

Hepatitis Disease Diagnosis Using Mixture of Expert

Suchitra Kumari

Dept. of Biotechnology, Birla Institute of Technology, Mesra, Ranchi, Jharkhand, India

Abstract

Automated diagnosis of diseases has forever been of interest as an interdisciplinary study among computer and medical science researchers. Detection of hepatitis is really a big problem for general practitioners. An expert doctor commonly takes decisions by evaluating the current test results of a patient or by comparing the patient with other patients with the same condition with reference to the previous decisions. In this study, various models were generated by using mixture of experts as a classification method. Further, Model having very good accuracy of 97.37% with least minimum square error was selected for the prediction of disease. This approach can be used for easy diagnosis of hepatitis for a large number of populations by incorporating the profile of more samples in the training stage.

Keywords

Hepatitis diagnosis; UCI repository; Mixture of expert; Computational intelligence; Minimum square error.

I. Introduction

Hepatitis is the fifth most death causing diseases after heart disease, stroke, chest disease and cancer [1]. Worldwide, 1.5 million deaths have been estimated in each year [2]. It is the serious health problem in which destruction of the liver cells causes cirrhosis and sometime extended to liver cancer too. Various risk factors for hepatitis includes blood transfusions, tattoos and piercing, drug abuse, haemodialysis, health workers, and sexual contact with hepatitis carrier [3].

Early stage diagnosis of hepatitis is very difficult in general population due to the lack of regular routine checkup as well as awareness. Therefore, Medical diagnosis is quite difficult and totally depends on visual task done by expert doctors based on their expertise. An expert doctor commonly takes decisions by evaluating the current test results of a patient or the expert doctor compares the patient with other patients with the same condition by referring to the previous decisions [2]. Therefore, various machine learning and data mining techniques have been widely exploited for the automatic diagnosis of hepatitis. However differences in accuracies of classification were observed in all the cases. Algorithms like C4.5, Naïve Bayes (NB), Bayesian network with naïve dependence and feature selection (BNMF), Bayesian network with naïve dependence (BNND) report 83.6%, 88.7%, 90.0% accuracy [4]. Further, Accuracy in methods like 15 Neural Network (15NN) with Manhattan, weighted 9 Neural Network (9NN) and 18NN with Manhattan was reported as 89.0%, 90.2% and 90.2% respectively [5]. Several method like ARTIST Sea Ice (ASI), Classification And Regression Tree (CART), Learning Vector Quantization (LVQ), Fisher discriminant analysis, Automatic Speech Recognition (ASR), Quadratic discriminant analysis (QDA), Naïve Bayes (NB) along with Semi-NB, and Linear discriminant analysis (LDA) had the accuracy of 82.0, 82.7, 83.2, 84.5, 85.0, 85.8, 86.3 and 86.4 % respectively [6]. Multilayer Perceptron with Backpropagation and Radial Bias Function (RBF) combined with Tooldiag method had shown the accuracy of 77.4% and 79.0% respectively, whereas the accuracy of 88.4% was observed in the case of Finite-State Machine (FSM) without rotations [7]. MultiLayer Perceptron has lead to 74.3% accuracy

[4] while GRNN gave 80.0% accuracy only [7]. Importantly, Fuzzy System-Artificial Immune Recognition System (FS-AIRS) alone has shown the accuracy of 92.5% [8], while FS-fuzzy-AIRS in combination with fuzzy algorithm was increased and resulted in 94.1% [9]. Linear Discriminant Analysis –Adaptive Network based on Fuzzy Inference System (LDA-ANFIS) had also shown the same accuracy as previous one [10]. By using Multilayer Neural Network (MLNN)(MLP) along with Levenberg Marquardt (LM), Principle Component Analysis (PCA) along with Least Square Support Vector Machine (LSSVM), Genetic Algorithm - Support Vector Machine (GA-SVM) and Support Vector Machine –Simulated Annealing (SVM-SA) obtained 91.8, 95.0, 89.6, 96.2 % accuracy [2, 11-13].

