Classification of Liver disease using Multilayer Perceptron Neural Network

Geeta Kadu[1], Dr. Ranjana Raut[2]

1 Dept of Electronics Engineering, Shri Datta Meghe Polytechnic, Nagpur, Maharashtra, India

2 Dept of Applied Electronics, SGB Amravati University, Amravati, Maharashtra, India

ABSTRACT

Artificial intelligence is now commonly used in the medical field for diagnosis of various diseases diagnosed due to malfunctioning of specific organs. Artificial Neural Networks (ANN) are effectively implemented for the classification of normal and diseased liver. In this study, The Multilayered Perceptron Neural Network (MLP NNT) Machine Learning Classifier has been used for the classification of three different types of Liver Diseases viz. 'healthy liver', 'fibrotic liver' and 'cirrhotic liver'. UCI liver disorder data obtained from UCI machine learning repository is used for the classification. Performance evaluation is done by varying the hidden layers between 10 to 100. Three parameters Accuracy, Regression and MSE are drawn. Finally data is plotted in graphical way and the optimal results are concluded.

KEY WORDS: *Artificial Neural Networks (ANN), Accuracy, Regression (r), Liver disorder, Multilayered Perceptron Neural Network (MLP NNT), MSE*

1. INTRODUCTION

Due to unhealthy life style, stress and pollution we are damaging our liver. Scarring of liver is fibrosis and when it covers most of the liver results in cirrhosis [1]. A scarred tissue reduces the efficiency of liver. In cirrhosis scar tissues replaces normal, healthy tissue, blocking the flow of blood through the liver and preventing *it* from working properly. Fibrosis is not detected at early stage and usually liver damage is noticed in cirrhotic liver. Though alcohol is not only responsible for cirrhosis but alcohol and presence of hepatitis C virus, for long period, increases the chances of cirrhosis. Liver damage or liver diseases can be detected using liver function tests. Automatic decision support system can be developed using ANN for classification of liver diseases. It can be implemented effectively for diagnosis at preliminary level. Thus, liver diseases can be effectively detected at an early stage. Mortality rate, due to liver diseases, can be reduced.

In this paper an ANN system for the classification of three types of liver states, 'healthy liver', 'fibrotic liver' and 'cirrhotic liver', is explained. MLP NNT uses UCI liver disorder data for classification. Performance of the system is checked by varying hidden layers between 10 to 100. Performance of the system is evaluated using statistical parameters Accuracy, Regression and MSE.

2. LITERATURE REVIEW

It was observed during literature survey that classification techniques can be effectively used for liver disease diagnosis, using clinical data, at the preliminary level. This helps in deciding the further line of treatment and the need of additional tests required.

Survey on different AI methods used for the diagnosis and detection of liver disorder acuteness done by Neshat et. al [3]. Different AI methods discussed in their study are ASAMC, TENN, HNN and fuzzy Hopfield beside fuzzy C-Means. Pattern recognition methods used in classification techniques, such as NB, DT, SVM, LRA, Fuzzy LSSVM, LS-SVM, C4.5, DFPA, FS-SFS, SFS, and LCF, were also discussed. New method for extracting fuzzy rules from trained SVM was explained.

Sindhuja et. al [4] explained data mining techniques, C4.5, Naive Bayes, Decision Tree, Support Vector Machine, Back Propagation Neural Network, for the diagnosis of liver diseases at an early stage. Advantages and disadvantages of each algorithm are explained.

Liang et. al [5] explained how neural networks can be used for binary classification in various fields. BUPA liver disorders data set is used for the prediction liver disorder in male patients. Comparison of five neural networks BPNN, RBFNN, GRNN, PNN and CMTNN explained in this work.

M. Neshat et. al [6] used Fuzzy expert system. Thus system formulates the rules for dividing liver diseases into three categories: low, medium and high. Fuzzy rules are formed by applying the opinion of expert, from medical field, on the clinical dataset.

H.Ratnamala Kirubala et. al [7] proposed Artificial intelligence and data mining techniques for the diagnosis of liver diseases. System is evaluated implementing twenty two classification machine learning techniques. Performance of the each technique was evaluated on the basis of accuracy and computational time. Best performance were given by C4.5 decision tree and the Random Tree algorithms.

Anil Kumar Tiwari et. al [8] implemented BP, RBF, SOM, SVM classifiers for liver disese diagnosis. Initially, liver patient data was evaluated, using univariate analysis and feature selection method, for selecting important parameters. Then artificial networks classifiers were trained and tested. Finally all the implemented classifiers were compared using performance parameters accuracy, mean absolute error, root mean square error and relative absolute error.

