An Automatic Diagnosis System for Hepatitis Diseases Based on Extreme Learning Machine

Derya AVCI

Firat University, Engineering Faculty, Department of Electrical and Electronic, 23119, Elazig, TURKEY

Abstract— Hepatitis is a major public health problem all around the world. This paper proposes an automatic disease diagnosis system for hepatitis using pattern recognition based on Extreme Learning Machines (ELM). The classifier used in this paper is single layer neural network (SLNN) and it is trained by ELM learning method. The hepatitis disease datasets are obtained from UCI machine learning database. The performance of proposed method is evaluated through statical methods such as classification accuracy, sensitivity and specivity analysis. The results of the proposed method are compared with the results of the previous hepatitis disease studies using same database as well as different database. When previous studies are investigated, it is clearly seen that the high classification accuracies have been obtained in case of reducing the feature vector to low dimension. However, proposed method gives satisfactory results without reducing the feature vector. The calculated highest classification accuracy of proposed method is found as 91.50 %.

Keywords— Hepatitis; Diagnosis Systems; Extreme Learning Machines (ELM).

I. INTRODUCTION

In medicine, the inflammation occured in liver is called hepatitis. It can be caused by infections with viruses, bacteria, fungi, exposure to toxins such as alcohol and autoimmunity. The hepatitis damages to cells of liver and tenderness, swelling and inflammation in the liver are some symptoms of this disease. Commonly, 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. The hepatitis can be acute or chronic and it is a common disease over the world. The different types of hepatitis are caused by different things, but they all produce inflammation of the liver [1]. Viral hepatitis refers to several common contagious diseases caused by viruses that attack the liver. The most important types of viral hepatitis are hepatitis-A, hepatitis-B, and hepatitis-C [2]. New forms of viral hepatitis such as D, E and G are also discovered. Hepatitis-A mainly infects by fecal contaminated substances taken by mouth. As a result of poor hygienic conditions, this results in epidemics based on water or food, especially in developing countries [1]. In infection of hepatitis A, any indication may be observed. Therefore, the patients could not become aware of the hepatitis-A.

Recently, a new learning algorithm called Extreme Learning Machine (ELM) which randomly selected all the hidden nodes parameters of generalized Single-hidden Layer Feedforward Networks (SLFNs) and analytically determines the output weights of SLFNs is proposed in [15-21]. Altough output weights are analitically calculated, there is no rule in determination of number of hidden neurons and type of the activation function. To obtain a good classification performance of ELM, these parameters should be determined properly.

This study proposes ELM based an optimal intelligent system for diagnose of diabetes. The training and testing dataset for proposed method is obtained from UCI dataset. These dataset compose of 155 data. The randomly selected 100 of 155 data are used for training of classifier whereas remaining data is used for testing of classifier. For different activation function and number of hidden neurons, the results of proposed method are given. Further, a comparison is performed with previous studies to show validity of proposed method. From results, the proposed method is quite powerful tool for automatic diagnosis of hepatitis and may work in real-time systems.

II. METHODOLOGY

1. Pattern Recognition Concept

Pattern recognition could simply be defined as the process of categorizing the input signal. This process can be divided into two main groups as classification and regression. In the classification process, input signal is seperated into two or more classes while pattern recognition system tries to find the appropriate mapping between the input and the output variables in the regression process. Generally, regression problems are harder to solve than classification problems due to several reasons [22]. Figure 1 shows conventional pattern recognition concept for classification. As shown in Figure 1, the pattern recognition concept composes of two stages. They include feature extraction and classification stages. The feature extraction stage is the most important part of pattern diagnosis. The appropriate features are extracted by a feature extractor. If the appropriate features are not selected, the classification performance will be poor even though using the best classifier. Therefore, feature extractor should reduce the dimension of pattern vector to a lower. The reduced feature vector should also cover useful information of the original vector. In last stage, reduced feature vector is given to inputs of classifier for the classification [20-21].

