

A Survey On Prediction of Heart Disease for Diabetic Patients

IPOONAM PAWAR2 DEEPALI VORA

B.E Computer and M.E,Information Technology,
Pursuing,

B.E Computer and M.E,Information Technology

VidyalankarInstitute of Technology, Wadala, MumbaiVidyalankar Institute of Technology, Wadala, Mumbai

Abstract-Data mining techniques have been widely used in clinical decision support systems for prediction and diagnosis of various diseases with good accuracy. Neural Networks are one of many data mining analytical tools that can be utilized to make predictions for medical data. One of the most important applications of systems is in diagnosis of heart diseases because it is one of the leading causes of deaths all over the world. Diabetes is a chronic disease that occurs when the body cannot produce enough insulin .Diabetes is a disease which causes serious health complications including heart disease, kidney failure and blindness .By using complex tests conducted in labs, almost all system can predict diabetes and heart diseases. Based on risk factor like age, diabetes, hypertension, smoking, high cholesterol, family history, tobacco, alcohol intake, fatness or physical inactivity, etc, none of the system predicts heart diseases. By using major risk factors the proposed system presents a technique for prediction of heart disease for diabetic patients. This technique includes two most successful data mining tools, genetic algorithm and neural networks. The proposed system will be implemented as a web based application, where user will give answers to the predefined questions. The system will retrieve the data from stored database collection and compares the user values with trained data set using Multilayer perceptron neural network. Back propagation algorithm will be used to train the network using the weights optimized by Genetic algorithm.

Keywords- Data mining, Neural Networks, Back propagation algorithm, Genetic algorithm

I. Introduction

Heart is the important part of our body. Life is dependent on workingof heart. Heart disease is a disease that affecting on the operation of heart. There

are number of reasons which increases risk of Heart disease. Heart diseases are the number one cause of death globally: more people die annually from Heart diseases than from any other cause [1][3][4]. Diabetes is a chronic disease which causes serious health complications including heart disease. Diabetes also increases the risk of micro-vascular damage and macro-vascular complications [9].

There are several methods in the literature individually to diagnosis diabetes or heart disease. There is no automated diagnosis method to diagnose Heart disease for diabetic patient based on risk factors such as age, diabetes, hypertension, smoking, high cholesterol, family history, tobacco, alcohol intake, fatness or physical inactivity.[1][5][7].

Heart disease patients have lot of these visible risk factors in which can be used very effectively for diagnosis. System based these risk factors would not only help medical professionals but it would give patients a warning about the probable presence of heart disease even before he visits a hospital or goes for costly medical checkups[1].In this proposed system, the advantages of neural network and genetic algorithm are combined to predict the riskof heart disease for diabetic patient . Genetic algorithm is an optimization algorithm and It can find out good solutions to problemsacceptably quickly. When datasets designed, a neural network have ability to give a solution because it can train itself from the datasets, so that By combining the optimization techniqueof genetic algorithm with neuralnetwork, ,a

system give better predication of heart disease based on risk factors.

II. Materials And Methods

A. Data Mining

- Data mining is used to discover patterns , analyzing enormous sets of data and extract medical data using complex algorithms to help make better decisions [1].
- Data mining techniques are used to diagnosis of many diseases like diabetes [9], heart disease, cancer ,stroke and many data mining techniques have been used in the diagnosis of heart disease with good accuracy.
- Data mining having different techniques such as kernel density, naïvebayes, decision tree, neural network, bagging, and support vector machine for prediction and diagnosis of heart diseases [7].

B. Neural Network

Neural networks is an information processing structure consisting of multiple numbers of processing elements called node, they are interconnected via unidirectional signal channels called connections.

The Neural Network can be classified in two main groups according to the way they learn,

I. Supervised learning:

The technique where the network is learned using known output is called Supervised learning e.g.:Single-layerperceptron,Multi-layer perceptron,Reinforcement Learning

2. Unsupervised learning:

Unsupervised networks learn by identifying special features in the problems they are exposed to.

e.g.: Self-organizing feature maps.[1][8]

C.Multi-layer Perceptron Neural Networks (MLPNN)

A neural network consists of layers of interconnected artificial neurons, as shown in Figure 1[6]. A neuron in a neural network is sometimes called a “node”; all these terms mean the same thing, and are interchangeable. A multilayer perceptron neural network consists of a layer of input node, one or more layers of hidden nodes, and one output layer of node.[4]

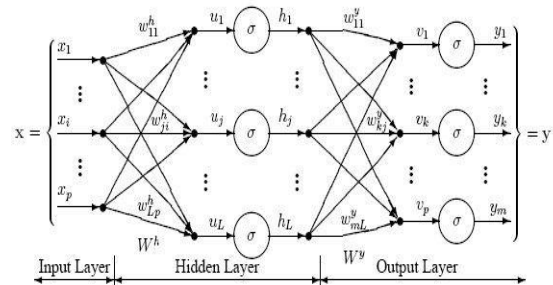


Figure 1: Structure of MLPNN [6]

This structure is called multilayer because it has a layer of processing node (i.e., the hidden nodes) in addition to the output nodes.. Typically, the layers are fully connected, meaning that all nodes at one layer are connected with all nodes at the next layer. So, this means that all input nodes are connected to all the nodes in the layer of hidden nodes, and all the nodes in the hidden layer are connected to all the output nodes [6].

