Medical Diagnosis using Back Propagation Algorithm in ANN

Meenakshi Verma
Department of computer Science
Maharaja Surajmal Institute
C-4, Janakpuri, New Delhi, India

***ABSTRACT*--Artificial Neural Networks are finding many uses in the Medical diagnosis applications. ANN plays a vital role in the medical field in solving various health problems like acute diseases and other mild diseases. The goal of this paper is to evaluate Artificial Neural Network in disease diagnosis. Three cases are studied. The first one is diabetes disease, data is the risk factors and their strength of association to the development of type 2 diabetes was used as relative weight of input variables. The second is the Hypertension disease, data is disease symptoms. Third is obesity disease i.e. body fat, data is disease symptoms. In all the above mentioned three diseases each patient is classified into two categories infected and non- infected. Classification and Prediction are important tool in medical diagnosis decision support. Feed Forward Back Propagation Neural Network Model is used as classifier to distinguish between infected and non-infected persons in all cases. In this study, the data were obtained from UCI machine learning repository in order to diagnosed diseases. The data is separated into inputs and targets. The targets for the neural network will be identified with 1's as infected and will be identified with 0's as non-infected. The Back-Propagation neural network model is systematically trained and with data sets.**

Keywords: **Artificial Neural Networks, Medical Diagnosis, Feed-forward back propagation network, Artificial Intelligence, and Decision Support Systems.**

I. INTRODUCTION

A. Artificial Neural Network

Artificial Neural Network provide a powerful tool to help doctors to analyse, model and make sense of complex clinical data across a broad range of medical applications. Most applications of ANN to medicine are a classification and prediction problem that is the task is on the basis of the measured features to assign the patient to one of a small set of classes. ANN is mathematical model widely used for classification and diagnosis in various thirst areas like effective decision making in medical field, signal processing and so on. A doctor uses a combination of a patient's case history and current symptoms to reach a health diagnosis when a patient is ill. In order to recognize the combination of symptoms and history that points to particular disease, the doctor's brain accesses memory of previous patients, as well as information that has been learned from books or other doctors. A neural network has the ability to mimic this type of decision-making process and use knowledge base of information and a training set of practice eases, to learn to diagnose diseases. The major problem in medical field is to diagnose disease. Human being always make mistake and because of their limitation, diagnosis would give the major issue of human expertise. A computer system can help in medical diagnosis, as it never gets tired or bored, can be updated easily in a matter of seconds and is rather cheap and can be easily distributed. Again a good percentage of visitors of a clinic are not sick or at least their problems are not serious, if an intelligent diagnosis system can refine that percent, it will set the doctors free to focus on nuclear and more serious cases. Neural Network is ideal in recognizing diseases using scans since there is no need to provide

ISSN: 2278 – 7798

a specific algorithm on how to identify the disease. Neural Network learns by example so the details of how to recognize the disease are not needed. What is needed is a group of examples that are representative of all the variations of the disease. The Back Propagation Neural Network algorithm used for training depends on a multilayer Neural Network with a very small learning rate. When the output of the network is continuous, it is performing prediction and when the output has discrete values, then it is doing classification. Diagnosis involves the process of collection, analysis, recognition and classification of data. Employing the technology especially ANN techniques in medical applications could result in reducing cost, time, medical error and need of Human expertise.

B. Brief Description of the Diseases to be diagnosed

Following is brief description for the three diseases taken for the purpose of the analysis of this paper:-

Diabetes

Diabetes represents a serious health problem in developed countries with estimated numbers reaching 366 million diabetes cases globally in 2030. It is one of metabolic diseases where a patient has high blood sugar either caused by the body failure to produce enough insulin or the cells failure to respond to the produced insulin.

Diabetes can be identified by studying on several readings taken from a patient such as fasting glucose, sodium, potassium, urea, creatinine, albumin and many others. Insulin is a natural hormone secreted in human body to convert sugar, starch and food into simpler molecules which is then utilized by the cells to generate energy required by daily life. Due to lack of insulin this conversion is affected and the sugar starts getting accumulated in blood stream and thus increases the blood glucose level and as a result the person develops diabetes mellitus.

Obesity (Body Fat)

Obesity is the condition whose principal characteristic is the excess of adipose tissue, in which the cells can increase in size as well as the number and results in the decrease in the quality of life and health of the individual who suffers from it.

Obesity is inherited and also acquired by an erroneous conduct of eating and lifestyle. The environment has great influence over habits and attitudes that are acquired through eating. A person with inherited tendencies to be overweight can change those tendencies if he lives in a disciplined eating environment, i.e. changes his habits by a choice of his will or helpful techniques such as hypnosis or auto hypnosis.

An overweight person who changes his old and deeply rooted habits and life patterns can obtain a healthier and better life in every aspect.