José Neves et. al [9] implemented artificial neural network for the classification of liver patients. From the available patient information and clinical data, using a formal framework based on Logic Programming, input variables are selected. Back propagation algorithm was applied for training and testing the network. Using coincidence matrix performance parameters sensitivity and specificity were found for evaluating the performance of the classifier.

Jyun Jie Lin et. al [10] designed a hybrid model for classification of liver diseases. Data is preprocessed first using casebased reasoning approach. Classification is done by designing a decision making system using particle swarm optimization algorithm.

Ebenezer Obaloluwa Olaniyi et. al [11] Implemented back propagation neural network and radial basis function neural for the classification of liver diseases. BUPA liver disorder dataset is used for modeling and testing the networks. Both models are compared on the basis of recognition rate.

Parminder Kaur et. al [12] designed random tree algorithm for the detection four types, fatty liver, Wilson, Inherited, Autoimmune, and Cholestatic, liver diseases. Weka tool is used to find operating characteristics such as: true positive, false positive, true negative and false negative.

Bendi Venkata Ramana et. al [13] explained the use of different machine learning algorithms for designing the automatic liver disease classification. Back propagation Neural Network algorithm, C4.5, Naïve Bayes classifier and Support Vector Machines algorithms are implemented and performance is compared using Accuracy, Precision, Sensitivity and Specificity.

Sana AnsariI et. al [14] implemented supervised and unsupervised neural network models for the diagnosis of hepatitis virus. Supervised NN models used are Feedforward Backpropagation Neural Network and Generalized Regression Neural Network. Unsupervised NN used is self organizing map. Networks are implemplemented and tested on VCI machine learning database.

Naiping Li et. al [15] used general regression neural network (GRNN) model for the classification of alcoholic liver diseases. Clinical parameters of the patients are given as input to the network. Three types of outputs are alcoholic cirrhosis with decompensated liver function, alcoholic liver cirrhosis with compensated liver function and alcohol hepatitis.

S. Madhusudhanan et. al [16] designed Fuzzy based classification system for the diagnosis of hepatitis. Classification rules are extracted using Fuzzy Based Ant Miner Algorithm (FACO). Important rules are generated using heuristic functions.

Mehdi Neshat et. al [17] given liver disease diagnosis systems using Hopfield neural network and fuzzy Hopfield neural network. Accuracy of fuzzy Hopfield neural network found to be more than the Hopfield neural network.

Ghumbre Shashikant Uttreshwar et. al [18] first designed the system based on logical inference for detecting whether the patient is affected with hepatitis B or not. Then generalized regression neural network is used for finding the severity level of hepatitis B.

3. PROPOSED SYSTEM

UCI liver disorder data obtained from UCI machine learning repository is processed for checking missing values. Multilayered Perceptron Neural Network is used as a classifier. Proposed system is represented by the block diagram shown in fig. 1 below:

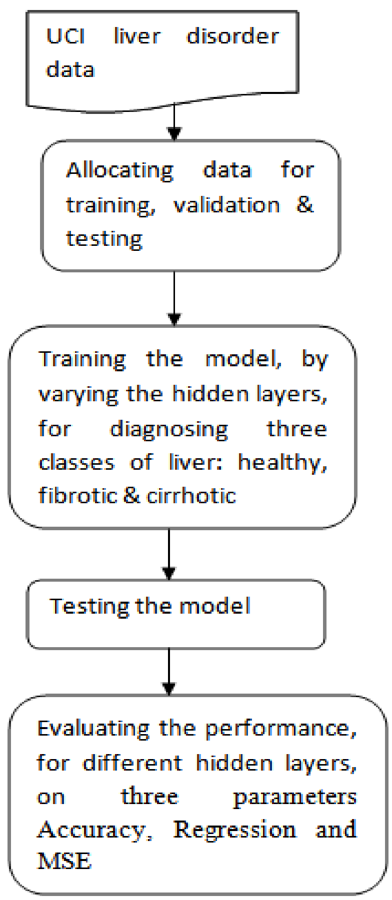

Figure 1: Block Diagram of the System

3.1Data

UCI liver disorder data obtained from UCI machine learning repository is used for the classification. Dataset contains five attributes used to predict the liver class healthy, fibrotic or cirrhotic. These five attributes are mcv, alkphos, sgpt, sgot and gammagt. Attribute Information is given in table 1.

