SOUTHEAST UNIVERSITY - MONASH UNIVERSITY

JOINTGRADUATE SCHOOL (SUZHOU)

An Application of the Neural Network (NN) Model on the Diagnosis of Hepatitis

A Literature Review on the Hepatitis Diagnosis Using Neural Networks (NNs)

Student name: Huang Hanling ID: 2731****

Student name: Yu Yunxiu ID: 2731****

13 May 2016

Abstract

Disease diagnosis has become an important application aspect of neural networks in medical area. This paper is about to review related studies in diagnosing hepatitis. These studies cover neural network models with different architectures, learning algorithms and activation functions, including the Multilayer Perceptron (MLP) network with the backpropagtion (BP) algorithm, the Radial Basis Function (RBF) network, the Generalized Regression Neural Network (GRNN), the multilayer neural network (MLNN) with the Levenberg Morquardt (LM) algorithm and the multilayer neural network (MLNN) with the approximation sigmoid activation function. All studies mentioned in this review use identical hepatitis data set from the UCI machine learning database. We will analyze each models' features and their accuracy in diagnosing hepatitis. And then we will compare different neural network models' performances.

Key Words: *Neural Network, Hepatitis diagnosis*

1. Introduction

Hepatitis disease is an inflammation and can damage patients' liver cell. As a major public health problem, hepatitis may be acute with recovery within 6 months or be chronic which may lead to death if the patients breathe with difficulties (Ozyilmaz & Yildirim, 2003). As a result, the diagnosis of hepatitis that whether the patients will survive or die has a significant meaning in medical area. The traditional medical diagnosis is usually done by expert doctors and it is quite difficult. Recently, the automatic diagnosis using neural networks solves this limitation of the traditional methods of diagnosing hepatitis. The neural network is a type of Artificial Intelligence (AI) that is widely used in disease diagnosis. It consists of processing units called neurons and its working mechanism is similar to the human brain (Ansari et al., 2011).

In order to build an accurate NN model to do the hepatitis diagnosis, lots of researchers use different algorithms to design different neural networks for hepatitis diagnosis. To have a better understanding of the NN technologies of hepatitis diagnosis and to find a better NN model to diagnose hepatitis, we have read more than ten papers in this field and finally choose several papers which are most relevant to our research. In these studies, they all choose the hepatitis data set of the UCI machine learning database. This data set consist of 155 samples of hepatitis patients. 32 samples of them belongs to class 1-die while 123 samples of them belongs to class 2-live. All samples have 19 features which are age, sex, steroid, antivirals, fatigue, malaise, anorexia, liver big, liver firm, spleen palpable, spiders, ascites, varices, bilirubin, alk phosphate, sgot, albumin, protime and histology. These studies use these

19 features as the inputs of their NN model and classify the results into two classes: die and live which means the type of hepatitis is life-threatening or non-life-threatening. By reading these relevant papers, we find that current studies have built NN models which use the physical and medical features of data set to classify the results and have realized the hepatitis diagnosis. These studies have shown that NN is a feasible method in the hepatitis diagnosis (Ozyilmaz & Yildirim, 2003). However, these studies have not shown the correlations among these features clearly. Meanwhile, the accuracy of the NN model would not be guaranteed.

This review will make critical comments on different NN models in these papers, describe their methods and evaluate their performance based on their accuracy and efficiency. The objective of this review is to extract the meaningful features of these NN models that will be helpful to build our own NN model and to make improvements to the NN model building by studying the correlations of features of inputs.

2. Scope and method

In this review, we focus on the neural network models in diagnosing hepatitis. According to papers we read, first, we will classify these models into four sections and describe models' features in each section. The four sections of these models are: (a) Two standard feedforward network: the Multilayer Perceptron (MLP) structure trained by the backpropagation (BP) algorithm and the Radial Basis Function (RBF) network structure with Ordinary Least Square (OLS) algorithm; (b) A variant of the RBF model: the Generalized Regression Neural Network (GRNN) model; (c) The Multilayer Neural Network (MLNN) structure model with the Levenberg Morquardt (LM) learning algorithm; (d) The Multilayer Neural Network(MLNN) model with the approximations of sigmoid activation function. Then, we are going to give their accuracy in hepatitis diagnosis and compare their performances. Finally, conclusion and expectation of using NN models in diagnosing hepatitis will be given.

To attest the quality of the research papers we find for this review, we use Monash University Library search engine for its accessibility to the most important publications from journals and conference proceedings.