In this study, we proposed a Mixture of Expert method for the hepatitis diagnosis with the help of user friendly Graphical User Interface tool. The advantages of using machine learning approaches in medical diagnosis have costs to decrease and caused diagnosis accuracy to increase.

II. Material & Methods

A. Datasets for Hepatitis Diagnosis

Hepatitis dataset collected from University of California (UCI) machine learning repository database was used in this study. This is the popular dataset used for study of machine learning algorithm in diagnosis of hepatitis disease. This hepatitis disease dataset requires determination of whether patients with hepatitis will either live or die. There are 155 large set of samples having 19 attribute or feature [Table 1] in this dataset. Out of 155 samples, 2 different classes were shown, die with 32 cases and live with 123 cases. This dataset is generally used by researcher who uses machine learning method. In this study 18 attributes (age attributes were not being considered in this study) were taken out of 19 attributes from UCI repository machine learning repository.

Table 1: Attributes of Hepatitis Dataset in UCI Repository

S.No.	Variable	Variable description	Value
1	Age	Ranges from 20 to 78 years old	Having discrete values
2	Sex	Male or female, and is represented by 1 or 2 respectively	Male, female
3	Steroid	Value 1 for no & value 2 for yes	No, yes
4	Antivirals	Value 1 for no & value 2 for yes	No, yes
5	Fatigue	Value 1 for no & value 2 for yes	No, yes
6	Malaise	Value 1 for no & value 2 for yes	No, yes
7	Anorexia	Value 1 for no & value 2 for yes	No, yes
8	Liver big	Value 1 for no & value 2 for yes	No, yes

9	Liver firm	Value 1 for no & value 2 for yes	No, yes
10	Spleen palpable	Value 1 for no & value 2 for yes	No, yes
11	Spiders	Red capillary tufts in the skin that blanch on pressure, and is represented by 1 or 2 respectively	No, yes
12	Ascites	Accumulation of fluid in the abdominal cavity, and is represented by 1 or 2 respectively	No, yes
13	Varices	Dilated veins, and is represented by 1 or 2	No, yes
14	Bilirubin	A bile pigment cleared from the blood by the liver	0.3 to 7.6 having discrete values
15	Alkaline phosphatase	Protein found in bile duct cell membranes	30 to 295 having discrete values
16	SGOT	Aspartate transaminase(SGOT) enzyme that catalyze protein transform actions within hepatocytes	14 to 648 having discrete values
17	Albumin	A protein in the serum that transports substances such as drugs and prevents leakage of fluid into the surrounding tissues	2.4 to 6.4 having discrete values
18	Protine	The pro-thrombin time in serum	0 to 100 having discrete values
19	Histology	Value 1 for no & value 2 for yes	No, yes

B. Remove Instances with Missing Attribute and Normalization of Data

Binary values in dataset at the place of 2 were changed to 0 and 1 kept as same. In the original dataset many missing values were observed. All the missing values were suitably replaced with the values of 0 and 1. Then, the data were divided two parts – live class dataset and die class dataset. After this, all missing discrete value with the average in the live dataset and die dataset were placed separately. In the normalization process, all the records of the attribute of live dataset and die dataset were placed in the range between the 0 and 1.

C. Sorting, Data Splitting and Generation of Models

Sorting has been done randomly in the whole dataset. Then the whole dataset of 155 samples were divided into train and test dataset having 117 and 38 samples respectively in each. Simple regression method using a mixture of experts model with Expectation Maximization (EM) training were used further to train the data. The architecture used is a “variance” mixture of experts with generalized logistic models (perceptrons) for the experts and the gate. Training is done within the Expectation Maximization framework. Total six parameters (no. of expert, expert output activation function, type of ME, no. of iterations in the M step, EM step, optimization in the M step) were considered in the development of different models in this method. Out of which five parameters were kept constant and one was kept variable to define various models.

