Journal of Theoretical and Applied Information Technology
10th October 2013. Vol. 56 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

BACKPROPOGATION NEURAL NETWORK FOR PREDICTION OF HEART DISEASE

NABEEL AL-MILLI
Financial and Business Administration and Computer Science Department
Zarqa University College
Al-Balqa' Applied University
E-mail: Nabeel.almilli@bau.edu.jo

ABSTRACT

Recently, several software's, tools and various algorithms have been proposed by the researchers for developing effective medical decision support systems. Moreover, new algorithms and new tools are continued to develop and represent day by day. Diagnosing of heart disease is one of the important issue and many researchers investigated to develop intelligent medical decision support systems to improve the ability of the physicians. Neural network is widely used tool for predicting heart disease diagnosis. In this research paper, a heart disease prediction system is developed using neural network. The proposed system used 13 medical attributes for heart disease predictions. The experiments conducted in this work have shown the good performance of the proposed algorithm compared to similar approaches of the state of the art.

Keywords: *Backpropagation, Heart Disease, Neural Network.*

1. INTRODUCTION

Heart is the significant part of our body. Life is itself dependent on efficient working of heart. If operation of heart is not good, it will influence the other body parts of human such as brain, kidney … etc. Heart disease is a disease that is based on the performance of heart. Several factors which increases risk of Heart disease such as: cholesterol, high blood pressure, lack of physical exercise, smoking and obesity.

The World Health Organization (WHO) has estimated that 12 million deaths occur worldwide, where heart disease is the major cause of deaths. For example, in 2008, 17.3 million people died due to Heart Disease. WHO estimated by 2030, almost 23.6 million people will die due to Heart disease.

In order to reduce the risk of heart disease, prediction should be done. Discovering of heart disease is usually based on symptoms, physical examinations and sings of patient body. Normally, doctors are predicting heart disease by knowledge and experience. Discovering and predicting diseases is a difficult task in medical environment. Discovering heart disease from several factors is a multilayered problem which may lead to negative presumptions and unpredictable effects. As a result, Healthcare industry today creates large amounts of complex data about patients, hospitals resources, disease diagnosis, electronic patient records, medical devices etc. The huge amount of data (records) is a key resource to be processed and analyzed for knowledge extraction that enables support for cost-savings and decision making. Figure 1, illustrates the difficulties that will be arrive during diagnosis which leads to negative presumptions and unpredictable effects.

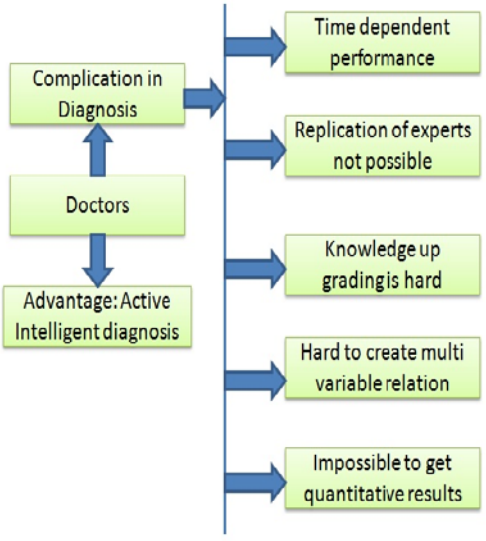

Figure 1: Complexity In Diagnosis With Doctor

The remainder of this paper is structured as follow. In section 2, we describe literature review of heart diseases. In Section 3, we illustrate

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Backpropagation Neural Network. The results and findings are presented in Section 5. Finally, conclusion remarks are made in Section 6.

2. LITERATURE REVIEW

Dilip Roy Chowdhury et al. [1] applied a backpropagation neural network in predicting neonatal disease diagnosis. The proposed method is used to recognize a pattern for the diagnosing and prediction of neonatal diseases. The authors applied backpropogation algorithm to train a neural network on different categories of neonatal diseases. The accuracy of the proposed model is 75% with higher stability.