Table 1: Attributes and their description

Sr no.	Attribute	Description
1	mcv	mean corpuscular volume
2	alkphos	alkaline phosphotase
3	sgpt	alamine aminotransferase
4	sgot	aspartate aminotransferase
5	gammagt	gamma-glutamyl transpeptidase

Class attribute gives classification as shown in table 2.

Table 2: Class Attributes

Class attribute	Class
Healthy liver	1
Fibrotic liver	2
Cirrhotic liver	3

3.2 MLP – Multilayered Perceptron Neural Network (NNT)

An Artificial neural networks (ANN) designs model which works similar to that of the biological brains. These models are used to solve difficult computational tasks like the predictive modeling tasks implemented in machine learning. A trained model is built which learns from the past data. Such a trained model can be successfully implemented, for classification, in the medical field for diagnosis [5].

A multilayer perceptron (MLP) is a feedforward artificial neural network model that maps sets of input data onto a set of appropriate outputs. Feedforward back-propagation neural network (FFBPNN) is also called as Back propagation neural network (BPNN).

MLP consists of multiple layers of nodes in a directed graph, with each layer fully connected to the next one. Except for the input nodes, each node is a neuron (or processing element) with a nonlinear activation function. MLP utilizes a supervised learning technique called back propagation for training the network [9]. MLP is a modification of the standard linear perceptron and can distinguish data that are not linearly separable. It consists of three layers, a visible

 ISBN:978-93-87793-54-5

input and an output layer and hidden layer, of nonlinearly-activating nodes. Hidden layer can be single or in multiples. Each node in one layer connects with a certain weight to every node in the following layer. The choice of activation function in the output layer is strongly constrained by the type of problem that are being modeled. Once the network is formed it is required to be trained for the given data set to produce required output. This is training stage. Using forward pass the network input propagates to produce the output value. This output is compared with the expected output and the error is calculated. This error is then propagated back through the network, one layer at a time, and the weights are updated accordingly. This is backpropogation algorithm. After many repetitions or epochs the network is trained. This trained network can be used for the prediction of new data [8].

4. DESIGN OF THE NETWORK

For the diagnosis of three types of liver diseases Multilayered Perceptron Neural Network or feedforward back-propagation neural network is implemented. Five attributes of UCI liver disorder data are taken as the neurons at the input layer. In the output layer there are three neurons indicating three liver states: healthy, fibrotic and cirrhotic. Levenberg Marquardt algorithm is used. Peculiarity of this algorithm is it requires more memory but less execution time. Training automatically stops when generalization stops improving, which is indicated by an increase in MSE of the validation samples. Implemented MLP network is shown in fig. 2. Following algorithm is used:

1. Variables in the data set are segregated as input and output.
2. Five inputs, mcv, alkphos, sgpt, sgot and gammagt, of the liver disorder data are given to the input layer of the network.
3. Three classes of output are converted to 1 0 0, 0 1 0 and 0 0 1
4. Data of 215 patients is divided into 50% data for training and 50% for testing
5. Network is trained, using 50% data separated for training,
6. Designed network is tested on the 50% for data separated for testing.
7. by varying hidden layers between 10 to 100.
8. Performance evaluation of MLP NNT is drawn by varying the hidden layer between 10 to 100.
9. Performance of the network is tested using three parameters: Accuracy, Regression (r) and MSE.

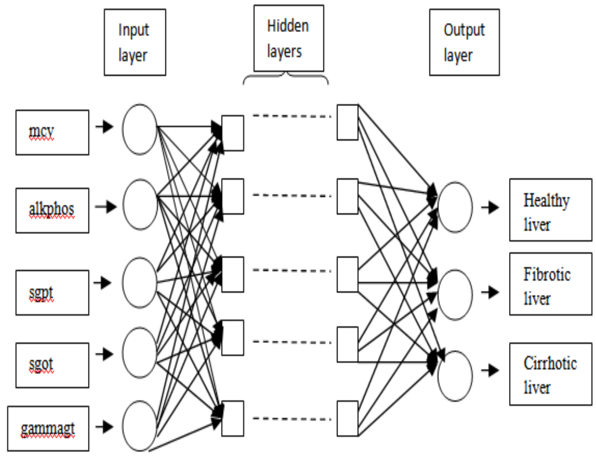

Figure 2: MLP Network

5. EVALUATION

Performance parameters of the selected classifier – MLP, are defined as follows:

Accuracy: Accuracy is the percent of correct classifications given by equation (1).