D. Mixture of Expert

Mixture of Expert is a supervised learning algorithm that divides a learning task into appropriate subtasks, each of which can be solved by simple expert network [14]. The global output of the ME system is derived as a convex combination of the outputs from a set of N experts, in which the overall predictive performance of the system is generally superior to that of any of the individual experts. The ME error function is based on the interpretation of MEs as a mixture model with conditional densities as mixture components (for the experts) and gating network outputs as mixing coefficients. ME architecture [Fig. 1] is composed of several expert networks and a gating network. Gating network produces a scalar output from the vector input x.

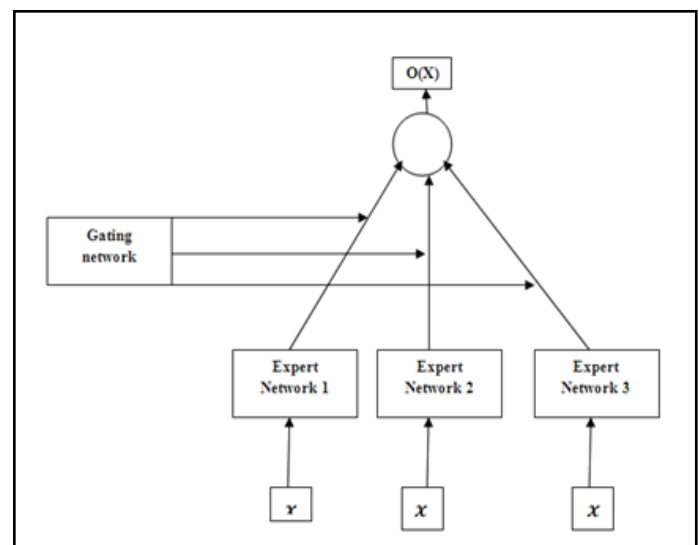


Fig. 1: Mixture of Expert Architecture

The gating network works on a generalized linear function in which the output for i^{th} input variable is given by

$$g(x, v_i) = e^{\xi_i} / \sum_{k=1}^N e^{\xi_k} = 1/e^{\xi_i} \quad (1)$$

Where, $\xi_i = v_i^T x$ and v_i is the weight vector. Each expert network produces an output vector for an input vector from the following generalized linear equation:

$$o_i(x) = f(W_i x) \quad (2)$$

Where, W_i is a weight matrix. The final output of ME is thus the sum of multiplication of the output from the gating and expert networks [15].

$$o(x) = \sum_{k=1}^N g(x, v_k) o_k(x) \quad (3)$$

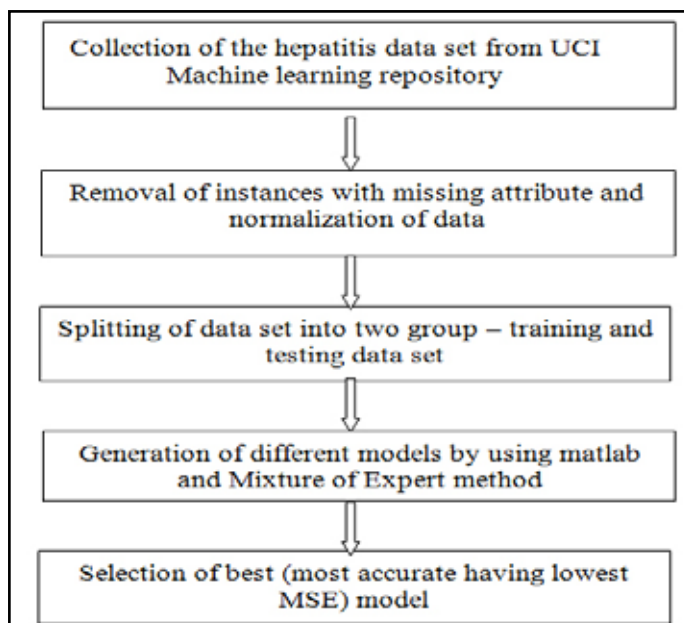


Fig. 2: Overall Block Diagram of the Methodology Used

III. Results

Various configurations were tested to attain the best optimum accuracy of the algorithm by the trial and error approach. The number of hidden nodes is joined among the mapping ability of the network. Large numbers of hidden nodes make the network more powerful. In contrast, if the number of hidden nodes is too large, then overtraining occurs [15].

