FIT5186 Intelligent Systems

SOUTHEAST UNIVERSITY - MONASH UNIVERSITY

JOINTGRADUATE SCHOOL (SUZHOU)

Diagnosis of Hepatitis Using Neural Networks

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

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

27 May 2016

Abstract

Hepatitis is one of the major public health problems worldwide. With the development of Artificial Neural Networks (ANNs), this technology has been widely used in the medical area. Therefore, many researchers have dedicated to build the hepatitis diagnosis neural network model using different architecture, training algorithm and activation function. However, they have not paid much attention to the input feature's influence on the diagnosis result. In this paper, we will use the backpropagation (BP) algorithm to train the neural network model. By designing and modifying its architecture, we will get a high accuracy hepatitis diagnosis model. Then, we are going to change the number of input features to analyse the influence on the diagnosis result and find the correlations between input features.

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 and time consuming. 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). A lot of researchers use different algorithms to design different neural networks for 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.

In this paper, we will first build a high accuracy NN model where the backpropagation (BP) algorithm is applied in the training of this multilayered FeedForward Neural Network (MFNN). To do the experiments, we use software NueroShell 2 to process the data set for it can provide flexibility in changing the architecture of the network and parameters. By change the number of input features, we try to find the features that has influence in the diagnosis result and their correlations. This paper will introduce several different experiments which have been conducted by different architectures and parameters.

2. Data Sets

We are going to use hepatitis data set from UCI machine learning database. It has been used by many researchers as data set to train and test their model. So this data set is reliable. And it provides data of 19 physical and medical features of 155 sampled hepatitis patients which 32 samples of them belongs to class 1-die while 123 samples of them belongs to class 2-live. The web link of the data set is: http://archive.ics.uci.edu/ml/datasets/Hepatitis.

Features information:

1. Class: DIE, LIVE
2. AGE: 10, 20, 30, 40, 50, 60, 70, 80
3. SEX: male, female
4. STEROID: no, yes
5. ANTIVIRALS: no, yes
6. FATIGUE: no, yes
7. MALAISE: no, yes
8. ANOREXIA: no, yes
9. LIVER BIG: no, yes
10. LIVER FIRM: no, yes
11. SPLEEN PALPABLE: no, yes
12. SPIDERS: no, yes
13. ASCITES: no, yes
14. VARICES: no, yes
15. BILIRUBIN: 0.39, 0.80, 1.20, 2.00, 3.00, 4.00
16. ALK PHOSPHATE: 33, 80, 120, 160, 200, 250
17. SGOT: 13, 100, 200, 300, 400, 500
18. ALBUMIN: 2.1, 3.0, 3.8, 4.5, 5.0, 6.0
19. PROTIME: 10, 20, 30, 40, 50, 60, 70, 80 and 90
20. HISTOLOGY: no, yes

In the data set, all features are numerical represented, we do not need to preprocess any attribute. When do the experiment, we use 80 samples of the data set which have no missing information and 13 of these samples belongs to first class and 67 belongs to the second class.

3. Training Issues

We did 5 rounds of experiments which including 30 experiments. In order to form the test set, 20 % of the data set has been randomly extracted. Round 1 includes experiments 1-3; Round 2 includes experiments 4-5; Round 3 includes experiments 6-24; Round 4 includes experiments 25; Round 5 includes experiments 26-30. The goal of Round 1 is to change parameters to build a high accuracy model. In this round, all of the 19 features chosen from the data set are used as inputs and 1output (i.e., Class die is represented as 1, Class live as 2). In experiment 1, we use the default parameters of the neuro shell which the hidden layer is 18, the learning rate and momentum both are 0.1. In experiment 2, we change the hidden layer to 30, and keep the learning rate and momentum as 0.1. In experiment 3, the hidden layer is 30, the learning rate is 0.9 and momentum is 0.6.

In experiments in Rounds 2-5, the hidden layer, the learning rate and the momentum are fixed to 30, 0.9 and 0.6 respectively. The goal of Round 2 is to decide the number of output neuros of the model. In this round, all of the 19 features are used as inputs. Experiment 4 has 1output neuron (i.e. Class die is represented as 1, Class live as 2). Experiment 5 has 2 output neurons (i.e. Class die is represented as 0 1, Class live as 1 0).

In experiments in Rounds 3-5, the output neuron is fixed to 1 (i.e., Class die is represented as 1, Class live as 2). The goal of Round 3 is to decide if there is a key feature which affect the result most. In this round, experiments 6-24 use 18 features as inputs which orderly remove one feature separately. The goal of Round 4 is to prove if there is disturbance features. In experiment 25, 18 features are used as the inputs and the hidden layer is changed to 18. The goal of Round 5 is to verify if there is correlation among some features. In experiment 26, only feature 8 of the data set is used as the input. In experiment 27, features 8 and 9 of the data set are used as the inputs. In experiment 28, features 8, 9 and 16 of the data set are used as the inputs. In experiment 29, features 8, 9, 16 and 19 of the data set are used as the inputs. In experiment 30, features 8, 9, 16, 19 and 18 of the data set are used as the inputs.

