IJCTA, 8(5), 2015, pp. 2137-2143

An Intelligent Decision Support System for Cardiac Disease Detection

Lokanath Sarangi[1], Mihir Narayan Mohanty[2]*, Srikanta Pattnaik[3]

Abstract: Quality service in time with less cost is a great challenge in almost all fields. In the field of medicine, research is carried out with the help of engineers. Disease diagnosis is very difficult task in the field of medicine. Many complex case patients appear in the hospital for survival. Heart disease is one of such task. It is predicted by examining the report of ECG, MRI, Blood pressure, diabetes in traditional manner. As it is a multilayer problem, authors have taken an attempt to diagnose through an intelligent decision system. It consists of two phases. In initial stage, the symptoms are considered to select the tests required. Some of the cases, it may not require for any test. For this purpose, the multi-layer perceptron network has been utilized. It has been trained with BP algorithm. Once the tests are chosen, further processing is done using the RBFN to detect the disease related to cardiac problem. RBFN with Gaussian Kernel function detects the problem accurately, that helps the physicians for further care.

Index Terms: Cardiac Disease, Detection, ECG, MLPNN, RBFN, Decision System.

1. INTRODUCTION

Day by day smart service is a greater demand everywhere. Similar cases in the field of medicine create challenge for physicians, laboratories of health care units. In recent years varieties of diseases are to be recognized which is a complex job. Heart disease is one of them. For leaving being, it is an important organ that controls the blood flow within the body. Any problem within the heart may cause the question of survival. In addition, it affects the other parts of the body. The cardiac care, diagnosis of the disease related to heart is an essential component.

For many purposes, the medical industries take the support of engineers and scientists for modern equipment. Researchers in both the fields have tried to develop the smart equipment since last two decades. In this work, the objective is to develop an intelligent decision support system for detection of cardiac problem. It can extract knowledge from the historical database. The system can help the medical practitioners for easy and faster diagnosis.

As the smart service is required with less cost, the computer-based systems may help to achieve the purpose. These types of system will help in case of data acquisition, preprocessing of data, patient information storage and also for monitoring purpose. Currently sufficient data are available to practitioners in recent years. It includes the clinical symptoms; variety biochemical information and imaging information are useful during diagnosis and treatment. For all these purposes, neural network based systems found to be effective and is used in this work to develop a decision support system.

The remaining part of the paper is organized as follows

Section-II describes the related work by researchers, Section-III explains the proposed method and the result is depicted in Section-IV, Section-V provides the concluding remark.

* Department of Electronics and Communication Engineering, ITER, Siksha 'O' Anusandhan University, Bhubaneswar, Odisha, India, *Email: lokanathsarangi@yahoo.com, mihirmohanty@soauniversity.ac.in, srikantapatnaik@soauniversity.ac.in*

2. RELATED LITERATURE

Many researchers have proposed many methods for detection of heart diseases including neural network method. Some of them are cited in this section.

In [1], the model has been composed of cascaded Multi-Layer Perceptron (MLP).The network is trained by LM algorithm Carlos Ordonez tried with an algorithm to reduce the number of rules to predict the degree of narrowing in arteries used for search limitation [2]. K.C Tan *et.al* proposed a hybrid GA-SVM model which acted as a good classifier with accuracy 84.07% [3]. Jesmin Nahar *et.al* got the result on heart disease data using various rules of data mining algorithms and found efficient result [4]. A proposal for feature selection method called Kernel F-score Feature Selection (KFFS) used as preprocessing step in the classification of medical data [5]. In [6], classification method based on preprocessing the data with Principal Component Analysis (PCA) and differential evolution classifier used to diagnose heart disease. Resul Das *et.al* proposed several tools and various methodologies for effective diagnosis using Neural Network (NN) and Statistical Analysis System (SAS) [7]. GA based system selects critical clinical features required heart disease diagnosis [8]. Various classifiers and knowledge extraction techniques were improved the accuracy in this direction [9-13]. The usefulness of MLP neural network architecture for heart disease detection brought a revolutionary change. Heart disease database consisting of 352 cases were assessed with help of three methods cross validation, hold out and boot strapping with a view to supporting Medical Decision Support System (MDDS) using MLP employing back propagation algorithm. The accuracy was less than 90% and less time consuming to diagnose [14]. Artificial Neural Network and data mining technique using back propagation algorithm a system was developed to predict heart disease with more accuracy. The MLPNN model had better result providing the patient with early diagnosis as in [15]. Multilayer Feed Forward Network was analyzed with back propagation algorithm gives high accuracy with momentum and variable learning rate to train the networks. MLP NN requires compact architecture in comparison to other NNs.