Hypertension

Hypertension is a disease that affects a wide range of the population, particularly the elderly after the age of 55. It is caused by high Blood Pressure. Blood Pressure is the force of blood pushing against blood vessels walls. The heart pumps blood into the arteries (blood vessels), which carry the blood throughout the body. If blood pressure is extremely high, there may be certain symptoms such as severe headache, fatigue, disorientation, vision problems, chest pain, difficulty in breathing, irregular heartbeat and blood in the urine.

Hypertension can cause stroke, heart attack, heart failure, kidney failure and vision problems. Men have a greater likelihood of developing high Blood Pressure than Women. This varies according to age and among various groups. Initial assessment of the Hypertensive patient should include a complete history and physical examination in to confirm a diagnosis of hypertension.

ISSN: 2278 – 7798

II. LITERATURE REVIEW

Jaafar and Ali [1] studied diabetes mellitus using Artificial Neural Network with the aim of determining whether someone is diabetes sufferer or not. A neural network of 8 inputs, 5 hidden layers and 1 output layer was used and result indicated high performance of patients diagnosed with diabetes.

Moein, Manadjemi and Moallen [2] analyzed the real procedure of medical diagnosis which usually is employed by physicians and converted to a machine implementable format. Then after selecting some symptoms of 3 different diseases, a data set contains the information of a few hundred cases was configured and applied to Multi layer neural network.

Gupta and Shreevastava [3] highlighted that almost all the physicians are confronted during their formation by the task of learning to diagnose.

Mythili, Kumar P. and Kumar R. [4] Proposed a method for diagnosing diabetes mellitus based on the risk factors. The diagnosis is accomplished using Back Propagation Neural Network Algorithm in NNTool Box of MATLAB.

Kadhim Qeethara [5] proposed and analysed the Artificial Neural networks in Medical Diagnosis using Back propagation Algorithm.

Sumanthi B. And SanthaKumaran A. [6] analysed the data of hypertension disease using the back propagation algorithm in ANN and resulted in finding out that whether patient is suffering from hypertension or not.

III. DESIGN OF PROPOSED METHOD

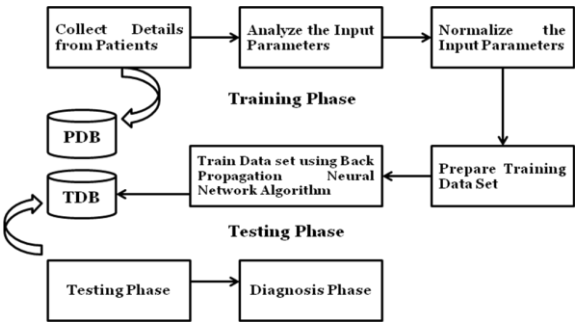

Figure 1: Framework for the Proposed Work

A. Procedure for Medical Diagnosis

(i) Data Collection and Classification

Around 122 data sets were collected in the survey for the proposed work. The inputs designed for the system are for 3 diseases, the input values range between 0 and 3 based on the condition or physical appearance of that person.

(ii) Normalizing Data

Before giving the inputs to the network input values should be normalised which is based on the activation function used.

The activation function used is log sigmoid

$$f(x)=1/1+\exp^{(-\sigma x)}$$

where σx is input of hidden layer. For the proposed study, the input values have been normalised between the range of 0 & 1.

B. Back Propagation Neural Network Algorithm

Back Propagation, an abbreviation for "backward propagation of errors", is a common method of training artificial neural networks. From a desired output, the network leans from many inputs, similar to the way a child learns to identify a dog from examples of dogs. Back propagation is a neural network learning algorithm. The feedback forward network structure is most important ANN structure. Design of a feed forward net for any specific application involves many issues, most of which require problem dependent solutions. The overall computational approach comprising of two parts: feed forward implementation of learned mapping and training of 3 layer network. Actual algorithm for a 3-layer network (only one hidden layer) as follows:

Initialize the weights in the network (often randomly)

Do For each example e in the training set (Symptoms)

O = neural-net-output (network, e); forward pass

ISSN: 2278 – 7798

T = Diagnosed output for e

Calculate error (T - O) at the output units

Compute delta_wh for all weights from hidden layer to output layer; backward pass

Compute delta_wi for all weights from input layer to hidden layer; backward pass continued

Update the weights in the network

As the algorithm's name implies, the errors propagate backwards from the output nodes to the inner nodes.

