using Feature Subset, Wrapping Selection and Extraction Methods. *Comput. Electr. Eng.* 2020, *84*, 106628. [CrossRef]
- 33. Shouman, M.; Turner, T.; Stocker, R. Using data mining techniques in heart disease diagnosis and treatment. In Proceedings of the IEEE Japan-Egypt Conference on Electronics, Communications and Computers, Alexandria, Egypt, 6–9 March 2012; pp. 173–177.
- 34. Subbalakshmi, G. Decision support in heart disease prediction system using naive bayes. *Indian J. Comput. Sci. Eng.* **2011**, *2*, 170–174.
- 35. Thiyagaraj, M.; 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: Singapore, 2020; Volume 2, pp. 141–152.
- 36. Yeh, D.Y.; Cheng, C.H.; Chen, Y.W. A predictive model for cerebrovascular disease using data mining. *Expert Syst. Appl.* **2011**, *38*, 8970–8977. [CrossRef]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).