4. Results

Round 1:

Experiments 1-3 have been done to build a high accuracy model. In this round, the parameters are changes three times to build three different models. The results are shown in Tables 1 to 3.

Table 1 Classification accuracy of Experiment 1

Model 1: hidden layer : 18 learning rate: 0.1， momentum: 0.1

	Classified Class 1	Classified Class 2	Row Accuracy
Actual Class 1	3	1	75%
Actual Class 2	1	11	91.67%
Column Accuracy	75%	91.67%	

Table 2 Classification accuracy of Experiment 2

Model 2: hidden layer : 30 learning rate: 0.1， momentum: 0.1

	Classified Class 1	Classified Class 2	Row Accuracy
Actual Class 1	2	2	50%
Actual Class 2	0	12	100%
Column Accuracy	100%	85.71%	

Table 3 Classification accuracy of Experiment 3

Model3: hidden layer : 30 learning rate: 0.9， momentum: 0.6

	Classified Class 1	Classified Class 2	Row Accuracy
Actual Class 1	2	2	50%
Actual Class 2	0	12	100%
Column Accuracy	100%	85.71%	

Conclusion:

After comparing the accuracy of the experiments in this round, we find model 1 has the low accuracy. Although Models 2 and 3 has same accuracy, we choose model 3 as our model to do the rest of our experiments because models using bigger learning rate and momentum have better performance when the test scale is large.

Round 2:

Experiments 4-5 have been done to decide the numbers of the output neurons. In this round, the hidden layer is 30, the learning rate is 0.9 and the momentum is 0.6. The difference between experiment 4 and 5 is that they use 1 and 2 outputs neurons respectively. The results are shown in Tables 4 and 5.

Table 4 Classification accuracy of Experiment 4 with 1 output neuron

	Classified Class 1	Classified Class 2	Row Accuracy
Actual Class 1	2	2	50%
Actual Class 2	0	12	100%
Column Accuracy	100%	85.71%	

Table 5 Classification accuracy of Experiment 5 with 2 output neurons

	Classified Class 1	Classified Class 2	Row Accuracy
Actual Class 1	2	2	50%
Actual Class 2	0	12	100%
Column Accuracy	100%	85.71%	

Conclusion:

After comparing the accuracy, we decide to use 1 output neuron and use Model 3 to do the rest experiments of Rounds 3-5.

Round 3:

Experiments 6-24 have been done using Model 3 to check if there is a single key feature that can significantly affect the result. 18 features are used as inputs which orderly remove one feature from the data set separately. The accuracy of experiments 6-15, 17-14 are the same and the experiment 16 has different accuracy. The results are shown in Tables 6 and 7.

Table 6 Classification accuracy of Experiments 6-15 and 17-24

	Classified Class 1	Classified Class 2	Row Accuracy
Actual Class 1	2	2	50%
Actual Class 2	0	12	100%
Column Accuracy	100%	85.71%	

Table 7 Classification accuracy of Experiment 16

	Classified Class 1	Classified Class 2	Row Accuracy
Actual Class 1	3	1	75%
Actual Class 2	0	12	100%
Column Accuracy	100%	92.31%	

Conclusion:

From the results in this round, we find that no matter which feature we remove to do the experiment, the accuracy do not decline. So we conclude that there is no key feature that can significantly affect the result individually. In addition, from the experiment 16 which removes the feature 12, we see that the accuracy increases rather than decreases. So, we deduct that the feature 12 is a disturbance feature. If it can be proved, this feature should be wiped out from the data set.

Round 4:

Experiment 25 has been done in this round to check if feature 12 is a disturbance feature. We change the layer to 18 to see if it still increases the accuracy. The result is shown in Table 8.

Table 8 Classification accuracy of Experiment 25

	Classified Class 1	Classified Class 2	Row Accuracy
Actual Class 1	2	2	50%
Actual Class 2	0	12	100%
Column Accuracy	100%	85.71%	

Conclusion:

From this experiment we find that once the hidden layer is changed, the accuracy back to the same with other experiments in this round. So we conclude that the deduction of that feature 12 is a disturbance feature is not true.