As a result, generalization of the network may lead to erroneous result, which may further result in poor performance of the network on test data. While trying different parameters, best accuracy of 97.37% was achieved with EM steps of 300 by applying scaled conjugate gradient in optimization. Classification accuracy of 86.84% was obtained with 350 numbers of EM steps. 94.7% classification accuracy was observed by applying 8 no. of iteration in the M step with same EM steps. The prediction error during the training process approached almost zero.

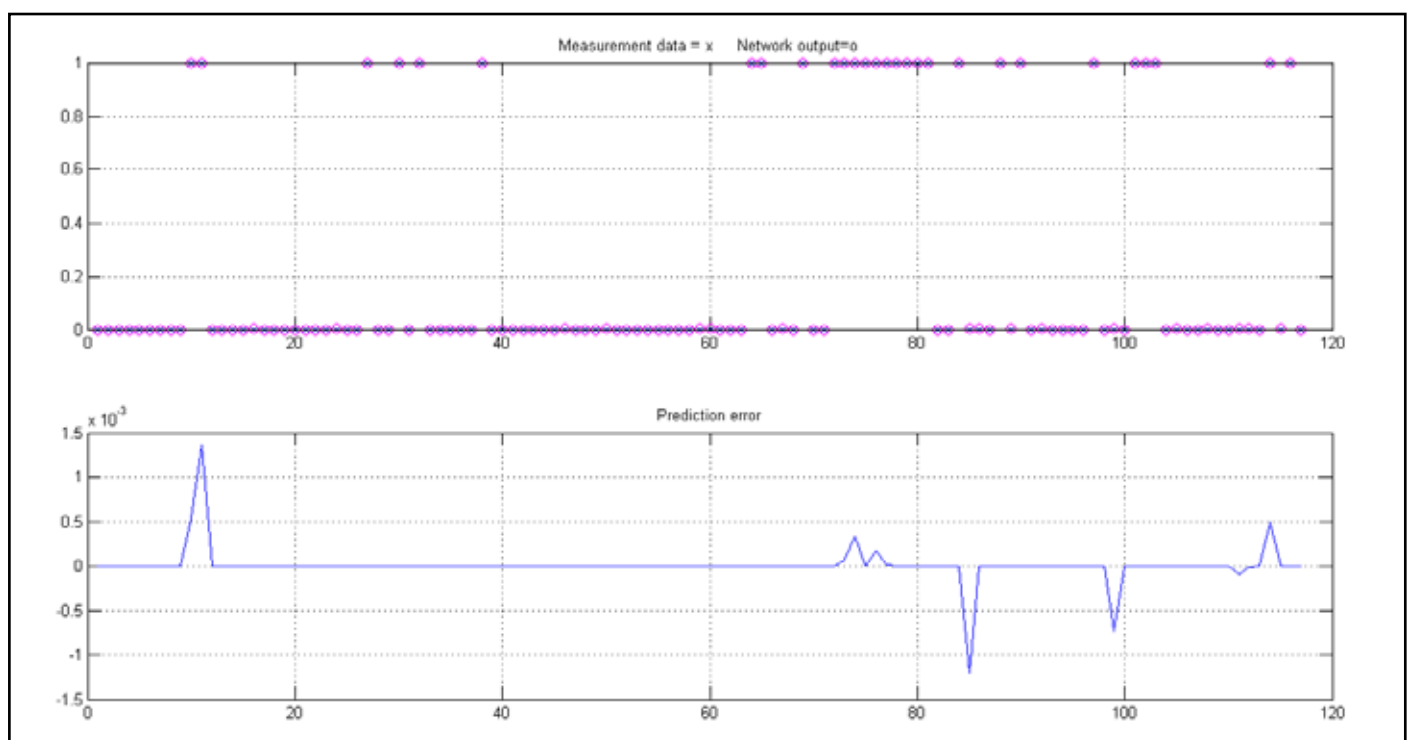


Fig. 3: Network Output of Training and Prediction Error in Training

Sensitivity, specificity, and total classification accuracy were used as measures to determine the performance of the ME classifier. These three terms can be defined as follow: sensitivity = (number of true-positive outputs/number of actually positive cases); specificity = (number of true-negative outputs/ number of actually negative cases); and total classification accuracy = (number of correct decisions/total number of cases). Thus sensitivity, specificity, and total classification accuracy were calculated as 100%, 75%, and 97.37%, respectively, from the confusion matrix [Table 2].

Table 2: Confusion Matrix

	Without Hepatitis	With Hepatitis
Without Hepatitis	34	0
With Hepatitis	1	3

VI. Discussion

Several Neural Network (NN) methods have already been tested in the hepatitis diagnosis. GRNN resulted in only 80% of accuracy whereas use of 1NN, 15 NN and 9NN methods, was increased the classification accuracy up to 85.3%, 89% and 90.3%. When artificial neural network combined with PCA (PCA-ANN), resulted in 89.6 % accuracy [16]. While neural network incorporated with MLPNN and