Milan Kumari et al. [2] solve cardiovascular disease dataset using different data mining algorithms, such as: Support Vector Machine, Artificial neural networks (ANNs), Decision Tree, and RIPPER classifier. The author's analyze the performance of these algorithms through several statistical analysis factors such as: sensitivity, specificity, accuracy, error rate, True Positive Rate and False Positive Rate. Accuracy of RIPPER, Decision Tree, ANN and SVM are 81.08%, 79.05%, 80.06% and 84.12% respectively. While the results of error rates for RIPPER, Decision Tree, ANN and SVM are 2.756, 0.2755, 0.2248 and 0.1588 respectively. The analysis shows that out of these four classification models SVM predicts cardiovascular disease with least error rate and highest accuracy.

A decision support system for diagnosis of Congenital Heart Disease has been proposed by Vanisree K et al. [3]. The core of the proposed system is based on Backpropagation Neural Network (multi layered Feed Forward Neural Network). The benchmark set used in this work are the signs, symptoms and the results of physical evaluation of a patient. The proposed system achieved an accuracy of 90%.

Niti Guru et al. [4] applied a neural network for prediction of heart disease, blood pressure and sugar. The benchmark consists of 78 records with 13 attributes are used for training and testing. The author used a supervised network for diagnosis of heart disease and trained it using back propagation algorithm. On the basis of unknown data is entered by doctor the system will find that unknown data from training data and generate list of possible disease from which patient can suffer.

A prototype Intelligent Heart Disease Prediction System (IHDPS) based on data mining techniques is proposed by Sellappan Palaniappan et al. [5]. The techniques used are Decision Trees, Naˆive Bayes and Neural Network. The proposed models are developed based on .NET platform. The benchmark dataset has several attributes such as age, sex, blood pressure and blood sugar which is used to predict the likelihood of patients getting a heart disease.

Shantakumar B.Patil et al [6] applied a proficient methodology for the extraction of significant patterns from the heart disease warehouses for heart attack prediction. The first step of this work, the data warehouse is pre-processed in order to make it suitable for the mining process. K-mean clustering algorithm has been applied for clustering the heart disease warehouse. Consequently the frequent patterns applicable to heart disease are mined with the aid of the MAFIA algorithm from the data extracted. In addition, the patterns vital to heart attack prediction are selected on basis of the computed significant weight age. The neural network is trained with the selected significant patterns for the effective prediction of heart attack.

3. BACKPROPAGATION NEURAL NETWORKS

The back propagation algorithm is a technique used in developing multilayer neural networks in a supervised manner. The back propagation algorithm, also known as the error back propagation algorithm, is based on the error-correction learning rule [7]. The algorithm has two passes through the different layers of the network: a forward pass and a backward pass. In the forward pass, an activity pattern is applied to the input nodes of the network, and its effect propagates through the network layer by layer. Finally, a set of outputs is produced as the actual response of the network. During the forward pass the synaptic weights of the networks are all fixed. During the backward pass, the synaptic weights are all adjusted in accordance with an error-correction rule. The actual response of the network is subtracted from a desired response to produce an error signal. This error signal is then propagated backward through the network. The synaptic weights are adjusted to make the actual response of the network move closer to the desired response in a statistical sense. The weight adjustment is made according to the generalized delta rule [8] to minimize the error. An example of a multilayer perceptron with two hidden layers is shown in Figure 2.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