Round 5:

Because the round 4 shows that there is not a key feature that affect the result individually, we want to test whether there are some joint features that can affect the result significantly. Experiments 26-30 use to verify if there is correlation among some features. The neuroshell shows each feature's contribution to the result, we do experiments with some top features to test the accuracy. The top 5 are features 8, 9, 16, 19 and 18. In experiment 26, only feature 8 of the data set is used as the input. In experiment 27, feature 8 and 9 of the data set are used as the inputs. In experiment 28, feature 8, 9 and 16 of the data set are used as the inputs. In experiment 29, feature 8, 9, 16 and 19 of the data set are used as the inputs. In experiment 30, features 8, 9, 16, 19 and 18 of the data set are used as the inputs. The result is shown in Tables 9 to 13.

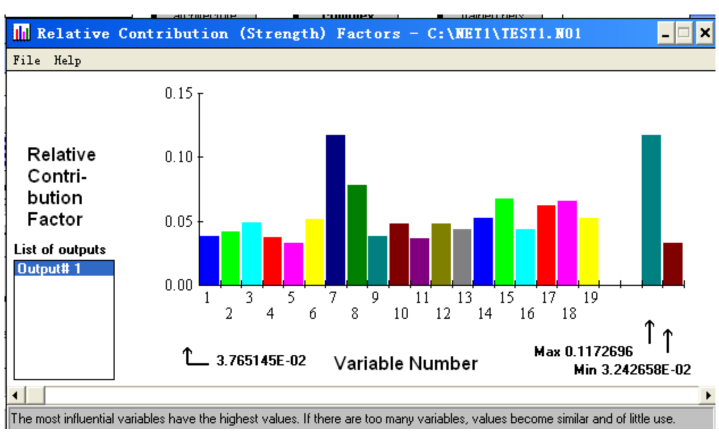

Figure 1 Top features

Table 9 Classification accuracy of Experiment 26 top1 as input

	Classified Class 1	Classified Class 2	Row Accuracy
Actual Class 1	0	4	0%
Actual Class 2	0	12	100%
Column Accuracy	0%	75%	

Table 10 Classification accuracy of Experiment 27 top 1 - 2 as input

	Classified Class 1	Classified Class 2	Row Accuracy
Actual Class 1	0	4	0%
Actual Class 2	0	12	100%
Column Accuracy	0%	75%	

Table 11 Classification accuracy of Experiment 28 top 1-3 as input

	Classified Class 1	Classified Class 2	Row Accuracy
Actual Class 1	0	4	0%
Actual Class 2	0	12	100%
Column Accuracy	0%	75%	

Table 12 Classification accuracy of Experiment 29 top1- 4 as input

	Classified Class 1	Classified Class 2	Row Accuracy
Actual Class 1	2	0	100%
Actual Class 2	0	14	100%
Column Accuracy	100%	100%	

Table 13 Classification accuracy of Experiment 30 top 1-5 as inputs

	Classified Class 1	Classified Class 2	Row Accuracy
Actual Class 1	2	2	50%
Actual Class 2	0	14	100%
Column Accuracy	100%	87.5%	

Conclusion:

Experiment 29 using top 1-4 has increased the accuracy to 100 and since we add one feature to experiment 29, the accuracy begins to go down. The result shows that joint features 8, 9, 16, 19 and 18 provide the perfect result and they are key features of the result.

5. Limitations

The limitation of our work lies in 3 aspects. First，the size of the data set is small and many of the samples in data are incomplete which result in smaller data set used for training and test of the model. So the result we get from the experiment may not be convincible. Second, according to the related paper we have read, there are many other

NN architectures and training algorithms used to build a high accuracy hepatitis diagnosis model. In our experiment, we just use the BP algorithm which does not have a better performance compared with other models. At last, in our experiment, not all possibility have been tested. In future studies, more experiments are needed to be done to ensure a better result of recognizing the influential features of the hepatitis diagnosis model.

6. Conclusion

This paper presents the procedure of deciding influential features in hepatitis diagnosis Neural Networks. A 3-layer MFNN network is built and trained by the BP algorithm. Through the adjustment of the hidden layer numbers and parameter of the network. We firstly get a high accuracy classification model. Based on this model, we have done a series of experiments trying to find the influential feature factor. After all experiments have been done, we find that using features ANOREXIA, LIVER BIG, ALK PHOSPHATE, and PROTIME as input will result in the highest diagnosis accuracy. Hence, we conclude that these four features are key features and influence the classification result most. However, due to the size of the data set is too small and the best classification algorithm has not been chosen, the result of these experiments may be relatively inconvincible. In spite of these limitations, we still can conclude that these experiments show the feasibility to do the classification using neural networks.

References

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

Roy, S. (1994). Factors influencing the choice of a learning rate for a backpropagation neural network. In the *1994 IEEE International Conference on Neural Networks and the 1994 IEEE World Congress on Computational Intelligence,* 1, 503-507.

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.