3. PROPOSED METHOD

Neural Networks are suitable for solution of engineering as well as biomedical problems widely. In this work, we have used the MLP network and RBF network for decision support system for heart disease detection and diagnosis. MLP network for improvement of learning algorithm indicates a faster and higher capability of detection. RBF network is similar to MLP. But in this case the features as pathological tests represented in ‘N’ dimensional space. The hidden layer has a number of radial units to model any function. It has a single hidden layer and a model any nonlinear function with a single hidden layer using Gaussian activation function. These networks are computationally efficient.

The block diagram of the proposed system is shown in Figure 1. It consists of two neural network models back to back. In the first stage, the MLP network decides the type of tests required or not from the patients’ symptoms. There are 13 symptoms consider in this work. It has used the back propagation-learning algorithm with the different symptoms from database.

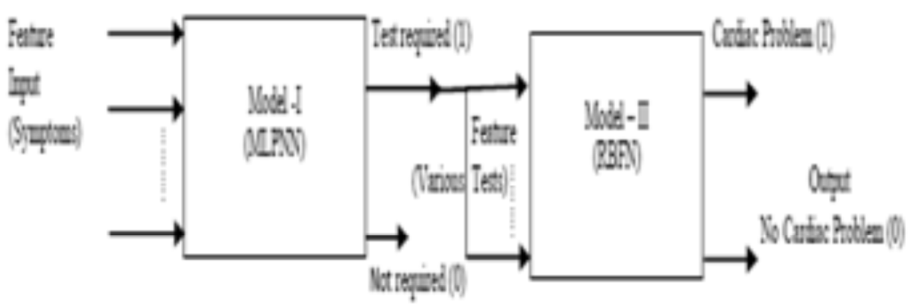

Figure 1: Block diagram of proposed system

The MLP network architecture is depicted in Figure 2. Once the test is defined by model-I can be performed and the different attributes from the test are fed to the RBF network for detection of heart disease. The attributes are eight test results. The RBF network used is shown in Figure 3.

3.1. Multilayer Perception

MLP is one of the most frequently used neural network architectures. The MLP architecture and the BP algorithm for learning MLP are used for test recommendation from symptoms.

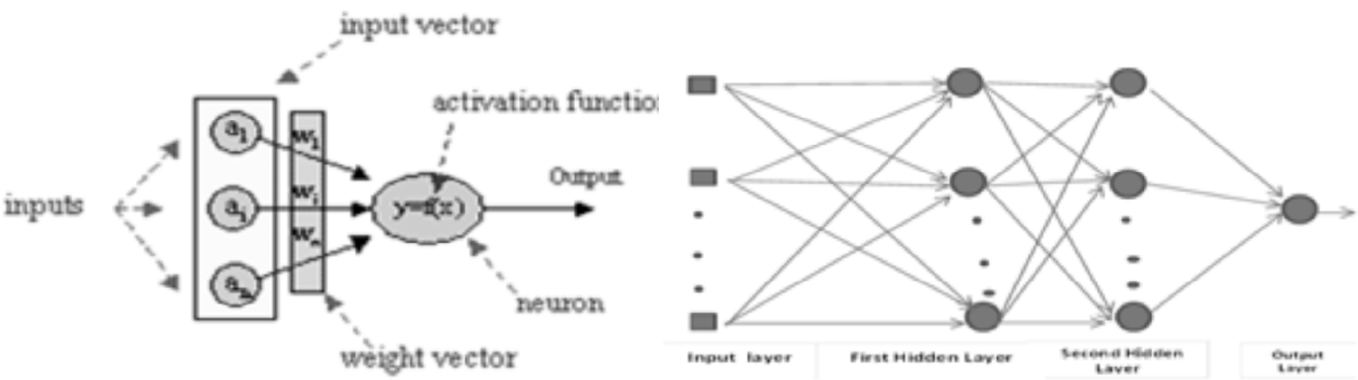

Figure 2: Representation of MLP network

The inputs received (as shown in Figure 2) can be represented as an input vector and can be defined as, $A = (a_1, a_2 \dots a_n)$, where a_i is the signal from the ith input. The weights connected to the jth neuron can be represented as a weight vector, $Wj = (w_{1j}, w_{2j} \dots w_{nj})$, where w_{ij} represents the weight associated to the connection between the processing element a_i, and the processing element a_j. While the threshold is determined by the weights associated with the inputs (Eq. 1), it modulates the response of a neuron to pre-defined range of values. Equation 2 defines the output y of a neuron as an activation function f of the weighted sum of *n+ 1 input*. The threshold is incorporated into the equation as the input

$$SUM = \sum_{i=1}^{n} x_i w_i \tag{1}$$

$$y = f\left(\sum_{i=0}^{n} x_i w_i\right) \tag{2}$$

$$f(x) = \begin{cases} 1 & if \quad \sum_{i=1}^{n} x_i w_i > 0 \\ 0 & if \quad \sum_{i=1}^{n} x_i w_i \le 0 \end{cases} \tag{3}$$