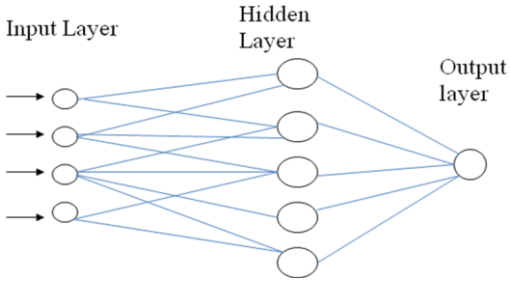

Figure 2: A Typical Neural Network

IV. RESULTS AND ANALYSIS

A. Data Analysis

Symptoms, signals are the data used in medical diagnosis. The data set is obtained from UCI machine learning repository.

Sets were collected in the survey and from UCI machine learning repository for the proposed work. The input values have been assigned based on the comparisons between the two men where one has diabetes and other is non diabetic. Based on this a probability of values is assigned and fed into the system as inputs. The inputs of the system have been designed on the basis of common symptoms of diabetes, hypertension and Obesity.

Before giving the data to the network, the input values should be normalized which is based on the activation function used. The activation function used is log sigmoid. In the proposed work the input values have been normalized between the range 0 and 1. The database is normalized to the input requirements of back propagation network.

The normalized data is used for training the network such that, the data will be in the range of 0 to 1.0. The accuracy of the technique is evaluated by splitting data into training and testing data set. When learning is stopped, the network is evaluated with the data from the test data set. The dataset is divided in to training set and test set. Training a neural network is the process of setting the best weights on the inputs of each of the units. The training set is a part of the input dataset used for neural network training that is for adjustment of network weights. The test set is a part of the input data set used to test how well the neural network will perform on new data. The test set is used after the network is trained, to test what errors will occur during future network application .this set is not used for training. The trained network has been tested with a test set, in which the outcomes are known but not provided to the network, to see how well the training has worked. Training patterns refers to the input data coupled with its desired outputs were prepared. When testing the network, various input data sets are applied to its input layer. The network generates the output. The network's output is compared with the desired outputs. These input data sets and desired outputs collectively known as test samples. The input values have been coded with numerical values ranging between 0 and 1 in order to give a contribution to the network independent from their real absolute values.

B. Performance Evaluation

Results showed that the proposed Medical diagnosis neural network could be useful for identifying the infected person. The Mean Square Error (MSE) and Regression values for the training, validation and Testing.