3. Body of the review

3.1 NN models' features description

(a) The MLP model with the BP algorithm and the RBF model with the OLS algorithm

Ozyilmaz and Yildirim (2003) conduct study of the MLP structure trained by the standard BP algorithm and the RBF network structure trained by the OLS algorithm. In this MLP network, the outputs of the units in one layer form the inputs to the next layer and the weight of network calculated through the training of network by the BP algorithm. The RBF model they use contains one hidden layer of neurons. In this network, the distance between the neuron center and the input vector is calculated at the input of each neuron. And the output of neuron is formed through the application of the Guassin bell basis function. The output of this network is gotten by a weighted sum of the neuron outputs and the unity bias. The network is trained by the OSL algorithm. 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. They choose 80 samples without missing attribute to form the data set and apply 5-fold cross-validation method in their study to avoid network memorizing training set so to guarantee the good generalization of the neural network. According to their study result, the MLP model reaches 81.375% average accuracy and the RBF model reaches 85% accuracy.

(b) The GRNN model

Ansari et al. (2011) use the GRNN model to diagnose hepatitis disease. The GRNN is a variant of the RBF network which is used by Ozyilmaz and Yildirim (2003) for the same purpose and both of them are based on the one pass learning algorithm. The GRNN in their study consists of 4 layers: input layer, pattern layer, summation layer, and output layer. In their study, the input layer has 19 neurons and one neuron at output layer. The pattern layer use the radial basis transfer function and the summation layer use the linear transfer function. Besides, the Euclidean distance weight function is used to calculate the weighted inputs. In the study of Ansari et al. (2011), they use 100 samples of which 87 samples belong to live class and 13 belong to die class to train the network and 55 samples of which 36 samples belong to live class and 19 belong to die class as the test set. According to their study result, the GRNN model reaches the diagnosis accuracy of 92%.

(c) The MLNN model with the LM learning algorithm

Bascil and Temurtas (2009) develop the MLNN for diagnosing hepatitis which use the LM algorithm as the training algorithm to update the weight of the network. The MLNN model they build consists of 2 hidden layers, the first hidden layer contains 60 neurons and the second hidden layer contains 20 neurons. The output layer of this model uses nonlinear sigmoid activation functions and has two neurons to indicate the two classes die and live respectively. They use 10-fold cross-validation to improve the precision of the model's classification result. According to their study result, they get 91.87% classification accuracy.

(d) The MLNN model with the approximations of sigmoid activation function

Çetin et al. (2015) make an improvement on the NN model in the study of Bascil and Temurtas (2009) which use the LM algorithm for learning. In their study, they use the approximations of sigmoid activation function instead of the most frequent used activation function:

$$y = \frac{1}{1+e^{-x}}$$

where y represents the output of neurons and x represents the number of neurons .After using the approximation of sigmoid activation function, it improves the calculation speed of the activation function and reduces the size of the hardware. So, to some extent, it will increasingly improve the application of the MLNN model in diagnosing hepatitis. The first approximation sigmoid function is a simple polynomial function:

$$\mathrm{y} = \frac{1}{2}\left(\frac{x}{1+|x|}+1\right)$$

And the second approximation sigmoid function use the piecewise linear technique to approximate the sigmoid function, the piecewise function is shown below:

$$\begin{array}{ll} \mathrm{Y} = 1 & |\mathrm{X}| \geq 5 \\ \mathrm{Y} = 0.03125 \cdot |\mathrm{X}| & 2.375 \leq |\mathrm{X}| < 5 \\ \mathrm{Y} = 0.125 \cdot |\mathrm{X}| + 0.625 & 1 \leq |\mathrm{X}| < 2.375 \\ \mathrm{Y} = 0.25 \cdot |\mathrm{X}| + 0.5 & 0 \leq |\mathrm{X}| < 1 \\ \mathrm{Y} = 1 - \mathrm{Y} & |\mathrm{X}| < 0 \end{array}$$

The third approximation function is formed through the Taylor series expansion, and the formula is shown below:

$$y = \begin{cases} 0.571859 + (0.392773)x + (0.108706)x^2 + (0.014222)x^3 + (0.000734)x^4 & -\infty < x \leq -1.5 \\ \frac{1}{2} + \frac{1}{4}x - \frac{1}{48}x^3 + \frac{1}{480}x^5 & -1.5 < x < 1.5 \\ 0.428141 + (0.392773)x + (0.108706)x^2 + (0.014222)x^3 - (0.00734)x^4 & 1.5 \leq x < \infty \end{cases}$$

The MLNN model they build contains 2 hidden layers and has 30 and 15 neurons respectively and use 10-fold cross-validation method to test the accuracy of it. According to their study result, compared with NN models using the sigmoid function which reach 93% classification accuracy, the first approximation sigmoid function have 91.8% accuracy, the piecewise linear approximation function reaches 92.5% accuracy and the Taylor series expansion approximation function reaches 93.1% accuracy. Hence, the Taylor series expansion approximation sigmoid function have the best approximation to sigmoid function.