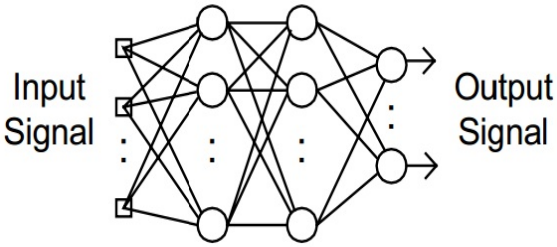

Figure 2:Multilayer Perceptron with Two Hidden Layers

Two commonly used neuron activation functions for the neuron in Figure 2 are sigmoidal and tansig functions. Both functions are continuously differentiable everywhere and typically has the following mathematical form:

$$\text{Sigmoidal f(x)} = \frac{1}{1+\exp(ax)}, a \succ 0$$

$$\text{tansig f(x)} = a\tanh(bx), a \,\&\, b \succ 0$$

4. PROBLEM DESCRIPTION

In this work the medical data related to Heart diseases is considered. This benchmark dataset was obtained from Cleveland database [9]. This is publicly available dataset in the Internet. Cleveland dataset concerns classification of person into normal and abnormal person regarding heart diseases. The dataset Benchmark datasets total 166 records. The dataset is divided into 2 sets training (116 records) and testing set (50 records). Matlab 2010 used for experiment. The data consists of 13 attributes (inputs) and 4 classes (outputs). Tables 1 and 2 illustrate the representation of the attributes and classes respectively.

5. SIMULATION RESULTS

The proposed algorithm was programmed using Matlab and simulations were performed on the Intel Pentium 4 2.33 GHz computer. Parameters settings used for the proposed algorithm after some preliminary experiments is shown in Table 3. We ran the experiments for 10000 iterations.

Table 1: Benchmark datasets attributes

Attribute	Description	Range
Age	Age in years	Continuous
Sex	(1=male; 0=female)	0,1
Cp	Value 1:typical angina Value 2: atypical anginal Value 3: non-anginal pain Value 4: asymptotic	1,2,3,4
trestbps	Resting blood pressure(in mm Hg)	Continuous
chol	Serum cholesterol in mg/dl	Continuous
fbs	(Fasting blood sugar .120mg/dl) (1=true; 0=false)	0,1
restecg	electrocardiography results Value 0: normal Value 1:having ST-T wave abnormality (T wave inversions and/or ST Elevation or depression of>0.05mV) Value 2:showing probable or definite left	0,1,2
Thalach	Maximum heart rate achieved	Continuous
Exang	Exercise induced angina(1=yes;0=no)	0,1
OldPeak	ST depression induced by exercise relative to rest	Continuous
Slope	The slope of the peak exercise ST segment Value 1: up sloping Value 2: flat Value 3:down sloping	0, 1, 2
Ca	Number of major vessels (0-3) Colored by fluoroscopy	Continuous
Thal	Normal, fixed defect, reversible defect	3,6,7

Table 2: Benchmark datasets classes

Class	Description
Class 0	Normal Person.
Class 1	First Stroke
Class 2	Second Stroke
Class 3	End of Life

Table 3: Parameters Setting

Parameter	Value
Iterations	1000
Number of Neurons in Input layer	13
Number of Neurons in Hidden layer	8,5,2
Number of Neurons in output layer	1

Table 4 presents the results obtained from our model for 11 runs with different seed number. Run10 has the highest variance from training and testing process. Figures 3 and 5 show the original and predicted output for training and testing respectively. Figure 5 shows the box plots that illustrate the distribution of solution quality for training and testing datasets. In both cases, there is less dispersion of the output data. We can see that

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

there are a close gap between the best, average and worse solution qualities which demonstrates that it is robust algorithm.

Table 4: Results Of Eleven Runs

Run	Vaf. Testing	Vaf Training
1	92%	81%
2	88%	73%
3	72%	61%
4	90%	80%
5	66%	51%
6	78%	59%
7	83%	68%
8	89%	77%
9	90%	83%
10	92%	86%
11	88%	80%

