Artificial Neural Networks for Diagnosis of Hepatitis Disease

Lale OZYILMAZ Tulay YILDIRIM
Yildiz Technical University
Electronics and Communications Eng. Dept.
34349 Besiktas-İSTANBUL / TURKEY
E-mail: ozyilmaz@yildiz.edu.tr, tulay@yildiz.edu.tr

Abstract- **Recently, neural networks have become a very important method in the field of medical diagnostic. The objective of this work is to diagnose hepatitis disease by using different neural network architectures. Standard feedforward networks and a hybrid network were investigated. Results obtained show that especially the hybrid network can be successfully used for diagnosing of hepatitis.**

I. INTRODUCTION

Diagnosis of the diseases is one of the most important problems in medicine. Medical diagnostics is quite difficult and visual task which is mostly done by expert doctors. Two problems are the most common in the field of automatic diagnostic: the selection of necessary parameter set for right diagnostics and forming of steady and powerful algorithm which doesn't require long time to run [1]. In recent times, neural networks have become a widely used method for this purpose. A reasonably good solution to medical problems could be given by the neural network algorithms. In this work, three neural network algorithms have been investigated for diagnosis of hepatitis diseases. Two standard feedforward network, Multilayer Perceptron (MLP) structure trained by standard backpropagation and Radial Basis Function (RBF) network structure trained by OLS algorithm, and a hybrid network, Conic Section Function Neural Network (CSFNN) with adaptive learning, were applied to hepatitis data. The results have been compared with the previous work [2].

II. HEPATITIS DATABASE

Hepatitis disease is inflammation and damage to hepatocytes in the liver and can be caused by infections with viruses, bacteria, fungi, exposure to toxins such as alcohol and autoimmunity. Usually, the liver can handle significant amounts of damage, and the liver function is still effective. However, it will decline if the disease is not fully controlled at an early stage. Hepatitis may be acute with recovery within six months. However, it also may lead to death if significant decompensation occurs. In addition, there are five types of hepatitis: hepatitis A, hepatitis B, hepatitis C, hepatitis D and hepatitis G. [2]

Hepatitis database [2,3] contains 155 samples in total with 75 of them having missing attributes. All samples have 19 attributes that are shown as follows:

1- Age : ranges from 20 to 72 years old;
2- Sex : male or female, and is represented by 1 or 2 respectively;
3- Steroid : value 1 for no, value 2 for yes;
4- Antiviral : value 1 for no, value 2 for yes;
5- Fatigue : value 1 for no, value 2 for yes;
6- Malaise : value 1 for no, value 2 for yes;
7- Anorexia : value 1 for no, value 2 for yes;
8- Liver Big : value 1 for no, value 2 for yes;
9- Liver Firm : value 1 for no, value 2 for yes;
10- Spleen Palpable : value 1 for no, value 2 for yes;
11- Spiders : red capillary tufts in the skin that blanch on pressure, and is represented by 1 or 2 respectively;
12- Ascites : accumulation of fluid in the abdominal cavity, and is represented by 1 or 2 respectively;
13- Varices : dilated veins, and is represented by 1 or 2 respectively;
14- Bilirubin : a bile pigment cleared from the blood by the liver;
15- Alkaline phosphatase : protein found in bile duct cell membranes;
16- Aspartate transaminase (SGOT) : enzymes that catalyze protein transformations within hepatocytes;
17- Albumin : a protein in the serum that transports substances such as drugs and prevents leakage of fluid into the surrounding tissues;
18- Pro-time : the pro-thrombin time in serum
19- Histology : value 1 for no, value 2 for yes;

There are two classes of the hepatitis disease: death and life. First class has 13 instances and second class has 67 instances.

III. NEURAL NETWORKS ARCHITECTURES

MLPs are feedforward neural networks trained with the standard backpropagation algorithm. They are supervised networks so they require a desired response to be trained.

0-7803-7898-9/03/$17.00 ©2003 IEEE

They learn how to transform input data into a desired response, so they are widely used for pattern classification. With one or two hidden layers, they can approximate virtually any input-output map. They have been shown to approximate the performance of optimal statistical classifiers in difficult problems. Most neural network applications involve MLPs.