3.2. Radial Basis Function Network

RBFN uses input layer, hidden layer and output layers similar to MLP. In this, features are represented in *N* dimensional space using center point and radius [16]. The hidden layer has a number of radial units to model any shape or function. The number of hidden layer need not be more than one as sufficient number of radial units can form a basis for pattern recognition. RBFN as against MLP can model any nonlinear function with a single hidden layer using Gaussian activation function. Therefore, they can be computationally simple and more efficient in a given scenario. The RBF network depicted in the second stage is shown in Figure 3.

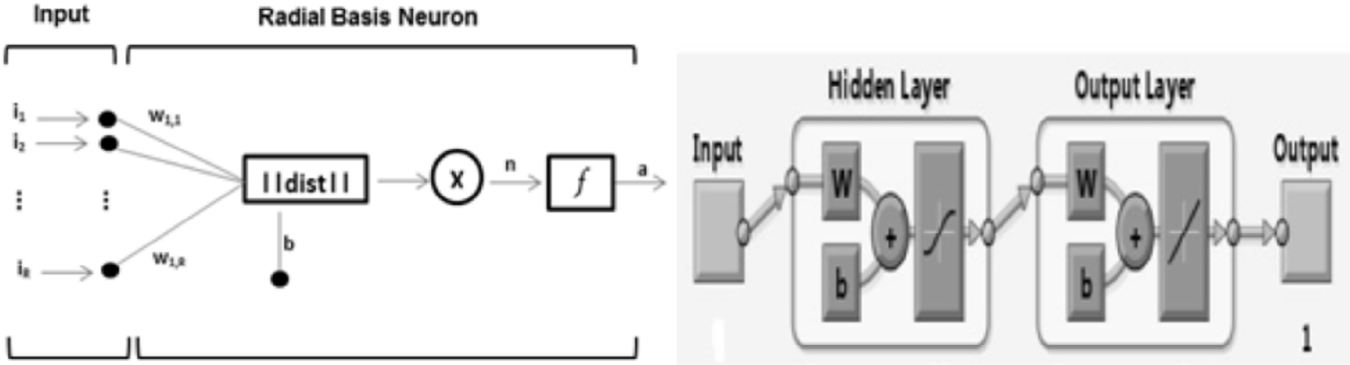

Figure 3: Radial Basis Function Network

The Gaussian activation function used in RBFN is given by

$$G_j(X) = \exp\left[-\left(X-\mu_j\right)^T \Sigma_j^{-1}\left(X-\mu_j\right)\right] \tag{4}$$

3.3. Training Algorithms

The neural network models the underlying function of a certain mapping through training. In supervised application, a set of data samples is provided for which the corresponding network outputs are to be trained (Haykin, 1999). In this case the network parameters are calculated in such a way so that they minimize a cost function and is represented as,

$$\min \Sigma_{i=1}^{Q}\left(Y_k\left(X_i\right)-F_k\left(X_i\right)\right)^T\left(Y_k\left(X_i\right)-F_k\right) \tag{5}$$

where Q is the total number of vectors from the training set, denotes the RBF output vector and represents the output vector associated with the a data sample from the training set. A large variety of training algorithms has been tested for training RBF networks. In initial stage, each data sample was assigned a basis function. In RBF, centres are uniformly distributed in the data space. The function to be modelled is obtained by interpolation. In less basis functions then given data samples are used. A least squares solution that minimizes the interpolation error is proposed.

$$\left\|X_i-\hat{\mu}_j\right\| = \min_{k=1}^{L}\left\|X_i-\hat{\mu}_j\right\| \tag{6}$$

A set of generator functions is proposed for hidden unit activation function selection.