D.Back Propagation

- The general idea with the back propagation algorithm is to use gradient descent to update the weights so as to minimize the squared error between the network output values and the target output values[1][2][4].
- 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.[10]

E.Genetic Algorithm

- Genetic Algorithm (GA) is an optimization techniques inspired by natural selection and natural genetics.
- Its main advantage is that GA only uses the fitness function but not gradient or other attached information in the optimizing process. The GA's fitness function is used to estimate the individuals' optimization degree by optimizing computation. Those individuals who have much greater fitness will have more opportunities to be duplicated to the next generation [1][2][3].
- Genetic Algorithm is consists of Selection, Cross over, Mutation, Accepting.
 - a) Select two parents' chromosomes from a population according to their fitness.
 - b) Create a pair of offspring chromosomes by applying the genetic operators - crossover and mutation.
 - c) Place the created offspring chromosomes in the new population[1][5]

III. Proposed System

In the proposed system data mining intelligent technique is used. The system predicts more accurately the disease with reduced number of attributes. The propose system will be take 5 main risk factors, diabetic is common risk factor and then system will perform preprocessing to remove noisy data, duplicate records, missing data and inconsistent data. System uses Normalization of the input values for each attribute measured in the training tuples, to speed up the learning process. Then Back propagation algorithm will be used to train the network using the weights optimized by Genetic algorithm.

The following is proposed system for prediction of Heart Disease for Diabetic Patients

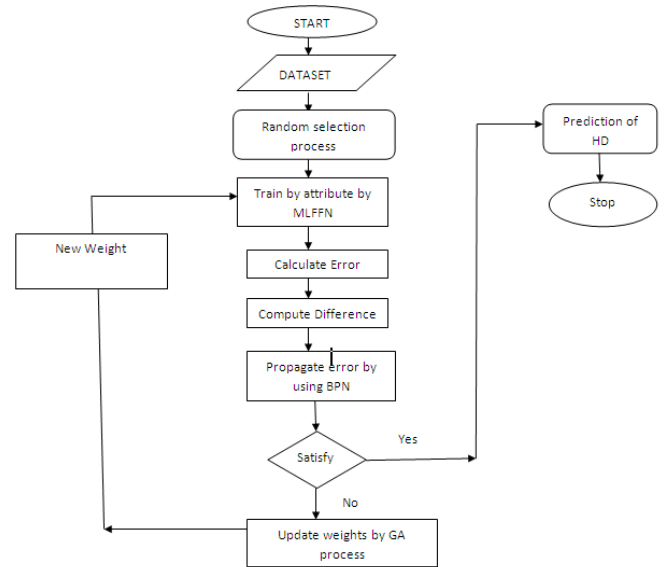


Figure 2: Proposed system

Proposed Algorithm

- 1) Initialize the process for predicting heart disease.
- 2) Extract the risk factors of patient's details, which consists of patient information.
- 3) Selection process start with assigning the weight randomly to each attributes.
- 4) Each attributes is given in the hidden layer of neural network and it will start process of training.
- 5) Compare the actual and desired output.
- 6) Calculate the error in each neuron.
- 7) Propagates error by using the Back propagation algorithm.
- 8) If the training process is satisfied, then it will go to the exact heart attack prediction of patients.
- 9) Otherwise it will be go to the genetic algorithm for adjusting the weight and return to the step 3.

Genetic Algorithm:-

- a) Initialize population.
- b) Evaluate the fitness of all the chromosomes of the population.
 - c) The finest chromosomes will be selected to reproduce, using mutation and crossover.

d) Substitute the worst chromosomes of the previous generation by the new produced chromosomes.

10) Finally the result of patients will be predicting the heart attack via affect patients or not affect the patient.

Conclusion

In this paper, we have survey on prediction of heart disease for diabetic patients by using Data mining, Neural Networks, Back propagation algorithm, Genetic algorithm to predict better prediction of heart disease based on risk factors.

REFERENCE

[1] Syed Umar Amin¹, Kavita Agarwal², Dr. Rizwan Beg "Genetic Neural Network Based Data Mining in Prediction of Heart Disease Using Risk Factors", Proceedings of 2013 IEEE Conference on Information and Communication Technologies (ICT 2013).

[2] Bhuvaneswari Amma N.G. "Cardiovascular disease prediction system using genetic algorithm and neural network", Computing, Communication and Applications (ICCCA), 2012 International Conference.

[3] Du-Yih Tsai, "Classification of heart diseases in ultrasonic images using neural networks trained by genetic algorithms," Signal Processing Proceedings, 1998. ICSP '98. 1998 Fourth International Conference on (Volume:2)

[4] H. Yan, et al., "Development of a decision support system for heart disease diagnosis using multilayer perceptron," in Proc. of the 2003 International Symposium on, vol. 5, pp. V-709- V-712.

[5] T.Manju¹, K.Priya², R.Chitra,"Heart Disease Prediction System Using Weight Optimized Neural Network", Computer Science Department, IJSCE, Vol 2 Issue 5 May 2013

[6] JyotiSoniUjma Ansari "Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction" International Journal on Computer Science and Engineering (IJCSE), 2011

[7] Shantakumar B.Patil and Y.S.Kumaraswamy, "Intelligent and Effective Heart Attack Prediction System Using Data Mining and Artificial Neural Network", European Journal of Scientific Research, Vol.31, No.4, pp.642-656, 2009.

[8] Miss. Chaitrali S. Dangare¹, Dr. Mrs. Sulabha S. Apte , " A Data Mining Approach For Prediction Of Heart Disease Using Neural Networks," IJCET, Volume 3, Issue 3, October - December (2012),

[9] T. Porter and B. Green, "Identifying Diabetic Patients: A Data Mining Approach," Americas Conference on Information Systems, 2009.

[10] Nabeel Al-Milli , " Backpropagation Neural Network For Prediction Of Heart Disease," Journal of Theoretical and Applied Information Technology 10th October 2013. Vol. 56 No.1