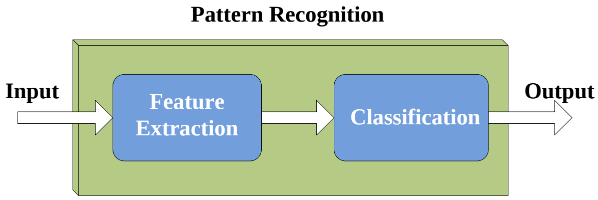

Figure 1. The block diagram of the pattern recognition concept

2. Extreme Learning Machine for Single Hidden Layer Feed Forward Networks

The NNs have been widely used in pattern recognition and regression problems. Commonly, the learning of NN has been performed by using gradient-based learning algorithms. However, such methods has several drawbacks such as difficult setting of learning parameters, slow convergence, slow learning and training failures [20-21].

To deal with the drawbacks of gradient-based learning methods, ELM was proposed by Huang, et al. [20]. In the ELM, the output weights of a single-hidden layer feedforward network (SLFN) are analytically computed by using the Moore-Penrose (MP) generalized inverse instead of iterative learning scheme. Figure 2 shows structure of a SLFN using ELM. In this figure, l1m, l2m and lrm, are weights vector connecting the ith hidden neuron and the input neurons, and w is the weight vector connecting the ith hidden neuron and output neuron and f(.) is activation function.

The most important properties of ELM are given as follows:

- The learning speed of ELM is extremely fast. Therefore, SLFNs can be trained by ELM much faster than classical learning methods.
- The ELM tends to reach both the smallest training error and the smallest norm of weights. Thus, the ELM tends to have good performance for neural networks.
- The ELM learning algorithm can be used to train SLFNs with non-differentiable activation functions.
- The ELM tends to reach the solutions straightforward without such trivial issues [20].

4. Data Description

The hepatitis data is obtained from UCI machine learning database. It contains 20 attributes including the class attribute [26]. The output shows whether patients with hepatitis are die or alive. Hepatitis dataset contains 155 data belonging to two different target classes. There are 20 features, 13 binary and 6 attributes with 6–8 discrete values. Also, the class distribution contains 32 cases for die and 123 cases for alive. The attribute information of dataset is given in Table I.

Table I. Attribute informations and statistical analysis of dataset

	Attribute information	**Values**
1	Age	10-80 by step 10 years
2	Sex	Male and female
3	Steroid	No, yes
4	Antiviral	No, yes
5	Fatigue	No, yes
6	Malaise	No, yes
7	Anorexia	No, yes
8	Liver big	No, yes
9	Liver firm	No, yes
10	Spleen palpable	No, yes
11	Spiders	No, yes
12	Ascides	No, yes
13	Varices	No, yes
14	Bilirubin	0.39, 0.80, 1.20, 2.00, 3.00, 4.00
15	Alk phosphate	33, 80, 120, 160, 200, 250
16	Sgot	13, 100, 200, 300, 400, 500,
17	Albumin	2.1, 3.0, 3.8, 4.5, 5.0, 6.0
18	Protime	10, 20, 30, 40, 50, 60, 70, 80, 90
19	Histology	No, yes
20	Class	Die, alive

5. Application of ELM for Diagnosis of Hepatitis

In these applications, a 3-fold cross-validation schema was applied where the two-fifth data were used for training the proposed ELM method and the remaining other data were used as the test data set. We applied this strategy for three times and we calculated the average values for determining the performance of proposed ELM method.