Finally, the combined network model for detection of disease from symptoms is shown in Figure 4. It is the self-organizing network. As the new neurons are added to the hidden layer, the learning algorithm

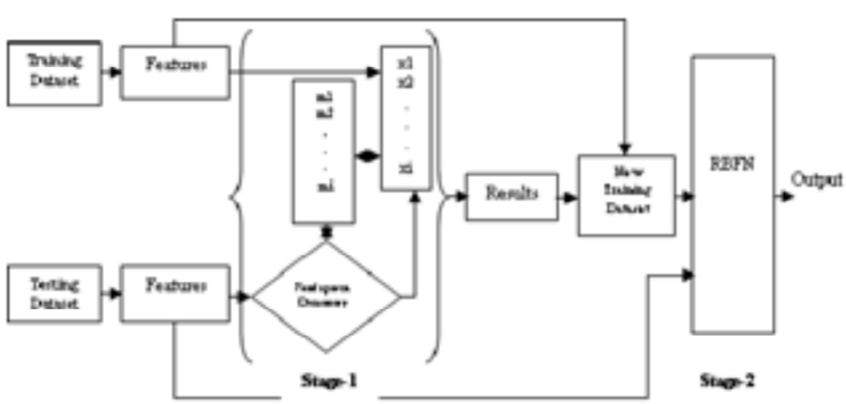

Figure 4: Proposed network for disease detection

attempts to maximize the magnitude of correlation among new neurons output and error of the network that is to be minimized. The numbers of input and output units depend on the application. In the cascaded learning process works as follows:

i) Initialize the input and output units based on the problem

ii) Set the hidden node initial value

iii) Train the network

iv) Continue this till the present number is reached

v) Test the accuracy with corresponding test dataset

vi) Repeat the procedure until the error reduction.

4. EXPERIMENTAL RESULTS

Dataset used for experiment is comprised the information of Heart Disease. The data collected from local hospitals and validated with data of UCI Repository. The data folder actually consisted of 8 different datasets with the variation in number of attributes. It is seemed to be more suitable with less number of 8 attributes (out of 64) with 300 instances. The description of the data set is illustrated in Table 1.

Table 1
Different datasets with the variation in number of attributes

S.N	*Attribute*	*Description*	*Range*
1	*Age*	Patient's Age	25-80
2	*Sex*	1= male; 0 = female	0-1
3	*CP (Chest Pain)*	Value 1: typical angina Value 2: non-anginal pain	1-2
4	*BP*	Blood pressure(in mm Hg)	90-190
5	*Cholstrl*	Serum cholesterol in mg/dl	120-560
6	*FBS*	(Fasting blood sugar 110mg/dl) (1 = true; 0 = false)	0-1
7	*ECG*	Electrocardiography results Value 0: normal Value 1: having ST-T wave abnormality	0-1
8	*Peak*	ST depression induced by exercise relative to rest	0-6.5

The method primarily based on the information collected from precedent experiences and from current circumstances, which visualizes something as it may occur in future is known as prediction. In this work, a total of 300 data samples are used for simulating the network architecture. 210 samples are used for training and 60 samples each are used for validation and rest are used for testing the system. Feature values of seven attributes of all samples have been used as input neurons and feature value of one attribute is used as output neuron in the first state of the classifier. In the second stage of the classifier the corresponding values are five and one respectively. Maximum numbers of epochs are set at 300 and the network took 50 epochs to attain the required residual error value. Initially, the signals are obtained from various attributes of the patients. These features are fed as input to the classifier for modelling the respective attributes to classify the disease accurately.

Figure 5 shows the convergence with 220 samples. Figure 6 shows the performance of RBF networks with 50 epochs.

5. CONCLUSION

In this paper, we have presented a decision support system based on the combination of MLPNN and RBFN architecture for heart disease diagnosis. The approach based on MLP Back Propagation neural