ISSN: 2278 – 7798

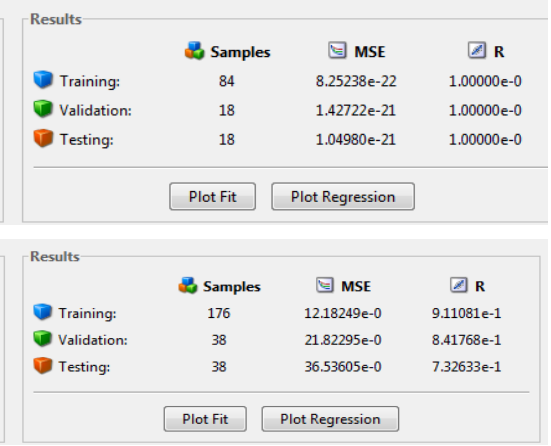

Figure 3: Result of Medical diagnose System

C. Performance Plot

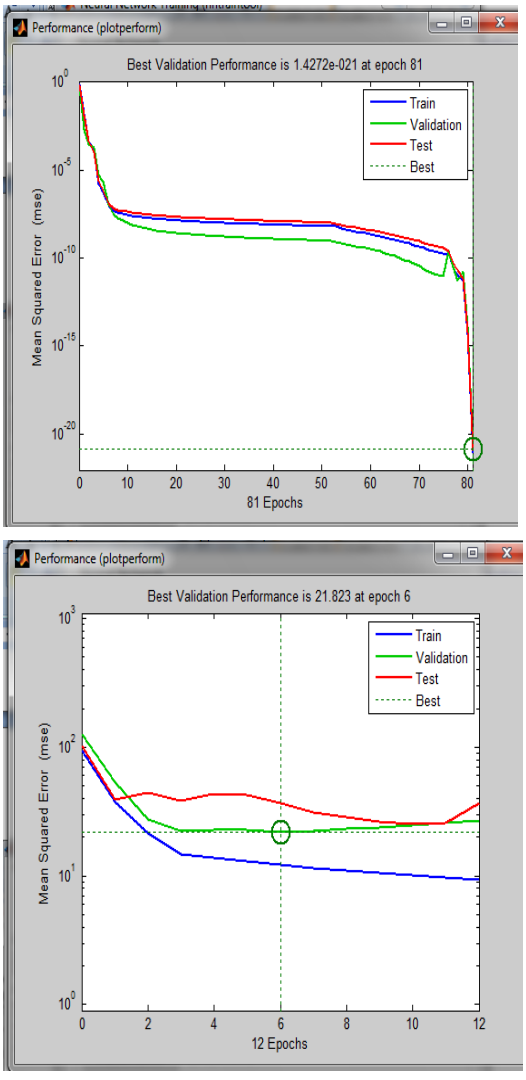

Figure 4: Performance plot of Medical Diagnose System

D. Training State Plot

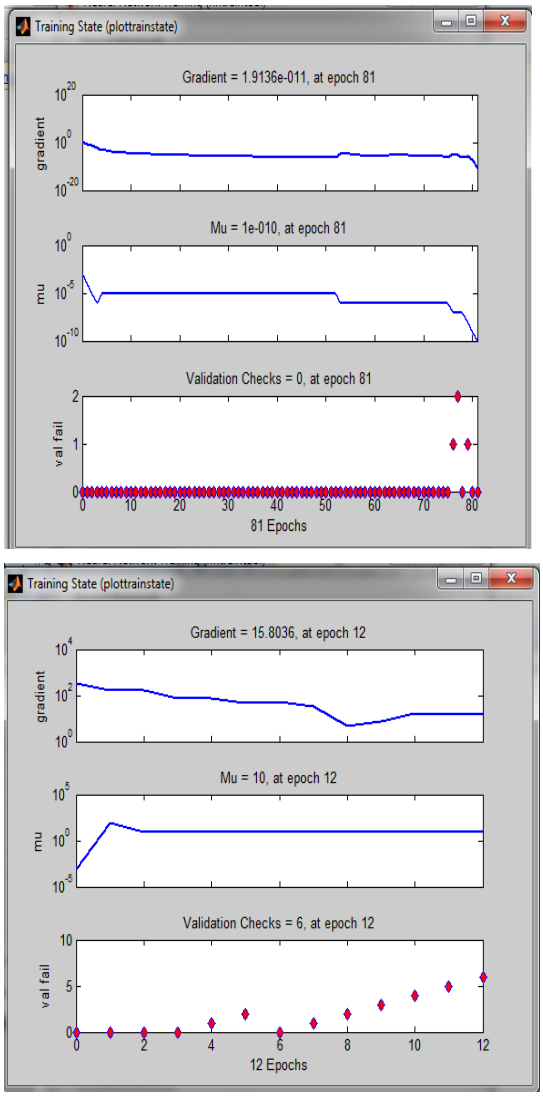

Figure 5: Training State plot of Medical Diagnose System

E. Output of Medical Diagnose System

Disease	Factors		Target Value
Diabetes	**2Hr Glucose**	**Fasting Glucose**	
Normal			0
Diabetes	<140 mg/dl >=200 mg/dl	<110 mg/dl >=126 mg/dl	1

ISSN: 2278 – 7798

	Weight		
Obesity			
Men			
Healthy	<94		0
Obese	>=102		1
Women			
Healthy	<80		0
Obese	>88		1
	Systolic	**Diastolic**	
Hypertension	**Pressure**	**Pressure**	
Normal			0
Hypertension	90-119 mmHg	60-80 mmHg	1
	>=140 mmHg	<90 mmHg	

Table 1: Output of the Study Medical Diagnose system

V. CONCLUSION

This study aimed to evaluate artificial neural network in disease diagnosis. The feed-forward back propagation neural network with supervised ;learning is proposed to diagnose the disease. Artificial neural networks showed significant results in dealing with data represented in symptoms. Results showed that the proposed diagnosis neural network could be useful for identifying the infected person.

REFERENCES

[1] Jaafar, S. F. & Ali, D. M. (2005). Diabetes Mellitus forecast using Artificial Neural Network. Paper presented at the Asian Conference on Sensors and the International Conference on New Techniques in Pharmaceutical and Biomedical Research processing. http:/www.computersociety/ieee/

[2] S. Moein, S.A. Monadjemi, & P. Moallem (2008) "A Novel Fuzzy- Neural Based Medical Diagnosis System", World Academic of Science, Engineering and technology 37 2008.

[3] Shreevastava M. And Gupta A. (2011) " Medical Diagnosis using Back Propagation Algorithm", International Journal of Emerging Technology and advanced Engineering, ISSN 2250-2459, Vol. 1, Issue 1, November 2011.

[4] Sumathy, Mythili, Kumar P., M T and Kumar R. (2010) "Diagnosis of Diabetes Mellitus based on Risk Factors", International Journal of computer Applications(0975-8887), Vol. 10-No.4, November 2010.

[5] Kadhim Qeethara (2011), "Artificial Neural Networks in Medical Diagnosis". International Journal of Computer Science, Vol. 8, Issue 2, March 2011.

[6] Sumathi B., Santhakumaran A. (2011), "Pre-Diagnosis of Hypertension using Artificial Neural Network", Global Journal of Computer Science and Technology, Vol. 11, issue 2, version 1.0, February 2011.

ISSN: 2278 – 7798