The basic MLP building unit is a simple model of artificial neuron. This unit computes the weighted sum of the inputs plus the threshold weight and passes this sum through the activation function (usually sigmoid). In a multilayer perceptron, the outputs of the units in one layer form the inputs to the next layer. The weights of the network are usually computed by training the network using the backpropagation algorithm. [4,5]

RBF network comprises one of the most used neural network models. It consists of one hidden layer of basis functions, or neurons. At the input of each neuron, the distance between the neuron center and the input vector is calculated. The output of the neuron is then formed by applying the basis function to this distance. The RBF network output is formed by a weighted sum of the neuron outputs and the unity bias. This corresponds to additional direct connections from the inputs to the output neuron. The parameters of the RBF network consist of the positions of the basis functions, the inverse of the width of the basis functions, the weights in output sum and the parameters of the linear part. [4,5]

In training, the RBF network parameters are tuned so that the training data fits the network model as well as possible. The basis function is usually chosen to be the Gaussian bell function. Although this function is the most commonly used basis function, other basis functions may be chosen.

MLP and RBF networks can be used to solve a common set of problems; such as, classification, function approximation, modeling of dynamic systems and time series, and so on.

CSFNN allows decision surfaces to be adapted between open boundaries as in MLP and closed ones as in RBF [6], providing unification between RBF and MLP networks. The propagation rule for CSFNN, which consists of RBF and MLP propagation rules, is derived using analytical equations for a cone. Let x be any point on the surface of the right circular cone, vertical angle ω which can be any value in the range $[-\pi/2,\pi/2]$, v vertex of the cone and a the unity vector defining the axis of the cone. Thus the equation of the circular cone is

$$(x - v)\,\vec{a} = \cos\omega \|x - v\| \tag{1}$$

The intersection of a three-dimensional cone with vertex v and opening angle 2ω by a plane forms a circle, a parabola, and a straight line in two-dimensional space by varying the opening angle 2ω. The angle changes depending on how high the vertex is. Straight line (hyperplane) and circle (hypersphere) represent the decision borders for MLP and RBF, respectively. Other type of decision borders, such as ellipses and parabolas, also can be obtained and these forms represent the intermediate conic section functions. The propagation rule of the Conic Section Function Neural Network is given by

$$y_j = \sum_{i=1}^{n+1} (x_i - c_{ij})\, a_{ij} - \cos\omega_j \sqrt{\sum_{i=1}^{n+1} (x_i - c_{ij})^2} \tag{2}$$

$$x_{n+1} \equiv 0$$

where a_{ij} refers to the weights for each connection between the input and hidden layer units in an MLP network, and c_{ij} refers to the centre coordinates in an RBF network, i and j are the indices referring to the units in the input and hidden layer, respectively. As can be seen easily, this equation consists of two major parts analogous to the MLP and the RBF. The equation simply turns into the propagation rule of an MLP network, which is the dot product (weighted sum) when the ω is $\pi/2$. Second part of the equation gives the Euclidean distance between the inputs and the centres for an RBF network.

IV. APPLICATIONS AND RESULTS

In this work, different neural networks were used to diagnose hepatitis disease. These were standard feedforward networks, MLP and RBF, and a hybrid network, adaptive CSFNN [7,8]. Training and test sets were formed separately using 80 samples which have no missing attributes. With the aim of demonstrating the future performance of the networks for untrained instances, 5-fold cross-validation method has been used for diagnosing of hepatitis disease. Therefore, database has been divided into 5 subsets: A1, A2, A3, A4 and A5. Subsets A1, A2 and A3 have 3 instances from first class and 13 from second class. Subsets A4 and A5 have 2 instances from first class and 14 from second class. Each subset is tested in turn and remaining four subsets are used for training.

Classification accuracies obtained after the training of subsets A2, A3, A4 and A5 and testing of subset A1 for all algorithms are given in Table I. Note that MLP results change due to random initialisation in every run. Therefore, the average of 10 runs is given in the Table I for MLP.

Table II shows the results when the subsets A1, A3, A4 and A5 were trained and the subset A2 was tested. In a similar way, when the subsets A1, A2, A4 and A5 were trained and the subset A3 was tested for all algorithms, obtained results were given in Table III.

Table IV shows the results when the subsets A1, A2, A3 and A5 were trained and the subset A4 was tested. Finally, results obtained from the training of subsets A1, A2, A3, A4 and testing of subset A5 for all algorithms were given in Table V.