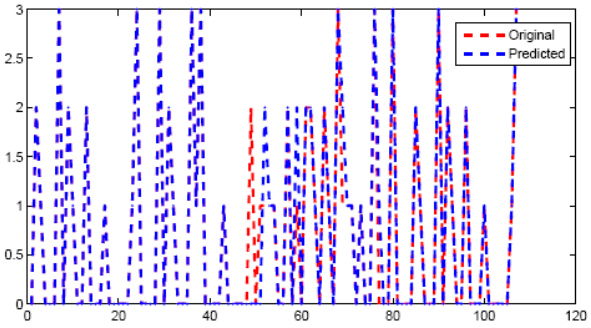

Figure 3. Original And Predicted Output For Training Dataset

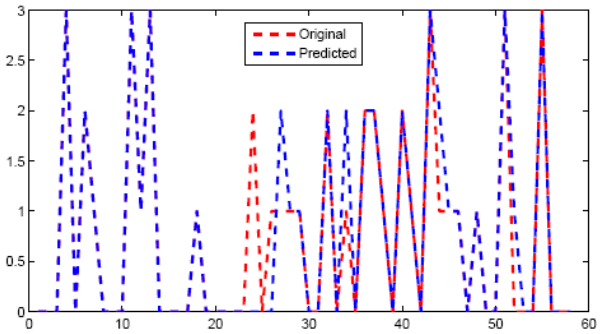

Figure 4: Original And Predicted Output For Testing Dataset

6. CONCLUSION AND FUTURE WORK

In this work, we present an approach that based on back propagation neural network to model heart disease diagnosis. In this research paper, a heart disease prediction system is developed using neural network. The proposed system used 13 medical attributes for heart disease predictions. The experiments conducted in this work have shown the good performance of the proposed algorithm compared to similar approaches of the state of the art.

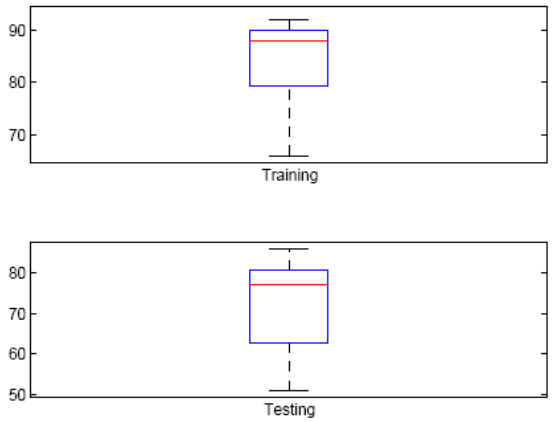

Figure. 5: Plotbox Diagram For Training And Testing

REFRENCES:

[1] Dilip Roy Chowdhury, Mridula Chatterjee R. K. Samanta, An Artificial Neural Network Model for Neonatal Disease Diagnosis, International Journal of Artificial Intelligence and Expert Systems (IJAE), Volume (2): Issue (3), 2011.

[2] Milan Kumari, Sunila Godara, Comparative Study of Data Mining Classification Methods in Cardiovascular Disease Prediction, IJCST Vol. (2), Issue (2), June 2011.

[3] Vanisree K, Jyothi Singaraju, Decision Support System for Congenital Heart Disease Diagnosis based on Signs and Symptoms using Neural Networks, International Journal of Computer Applications (0975 8887) Volume 19 No.6, April 2011.

[4] Niti Guru, Anil Dahiya, Navin Rajpal, Decision Support System for Heart Disease Diagnosis Using Neural Network, Delhi Business Review, Vol. 8, No. 1, January-June 2007.

[5] Sellappan Palaniappan, Rafiah Awang, Intelligent Heart Disease Prediction System Using Data Mining Technique, 978-1-4244-1968-5/08/, 2008 IEEE.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[6] Shantakumar B.Patil, Y.S.Kumaraswamy, Intelligent and Effective Heart Attack Prediction System Using Data Mining and Artificial Neural Network, European Journal of Scientific Research, ISSN 1450-216X, Vol.31 No.4 (2009), pp.642-656.
[7] J. Principe, N. Euliano, W. Lefebvre, Neural and Adaptive System Fundamentals Through Simulations, Wiley, 2000.
[8] S. Haykin, Neural Networks A Comprehensive Foundation, 2nd Edition, Prentice Hall, 2000.
[9] Cleveland database: http://archive.ics.uci.edu/ml/datasets/Heart+Disease