In the test process of ELM, the optimum parameters for the number of hidden layer neurons and type of activation function were used in ELM structure. As said previously, the dataset has 19 relevant features except for class attribute and includes a total of 155 cases. Thus, it is a matrix with dimension of 155 x 19. The type of activation function and the number of hidden neurons for ELM classifier are determined by doing many test in this study. The activation functions covers sigmoid, sinus, tangent sigmoid, sigmoid, radial basis, triangular, poly, hardlim. In addition, the numbers of hidden neurons are determined a value between 5 and 132. Performance of the proposed method is calculated by three evaluation methods as classification accuracy, sensitivity and specificity analysis. The classification accuracies for the datasets are found using the following equation.

$$Correct\ accuracy\ \ (C)=\frac{\sum_{k=1}^{|C|} assess(c_k)}{|C|},\ \ c_k\in C$$

$$assess(C)=\begin{cases}1 & if\ classify\,(c)=c.m\\ 0 & otherwise\end{cases} \tag{1}$$

$$assess(c)=\begin{cases}1, & if \quad classify(c)=c.d\\ 0, & otherwise\end{cases} \tag{2}$$

Where, C is the set of hepatitis data to be classified (the test set), c.m is the class of item c and classify(c) returns the classification of the by ELM. The sensivity and specificity analysis are obtained as follows:

$$sensivity(\%) = \frac{TP}{TP + FN}$$
$$specificity(\%) = \frac{TN}{FN + TN} \qquad (3)$$

Where; TP, TN, FP and FN are true positives, true negatives, false positives and false negatives, respectively.

6. Results and Discussion

In this study, an automatic intelligent system for diagnosis of hepatitis disease is presented. Performance of the proposed method is also evaluated by classification accuracy, sensitivity and specificity analysis. For different type of activation function and the number of hidden neurons, the classification accuracies of the proposed method are given in Table II. As shown in this table, the best classification accuracy of the proposed method is found as 91.50 % in case of activation function with tangent sigmoid and hidden neuron with 17. This table presents only the optimum results of proposed method. In addition, the worst classification accuracy of this method is also obtained as 61.7524% for radial basis activation function and hidden neuron with 240.

Table II. Some optimum results obtained from testing of the ELM for hepatitis diagnosis

Type of the activation function	The number of hidden neurons	Accuracy (%)
Sigmoid	87	89.0909
Poly	57	89.0909
Tangent sigmoid	***259***	***90.7273***
Sigmoid	143	90.9091
Poly	191	89.0909
Tangent sigmoid	***17***	***91.50***
Sigmoid	81	89.0909
Tangent sigmoid	204	92.7273

References

[1] http://www.hepcprimer.com/what.html

[2] Lale Ozyilmaz, Tulay Yildirim, Artificial Neural Networks for Diagnosis of Hepatitis Disease, International Joint Conference on Neural Networks, July 20-24, Portland, Oregon, USA, 2003.

[3] Tahseen A. Jilani, Huda Yasin, Madiha Mohammad Yasin, PCA-ANN for Classification of Hepatitis-C Patients, International Journal of Computer Applications (0975 – 8887), Volume 14– No.7, February 2011.

[4] Ultrasound and Magnetic Resonance Medical Images Textures Analysis Using Neural Networks, http://www.intermag.kiev.ua/projects/neuraln.html, (last accessed: 20 January 2012).

[5] M. Serdar Bascil, Feyzullah Temurtas, A Study on Hepatitis Disease Diagnosis Using Multilayer Neural Network with Levenberg Marquardt Training Algorithm, J Med Syst (2011) 35:433 – 436.

[6] Mahdieh Adeli, Hassan Zarabadipour, Automatic disease diagnosis systems using pattern recognition based genetic algorithm and neural networks, Recent Advances in Applied & Biomedical Informatics and Computational Engineering in Systems Applications,

[7] Esin Dogantekin, Akif Dogantekin, Derya Avci, Automatic hepatitis diagnosis system based on Linear Discriminant Analysis and Adaptive Network based on Fuzzy Inference System, Expert Systems with Applications 36 (2009) 11282–11286.

[8] Peng Guan, De-Sheng Huang, Bao-Sen Zhou, Forecasting model for the incidence of hepatitis A based on artificial neural network, China World Journal of Gastroenterol; 10(24), 2004, pp. 3579-3582.