Accuracy = 1 - Error rate

OR

Accuracy = (TP+TN) / (TP+TN+FP+FN) ----------------- (1)

Where:

TP - True Positive

TN - True Negative

FP – False positive

FN – False negative

Mean square error (MSE): It is the average squared difference between the outputs and targets. Formula MSE is given by equation (2):

$$MSE = 1/n \sum_{i=1}^{n}\left(Yi - \hat{Y}_i\right)^2 \qquad \text{---------------} \quad (2)$$

Where:

$\hat{Y}$ - Vector of n predictions generated from a sample of *n* data points for all variables

Y –Vector of observed values of the variable being predicted

Regression: It is a statistical tool which gives the degree to which the variables are correlated to each other. In other words it gives the probable change in one variable for the given amount of change in another. R value of '1' indicates close relationship and '0' indicates random relationship. The degree to which the variables are correlated can be given by the regression line with equation given by equation (3):

$$Y_e = a + b X \qquad \text{----------------} \qquad (3)$$

Where:

Y_e - dependent variable

X - independent variable

a & b - unknown constants that determine the position of the line.

6. RESULT & DISCUSSION

The developed application software efficiently performed for the classification of three types of liver states viz: healthy, fibrotic and cirrhotic. In this application software MLP is used. Performance of MLP is evaluated using performance parameters accuracy, MSE, and regression by changing hidden layers from 10 to 100. Values of performance parameters accuracy, regression (r) and MSE, for hidden layers ranging from 10 to 100, are given in table 3.

Table 3: Performance of MLP

Sr. No	No of Hidden Layer	Accuracy	Regression	MSE
1	10	98.1308	0.9833	0.028611
2	20	95.3271	0.95858	0.040727
3	30	94.3925	0.93972	0.094877
4	40	95.3271	0.93754	0.057617
5	50	96.2617	0.96149	0.033415
6	60	96.2617	0.94186	0.042263
7	70	94.3925	0.9309	0.061936
8	80	88.785	0.92799	0.091095
9	90	71.9626	0.67278	0.11134
10	100	95.3271	0.92625	0.078641

Values of the parameter regression for different hidden layers are graphically represented in fig. 3.

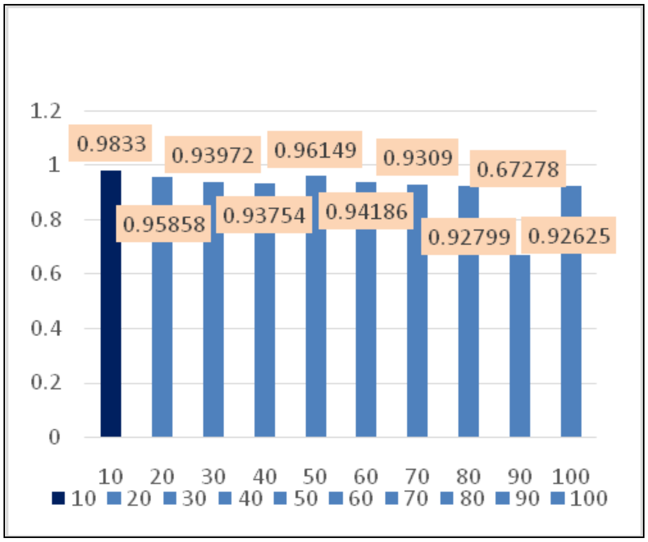

Figure 3: Comparison of Regression for different hidden layers

Values of the parameter accuracy for different hidden layers are graphically represented in fig 4.

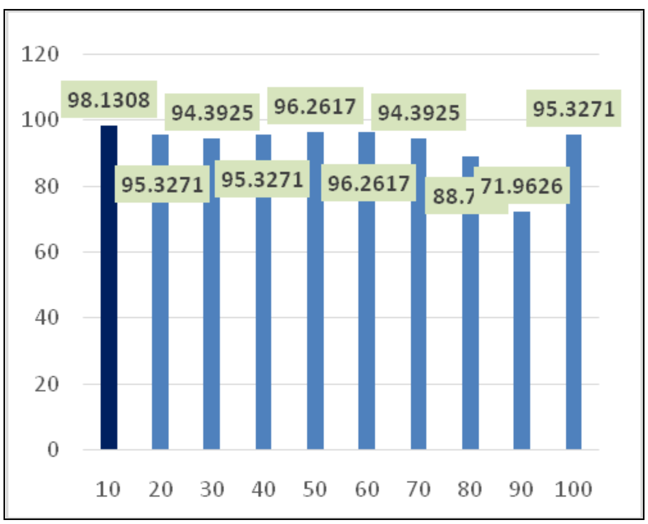

Figure 4: Comparison of accuracy for different hidden layers

ISBN:978-93-87793-54-5

Network is tested for different hidden layers. Best performance is achieved for hidden layer layer 10. The optimal performance, accuracy = 98.1308%, R=0.9833 and MSE = 0.028611, has been achieved for hidden layer = 10. Optimal performance of the network is shown in fig 5.