LM, classification accuracy was more enhanced than the simple neural network and gives 91.87% classification accuracy. Interestingly, similar classification accuracy of 96.77%, was observed by Chen et al., (2011) [17] and Adeli et al., (2011) [18] respectively by using two different methods LFDA-SVM and Genetic algorithm with Radial Basic Network Exact Fit (GA-RBEF) which is based on pattern recognition. In the present study, method of Mixture of Expert give classification accuracy of

97.37%, which is perhaps highest than the previous methods using computational intelligence techniques. This study demonstrates remarkably high accuracy on the UCI which further confirms the outstanding performance of this unique algorithm.

V. Conclusion

Mixture of expert method can help the physician for the hepatitis diagnosis. The results show that Mixture of expert method can assist in the diagnosis of hepatitis diseases. In future studies on the hepatitis diagnosis diseases, different learning methods will be used for increasing the accuracy of the systems. In future studies on the hepatitis diagnosis diseases, this method can be further modified to get more accuracy. The accuracy may also be increased by increasing the number of subjects in the dataset. Diagnosis of types of hepatitis can be included in the present study for prediction of type of hepatitis like hepatitis A, hepatitis B and hepatitis C, etc.

VI. Acknowledgement

Sincere thanks to Dr. Ambarish S. Vidyarthi, Professor and Head, Department of Biotechnology, B.I.T Mesra. I would also like to express my sincere gratitude to Dr. Shankaracharya, Department of Biotechnology, Birla Institute of Technology, Mesra.

VII. Author Disclosure Statement

No competing financial interests exist.