[9] Polat K., Gunes S., Hepatitis disease diagnosis using a new hybrid system based on feature selection (FS) and artificial immune recognition system with fuzzy resource allocation, Digital Signal Processing 16 (2006), pp. 889–901.

[10] Duygu Çalisir, Esin Dogantekin, A new intelligent hepatitis diagnosis system: PCA–LSSVM, Expert Systems with Applications, Volume 38, Issue 8, pages: 10705-10708, 2011.

[11] D. Rumelhart, and J. McClelland, "Parallel Distributed Processing", Vol. 1, MIT Press, 1986.

[12] Suresh, S., Omkar, S.N., and Mani, V., Parallel Implementation of Back-propagation Algorithm in Networks of Workstations, IEEE Transactions on Parallel and Distributed Systems, 16, pp. 24-34, 2005.

[13] Hsu, C.T. Kang, M.S. and Chen, C.S, Design of Adaptive Load Shedding by Artificial Neural Networks, IEE Proceed., Generation, Transmission and Distribution, 152, pp. 415-421, 2005.

[14] Cho, J. H., Chun, M. G., Lee, D. J., Parameter Optimization of Extreme Learning Machine Using Bacterial Foraging Algorithm, EESRI, pp: 742-747, 2007.

[15] G. B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme Learning Machine for Regression and Multiclass Classification," (in press) IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics, 2011.

[16] G. B. Huang, D. H. Wang, and Y. Lan, "Extreme Learning Machines: A Survey," International Journal of Machine Leaning and Cybernetics, 2 (2011) 107-1221.

[17] G. B. Huang, X. Ding, and H. Zhou, "Optimization Method Based Extreme Learning Machine for Classification", Neurocomputing, 74 (2010) 155-163.

[18] G. B. Huang, L. Chen and C. K. Siew, "Universal Approximation Using Incremental Constructive Feedforward Networks with Random Hidden Nodes", IEEE Transactions on Neural Networks, 17 (2006) 879-892.

[19] G. B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme Learning Machine: A New Learning Scheme of Feedforward Neural Networks," 2004 International Joint Conference on Neural Networks (IJCNN'2004), (Budapest, Hungary), July 25-29, 2004.
[20] G. B. Huang, Q.-Y. Zhu and C.-K. Siew, "Extreme Learning Machine: Theory and Applications", Neurocomputing, vol. 70, pp. 489-501, 2006.
[21] N.-Y. Liang, G.-B. Huang, P. Saratchandran, and N. Sundararajan, "A Fast and Accurate On-line Sequential Learning Algorithm for Feedforward Networks", IEEE Transactions on Neural Networks, 17 (2006) 1411-1423.
[22] L. Guo, D. Rivero, J. Dorado, C. R. Munteanu, A. Pazos, Automatic feature extraction using genetic programming: An application to epileptic EEG classification Expert Systems with Applications, Volume 38, Issue 8, August 2011, Pages 10425-10436.
[23] B. Biswal, P.K. Dash, B.K. Panigrahi, Non-stationary power signal processing for pattern recognition using HS-transform, Applied Soft Computing, Volume 9, Issue 1, January 2009, Pages 107-117.
[24] L. Jiang, Y. Liu, X. Li, S. Tang, Using bispectral distribution as a feature for rotating machinery fault diagnosis Measurement, Volume 44, Issue 7, August 2011, Pages 1284-1292.
[25] Zhang, L. and Zhou, W. D., Density-induced margin support vector machines Pattern Recognition, 44, pp. 1448-1460, 2011.
[26] UCI Machine Learning Repository. Available from: http://www.ics.uci.edu/~mlearn/MLRepository.html (last accessed: 20 January 2006).

Derya Avci – She was born in Malatya Turkey. She received a master's degree from Firat University Electronic and Computer Education. She received Ph.D. degree from Firat University Electrical and Electronic Engineering. Her interesting areas are pattern recognition, image processing, intelligent systems, artificial intelligence systems.