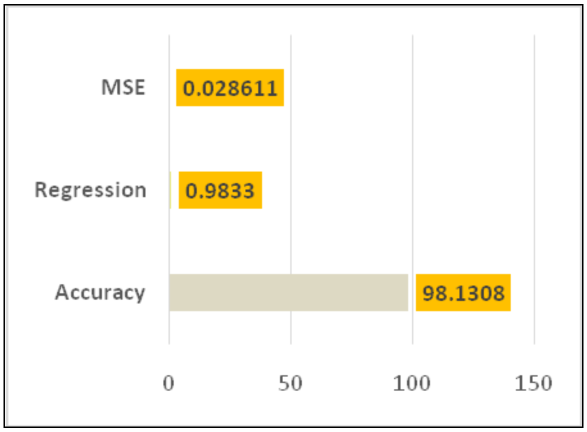

Figure 5: Optimal performance for the hidden layer 10

Mean square error was lowest, 0.028611, for hidden layer 10 at epoch 9. Best validation performance is shown in fig. 6.

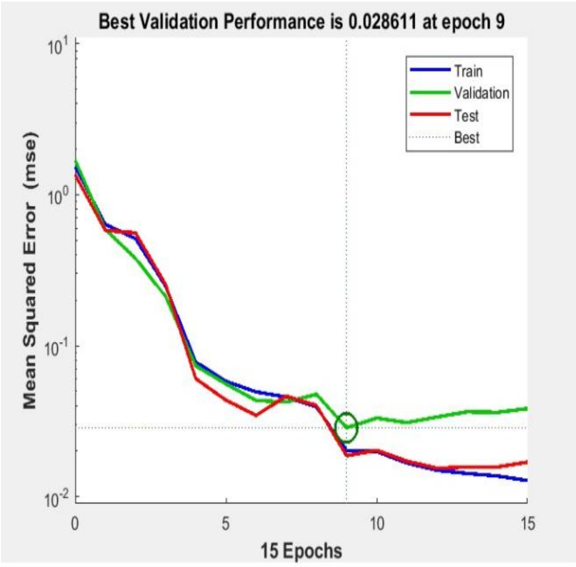

Figure 6: Best validation performance

Fig. 7 represents the highest regression value achieved for hidden layer 10.

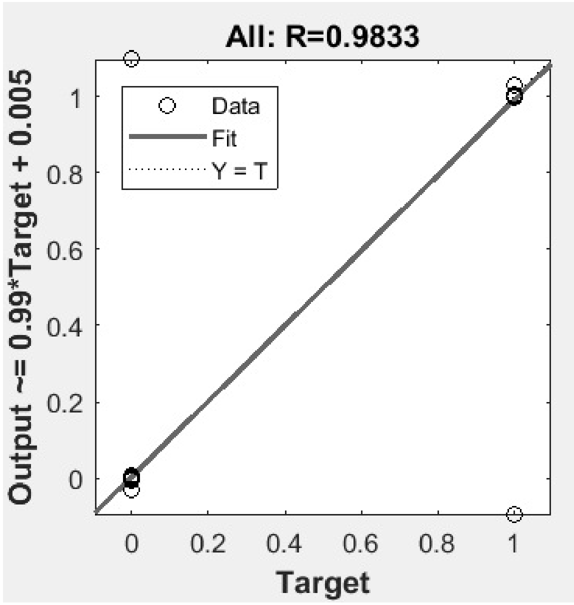

Figure 7: Best correlation between output and target

7. CONCLUSION

Incorrect diagnosis, after examining the clinical reports, may lead to the incorrect medication. In this paper an artificial intelligence based automatic system is explained for the diagnosis of liver diseases by inputting the clinical data of the patient to the system. In this paper, the MLP Classifier has been used for the classification of three different types of liver healthy, fibrotic and cirrhotic. We can state that similar systems can be implemented for the diagnosis of more type of liver diseases or can also be used for the diagnosis of other diseases. Such type of automatic systems can be implemented where expert diagnosis is not available for primary diagnosis. Algorithm used in this paper gives better accuracy as compared to the similar machine learning algorithms implemented in previous work.