References

- [1] Mougiakakou, G.S., Valavanis, K.I., Nikita, A., et al., "Diagnostic Support Systems and Computational Intelligence: Differential Diagnosis of Hepatic Lesions from Computed Tomography Images", IGI, 2009.
- [2] Sartakhti, S.J., Zangoeei, M. H., Mozafari, K., "Hepatitis disease diagnosis using a novel hybrid method based on support vector machine and simulated annealing (SVM-SA)", *Comput. Methods Programs Biomed.*, Vol. 10, pp. 3266-3277, 2011.
- [3] Shankaracharya, Kumari, S., Vidyarthi, S.A., "Development of java based graphical user interface for diagnosis of hepatitis using mixture of expert", *Nature proceeding*. DOI:10.1038/npre.2012.7093.1:Posted 2 Apr 2012.
- [4] Ozyilmaz, L., Yildirim, T., "Artificial neural networks for diagnosis of hepatitis disease", In: *Proceedings of the International Joint Conference on Neural Networks*, Vol. 1, pp. 586-589, 2003.
- [5] Dich, W., Grudzinski, K., Dierksen, F.H.G., "Minimal distance neural methods, In: *IEEE World Congress on Computational Intelligence*", the 1998 IEEE International Joint Conference on Neural Networks Proceedings, Vol. 2, pp. 1299-1304, 1998.
- [6] Ster, B., Dobnikar, A., "Neural Networks in Medical Diagnosis: Comparison with Other Methods", *proceeding of the international conference EANN96*, Vol. 1, pp. 427-430, 1996.
- [7] Duch, W.; Adamczak, R.; Grabczewski, K., "Optimization of logical rules derived by neural procedures", In: *IJCNN'99, International Joint Conference on Neural Networks*, Vol. 1, pp. 669-674, 1999.
- [8] Polat, K.; Günes, S., "Hepatitis disease diagnosis using a new hybrid system based on Feature Selection (FS) and artificial immune recognition system with fuzzy resource allocation", *DSP*, Vol. 16, pp. 889-901, 2006.
- [9] Polat, K.; Günes, S., "Medical decision support system based on Artificial Immune Recognition Immune System (AIRS), fuzzy weighted pre-processing and feature selection", *Expert Syst. Applicat.*, Vol. 33, pp. 484-490, 2007.
- [10] Dogantekin, E., Dogantekin, A.; Avci, D., "Automatic hepatitis diagnosis system based on Linear Discriminant Analysis and Adaptive Network based on Fuzzy Inference System", *Expert Syst. Applicat.*, Vol. 36, pp. 11282-11286, 2009.
- [11] Bascil S.M., Oztekin H., "A Study on Hepatitis Disease Diagnosis Using Probabilistic Neural Network", *Journal of Medical Systems*, 2010.
- [12] Çalışır, D.; Dogantekin, E., "A new intelligent hepatitis diagnosis system: PCA-LSSVM", *Expert Syst. Applicat.*, Vol. 38, pp. 10705-10708, 2011.
- [13] Tan, C.T.; Teoh, J.E.; Yu, Q.; et al., "A hybrid evolutionary algorithm for attribute selection in data mining", *Expert Syst. Applicat.*, Vol. 36, pp. 8616-8630, 2009.
- [14] Shankaracharya, Odedra, D., Samanta, S.; et al., "Computational Intelligence in Early Diabetes Diagnosis", *Rev Diabet Stud.*, Vol. 4, pp. 252-262, 2010.
- [15] Shankaracharya; Devang, O.; Medhavi, M.; et al., *Java-Based Diabetes Type 2 Prediction tool for Better Diagnosis*, *Diab tech ther.*, Vol. 14, pp. 251-256, 2012.
- [16] Jilani, A.T.; Yasin, H.; Yasin, M.M., "PCA-ANN for Classification of Hepatitis-C Patients", *IJCA*, Vol. 7, pp. 1-6, 2011.
- [17] Chen, L.H.; Liu, Y.D.; Youn, B.; et al., "A new hybrid method based on local fisher discriminant analysis and support vector machines for hepatitis disease diagnosis", *Expert Syst. Applcat.*, Vol. 38, pp. 11796-11803, 2011.
- [18] Adeli, M.; Zarabandipour, H., "Automatic disease diagnosis systems using pattern recognition based genetic algorithm and neural networks", *IJPS*, Vol. 6, pp. 6076-6081, 2011.



Suchitra Kumari received her B.Sc. degree in Biotechnology from Ranchi Women's College, Ranchi, Jharkhand, India in 2010, the M.Sc. degree in Bioinformatics from Birla Institute of Technology, Mesra, Ranchi, Jharkhand, India, in 2012. She won young scientist award on International conference in Microbial Plant and animal research 2012. Her research interests include Disease model study, drug designing, immunoinformatics. At present, she is working as Research Trainee at NTHRYS Biotech Labs, Hyderabad, India.