TABLE I
RESULTS FOR TESTING A1 SUBSET WITH DIFFERENT NEURAL NETWORK

Type of neural network	Number of correct classified patterns in test set	Classification accuracy %	Training epochs
MLP	13,4 (average)	83,75 (average)	500
RBF	14	87,5	55
CSFNN	14	87,5	6

TABLE II
RESULTS FOR TESTING A2 SUBSET WITH DIFFERENT NEURAL NETWORK

Type of neural network	Number of correct classified patterns in test set	Classification accuracy %	Training epochs
MLP	13,6 (average)	85 (average)	500
RBF	12	75	59
CSFNN	15	93,75	10

TABLE III
RESULTS FOR TESTING A3 SUBSET WITH DIFFERENT NEURAL NETWORK

Type of neural network	Number of correct classified patterns in test set	Classification accuracy %	Training epochs
MLP	13 (average)	81,25 (average)	500
RBF	14	87,5	59
CSFNN	14	87,5	34

TABLE IV
RESULTS FOR TESTING A4 SUBSET WITH DIFFERENT NEURAL NETWORK

Type of neural network	Number of correct classified patterns in test set	Classification accuracy %	Training epochs
MLP	13,9 (average)	86,875 (average)	500
RBF	14	87,5	58
CSFNN	14	87,5	5

TABLE V
RESULTS FOR TESTING A5 SUBSET WITH DIFFERENT NEURAL NETWORK

Type of neural network	Number of correct classified patterns in test set	Classification accuracy %	Training epochs
MLP	11,2 (average)	70 (average)	500
RBF	14	87,5	57
CSFNN	15	93,75	4

MATLAB Neural Network toolbox was used to obtain the results for MLP and RBF. Another MATLAB code was written for CSFNN.

The comparison of obtained average results for 5-fold cross-validation method and previous results was given in Table VI. Abbreviations in the Table VI are given as follows.

C4.5 : C4.5 decision tree
NB : Naïve Bayes classifier
TAN : Tree Augmented Naïve Dependence
BNND: Bayesian Network with Naïve Dependence
BNNF: Bayesian Network with Naïve Dependence and Feature selection

The results showed that the hybrid structure (CSFNN) has given the best classification accuracy. Since the accuracies of TAN and BNND changes normally in a maximum and minimum range and their maximum results are given they can not be considered as best classifiers (see Table VI).

V. CONCLUSIONS

In the work presented here, three neural network algorithms have been investigated for diagnosis of hepatitis diseases. The results were compared with some statistical methods used in a previous work.

MLP is not a good choice for such problem as the classification accuracy is low. It also does not guarantee the same performance for single run depending on random weight initialisation in training. RBF gives promising results. However, CSFNN has the best classification accuracy for hepatitis diagnosis. This results shows that using a hybrid network CSFNN that combines MLP and RBF is more reliable for the diagnosis.

The compared results shows that neural networks can be used in the problem of diagnosis for hepatitis diseases as efficiently as statistical methods.

TABLE VI
RESULTS FOR 5-FOLD CROSS VALIDATION METHOD

Type of neural network	Classification accuracy %
MLP	81,375 (average)
RBF	85
CSFNN	90
C4.5	83,6 (max)
NB	87,83 (max)
TAN	90,1 (max)
BNND	90 (max)
BNNF	88,76 (max)

REFERENCES

[1] http://www.intermag.kiev.ua/projects/neuraln.html

[2] N.Cheung, *Machine Learning Techniques for Medical Analysis*, Bs.C. thesis, University of Queensland, 2001.

[3] http://www.ics.uci.edu/pub/ml-repos/machine-learning-databases/, 2003.

[4] R.J. Schalkoff, *Artificial Neural Networks*, McGraw-Hill Inc., Singapore, 1997.

[5] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Macmillan College Publishing, New York, 1994.

[6] G.Dorffner, "Unified framework for MLPs and RBFNs: Introducing conic section function networks," *Cybernetics and Systems*, Vol.25, pp.511-554, 1994.

[7] T.Yıldırım, J.S.Marsland, "Improved back propagation training algorithm using conic section functions," *Proceedings of the IEEE Int. Conf. On Neural Networks (ICNN'97)*, Houston, Texas, USA, June 9-12, 1997.

[8] L.Ozyilmaz Basturk, *Developing of Learning Algorithm and Investigating of the Network Performance and Sensitivity for Various Problems in Conic Section Function Neural Networks*, Ph.D. Thesis, Yildiz Technical University, Istanbul, 2000.