REFERENCES

[1] https://www.hepatitis.va.gov/patient/basics/fibrosis-cirrhosis.asp date: 3/3/18

[2] Miss. Hemalata Vaidya, Miss. S. K. Chaudhari and Mr. H. T. Ingale, "Literature review on liver disease classification", IJARIIE, Vol-3 Issue-3 2017

[3] Mehdi Neshat, Ali Adeli, Azra Masoumi, "A Survey on Artificial Intelligence and Expert System for Liver Disorders", ARPN Journal of Systems and Software, VOL. 3, NO. 2, February 2013

[4] D.Sindhuja, R. Jemina Priyadarsini, "A Survey on Classification Techniques in Data Mining for Analyzing Liver Disease Disorder", IJCSMC, Vol. 5, issue 5, May 2016, pg.483 – 488

[5] P. Jeatrakul and K.W. Wong, "Comparing the Performance of Different Neural Networks for Binary Classification Problems", IEEE Eighth International Symposium on Natural Language Processing, 2009

[6] M.Neshat, M.Yaghobi, M.B.Naghibi and A.Esmaelzadeh, "Fuzzy Expert System Design for Diagnosis of liver disorders", International Symposium on Knowledge Acquisition and Modeling, 2008

[7] H. Ratnamala Kiruba, Dr G. Tholkappia Arasu, "An Intelligent –Agent Based Framework for Liver Disorder Diagnosis Using Artificial Intelligence Techniques", Journal of Theoretical and Applied Information Technology, 10th November 2014. Vol. 69

[8] Anil Kumar Tiwari, Lokesh Kumar Sharma and G. Rama Krishna, "Comparative Study of Artificial Neural Network based Classification for Liver Patient", Journal of Information Engineering and Applications, Vol.3, No.4, 2013

[9] José Neves, Adriana Cunha, Ana Almeida, André Carvalho, João Neves, António Abelha, José Machado and Henrique Vicente, "Artificial Neural Networks in Diagnosis of Liver Diseases", Springer International Publishing Switzerland, pp. 71–80, 2015

[10] Jyun Jie Lin and Pei-Chann Chang, "A Particle Swarm Optimization Based Classifier for Liver Disorders Classification", Memetic Computing Society, ICCP 2010 Proceedings

[11] Ebenezer Obaloluwa Olaniyi, Khashman Adnan, "Liver Disease Diagnosis Based on Neural Networks", Advances in Computational Intelligence, ISBN: 978-1-61804-343-6

[12] Parminder Kaur and Aditya Khamparia, " Classification of Liver Based Diseases Using Random Tree ", International Journal of Advances in Engineering & Technology, June, 2015

[13] Bendi Venkata Ramana, Prof. M.Surendra Prasad Babu and Prof. N. B. Venkateswarlu, "A Critical Study of Selected Classification Algorithms for Liver Disease Diagnosis", International Journal of Database Management Systems (IJDMS), Vol.3, No.2, May 2011

[14] Sana AnsariI, Imran ShafiI, Aiza Ansari, Jamil AhmadI and Syed Ismail Shah, "Diagnosis Of Liver Disease Induced By Hepatitis Virus Using Artificial Neural Networks", IEEE 2011

[15] Naiping Li, Yongfang Jiang, JinMa, Bo He,Wei Tang,Mei Li, Qing Huang and Ting Yuan, "Computational Classification and Diagnosis of Alcoholic Liver Diseases Using General Regression Neural Network", Hindawi Publishing Corporation, Mathematical Problems in Engineering, Volume 2014, Article ID 524621

[16] S.Madhusudhanan, Marcus Karnan and K.Rajivgandhi, "Fuzzy Based Ant Miner Algorithm in Data Mining for Hepatitis", 2010 International Conference on Signal Acquisition and Processing

ISBN:978-93-87793-54-5

[17] Mehdi Neshat and Abas E.Zadeh, "Hopfield Neural Network and Fuzzy Neural Network For Diagnosis of Liver Disorders" IEEE 2010

[18] Ghumbre Shashikant Uttreshwar and Dr. A.A. Ghatol," Hepatitis B Diagnosis Using Logical Inference And Generalized Regression Neural Networks", IEEE International Advance Computing Conference (IACC 2009)

[19] Sana Ansari, Imran ShafiI, Aiza Ansari, Jamil Ahmad and Syed Ismail Shah, "Diagnosis Of Liver Disease Induced By Hepatitis Virus Using Artificial Neural Networks", IEEE 2011