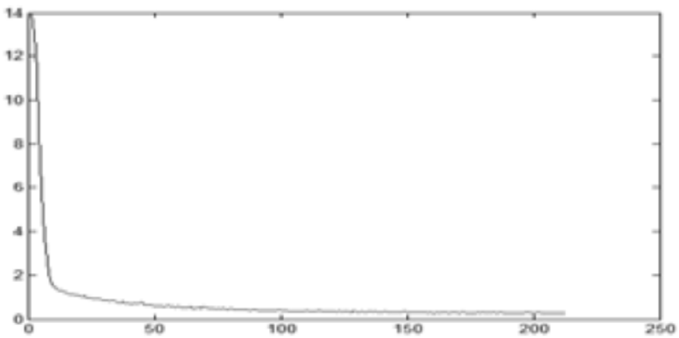

Figure 5: Convergence curve with Training Samples

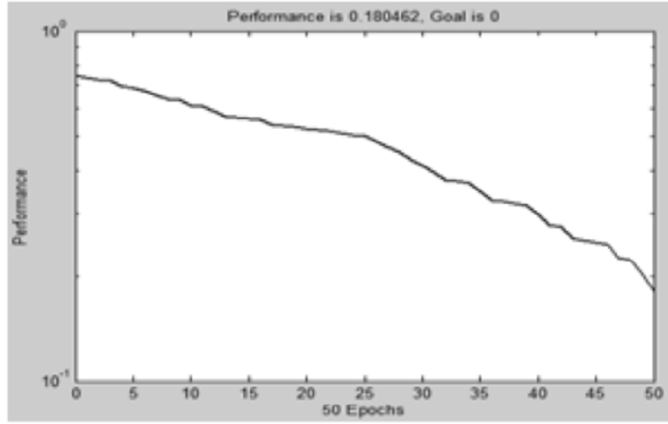

Figure 6: Performance of RBFN with 50 epochs

network is proposed for requirement of the test for Heart Disease. The Proposed system used nearly 8 significant medical attributes for Heart Disease detections with 13 attributes along with the symptoms. The experiment performs well with the proposed algorithm and is validated with standard UCI repository dataset. The training of Multilayer Perceptron Algorithm with BP algorithm found the best training function. The experimental results show the encouraging results using proposed mode with improved accuracy.

References

[1] L. Sarangi, M.N. Mohanty and S. Pattanayak, "Design of MLP Based Model for Analysis of Patient Suffering from Influenza" International conference on Intelligent Computing, Communication & Convergence pp. 39, Jan 2016.

[2] Carlos Ordonez," Association Rule Discover with the Train and Test Approach for the Heart Disease Prediction", IEEE Transactions on Information Technology in Biomedicine, vol. 10, no. 2, pp. 334-343, April 2006.

[3] K.C. Tan and E.J. Teoh *et al*, "A hybrid evolutionary algorithm for attribute selection in data mining", Journal of Expert system with applications, vol. 36, pp. 8616-8630, 2009.

[4] Jesmin Nahar and Tasadduq Imam *et al*," Association rule mining to detect factors which contribute to heart disease in males and females", Journal of Expert Systems with Applications, vol. 40, pp. 1086–1093, 2013.

[5] Kemal Polat and Salih Gunes," A new feature selection method on classification of medical datasets: Kernel F-score feature selection", Journal of Expert Systems with Applications, vol. 36, pp. 10367–10373, 2009.

[6] Pasi Luukka and Jouni Lampinen, "A Classification Method Based on Principal Component Analysis and Differential Evolution Algorithm Applied for Prediction Diagnosis from Clinical EMR Heart Data Sets", Journal of Computer Intelligence in Optimization Adaption, Learning and Optimization, vol. 7, pp. 263-283, 2010.

[7] Resul Das and Ibrahim Turkoglu, *et al*, "Effective diagnosis of heart disease through neural networks ensembles", Journal of expert system with applications, vol. 36, pp. 7675–7680, 2009.

[8] Hongmei Yan and Jun Zheng, *et al*, "Selecting critical clinical features for heart diseases diagnosis with a real-coded genetic algorithm", Journal of Applied Soft Computing, vol. 8, pp. 1105-1111, 2008.

[9] Akin Ozcift and Arif Gulten, "Classifier ensemble construction with rotation forest to improve medical diagnosis performance of machine learning algorithms", Journal of Computer Methods and Programs in Biomedicine, vol. 104, pp. 443-451, 2011.

[10] Chih-Lin Chi and W. Nick Street, *et al*, "A decision support system for cost-effective diagnosis", Journal of Artificial Intelligence in Medicine, vol. 50, pp. 149-161, 2010.

[11] Yoon-Joo Park and Se-Hak Chun, *et al*, "Cost-sensitive case-based reasoning using a genetic algorithm: Application to medical diagnosis", Journal of Artificial Intelligence in Medicine, vol. 51, pp. 133-145, 2011.

[12] Debabrata Pal and K. M. Mandana, *et al*, "Fuzzy expert system approach for coronary artery disease screening using clinical parameters", journal of knowledge based system, vol. 36, pp. 162-174, 2012.

[13] Ismail Babaoglu and Og¢uz Findik *et al*, "A comparison of feature selection models utilizing binary particle swarm optimization and genetic algorithm in determining coronary artery disease using support vector machine", Journal of Expert System With Applications, vol. 37, pp. 3177-3183, 2010.

[14] Hongmei yan, Yingtao Jiang, Jun Zheng, Chenglin Peng, Qinghui Li, "A Multilayer Perceptron-Based Descision Support System for heart Disease Diagnosis", vol. 30, pp. 272-281, 2006.

[15] Miss. Chaitrali S. Dangarw, Dr. Mrs. Sulabha S. Apte, "A Data Mining Approach For Prediction of Heart Disease Using Neural Networks" vol. 3, Issue 3, October-december, pp. 30-40, 2012.

[16] S Haykin, Neural Network, 2nd ed., 1999.