3.2 Performances Analysis

When comparing the classification accuracy of NN models in different studies, the effectiveness of each model in diagnosing hepatitis is obvious. It can be found that the MPL model with the standard BP algorithm has the least accuracy compared with other NN models. Although this NN model is widely used in classification, it applies the steepest descent method to update the weights and it suffers from a slow convergence rate and often yields suboptimal solutions (Gori & Tesi, 1992) which influences the accuracy of the classification. As a result, this NN model is not suitable to be used in this hepatitis diagnosis application field. Compared with the MLP with the BP algorithm, the RBF network model with the OLS algorithm gets a better accuracy result under the same training and test data set condition. As a variant of the RBF network, the GRNN performs better in the hepatitis diagnosis and since both of them are based on one pass algorithm, the training of the network is much faster than the MLP with the BP algorithm (Heddam et al., 2011). The LM algorithm used in the study of Bascil and Temurtas (2009) and Çetin et al. (2015) provides a high accuracy in diagnosing hepatitis for it provides faster convergence and better estimation results than the BP algorithm (Bascil & Temurtas, 2009). Using of the LM as training algorithm of NN would get better classification results than other models using different training algorithms. The idea of using the approximation sigmoid activation function to replace the sigmoid activation function in the study of Çetin et al. (2015) has very important realistic meaning that ensures the comparatively high classification accuracy and faster calculation speed of the activation function at the same time.

4. Interpretation and conclusion

By reading these papers and evaluating the performances of these NN models, we could conclude that NN is an enhanced artificial intelligent method in diagnosing hepatitis with enormous potential. It solves two serious problems of the traditional medical diagnosis methods. One is that it is hard to implement diagnosis. The other is that the visual task is mostly done by some expert doctors (Ozyilmaz & Yildirim, 2003). Using NN models, it becomes easy to diagnose hepatitis because that diagnosis result can be gotten as long as the input data is given. Furthermore, it is easy to read the result by people no matter they are expert or not, because the output is clearly defined as die or live.

According to these papers, several NN models with different algorithms can be used in the area of hepatitis diagnosis. However, the efficiency and accuracy of these NN models are different. The MPL model with the standard BP algorithm is not that suitable for the hepatitis diagnosis because of its comparatively low accuracy and implementing efficiency. Comparing with the MPL model with the BP algorithm, the RBF network model with the OLS algorithm gets a better performance in accuracy. In addition, the GRNN as a variant of the RBF network gets faster implementing speed. So, it is more suitable to be used in the hepatitis diagnosis area. By comparing the training algorithms used in these NN models, we find that the LM algorithm has higher accuracy than the BP algorithm. After changing the sigmoid activation function of the LM algorithm to the approximation sigmoid activation function, both high accuracy and high efficiency can be realized.

NN is a practical technology that deserves deeper study and wider development. The contributions of this study area are becoming more and more obvious. However, there still exists improving space in this area because current studies have not shown the influential features that affect the diagnosis result and the correlations among these features clearly. That is the direction we are going to do our research. In conclusion, these papers are helpful to our research not only because they demonstrate good NN models in hepatitis diagnosis, but also because it inspires us that we could make some improvements in this field by finding the influential features that affect the diagnosis result and correlations among these features.

References

Ansari, S., Shafi, I., Ansari, A., Ahmad, J., & Shah, S. I. (2011). Diagnosis of liver disease induced by hepatitis virus using artificial neural networks. Proceedings of the 2011 IEEE 14th International Multitopic Conference (INMIC), 8-12.

Bascil, M. S., & Temurtas, F. (2011). A study on hepatitis disease diagnosis using multilayer neural network with Levenberg Marquardt training algorithm. Journal of Medical Systems, 35(3), 433-436.

Çetin, O., Temurtaş, F., & Gülgönül, Ş. (2015). An application of multilayer neural network on hepatitis disease diagnosis using approximations of sigmoid activation function. Dicle Tıp Dergisi, 42(2), 150-157.

Gori, M., & Tesi, A. (1992). On the problem of local minima in backpropagation. IEEE Transactions on Pattern Analysis & Machine Intelligence, (1), 76-86.

Heddam, S., Bermad, A., & Dechemi, N. (2011). Applications of radial-basis function and generalized regression neural networks for modeling of coagulant dosage in a drinking water-treatment plant: Comparative study. Journal of Environmental Engineering, 137(12), 1209-1214.

Ozyilmaz, L., & Yildirim, T. (2003). Artificial neural networks for diagnosis of hepatitis disease. Proceedings of the 2003 IEEE International Joint Conference on Neural Networks, 1, 586-589.