Study and Practice of Sleep Quality based on BP Neural Network Model

Qi Hao[1], Chao Ma[2], Nuo Xu[1], Zhi-hong Ma[3] Corresponding author
[1]College of Computer and Information Engineering, Tianjin Agriculture University, Tianjin, 300384, China,
[2]College of Horticulture and Landscape Architecture, Tianjin Agriculture University, Tianjin, 300384,China
[3]College of basic science, Tianjin Agricultural University, Tianjin, 300384, China

Abstract: As the number of insomnia in adults increases, the severe insomnia makes people study and work be relatively low. Aiming at the problem of how to diagnose disease from the sleep indicators, put forward an analytic model based on BP natural network. First, analyzing the relationship between indicators and quality of sleep, it can be found that each patient has at least one relevant score for a sleep indicator that is very high. Second, by establishing a sleep classification model based on BP neural network, study the correlation between various indicators of sleep and diagnosis results. Finally, it can accordingly provide the diagnosis result for the patients, on the basis of the analysis of sleep index by using the model.

Keywords: Stepwise regression; SPSS; BP; Neural network

1. Introduction

We mainly use MATLAB to organize the data, and find the relationship between sleep and diagnosis.

2. The Association of Sleep and Diagnosis

Data processing and visualization:
We use MATLAB software to preprocess the data and exclude several diagnostic results which take the smallest proportion in all the results. (The number 0.5% of the patient population is used as a threshold value, which means the disease with the number of patients above 0.5% will be remained and that below 0.5% will be eliminated.) Finally, it was found that the major diagnosis results were mainly concentrated in the following situations (The code is shown in Annex):

Table 1. 15 Types of Major Diseases Table

Value	Coun	Percent
Adjustment Disorder	58	1.07%
Anxiety	402	7.44%
Anxiety disorder	861	15.94%
Anxiety, Sleep disorder	50	0.93%
Bipolar Affective Disorder	132	2.44%
Consultation	36	0.67%
Depression	1472	27.25%
Mixed Anxiety And Depression	368	6.81%
Mixed Anxiety And Depression Disorder	32	0.59%
Non-Organic Insomnia	106	1.96%
Recurrent Depressive Disorder	74	1.37%
Schizophrenia	46	0.85%
Sleep disorder	1699	31.45%
Sleep disorder, Depression	33	0.61%
Sleep disorder, Mixed Anxiety And Depression	33	0.61%

Classify types, according to the diagnostic results, we found that some of the indicators are repeated. Therefore, in order to avoid excessive partial offset, we use MATLAB, based on the similarity of diagnostic results, to analyze all the indicators comprehensively. Finally, the diagnostic results are divided into nine categories, and we use MATLAB to calculate the average of the patient's sleep indicators. Then analyzing the data by Excel, we conclude each diagnostic result. So that we can visually see each correlation between the indicators and diagnostic result. Take some diseases as an example: Adjustment Disorder, Bipolar Affective Disorder, Depression and Recurrent Depressive Disorder are easily triggered when the Hypnagogue, a indicator of the patient, is too low. At the same time, if the indicator of patient's sleep latency reached very poor, the symptoms he suffered were Bipolar Affective Disorder; if the indicator of patient's sleep time is poor, the symptoms he suffered were Adjustment Disorder; but if the indicator of patient's sleep time is good, he may suffer with Depression or Recurrent Depressive Disorder.
Next, Excel software is used to compare various sleep indicators of each diagnosis results, to obtain the difference of the indicators, which affect the diseases. Draw as follows:

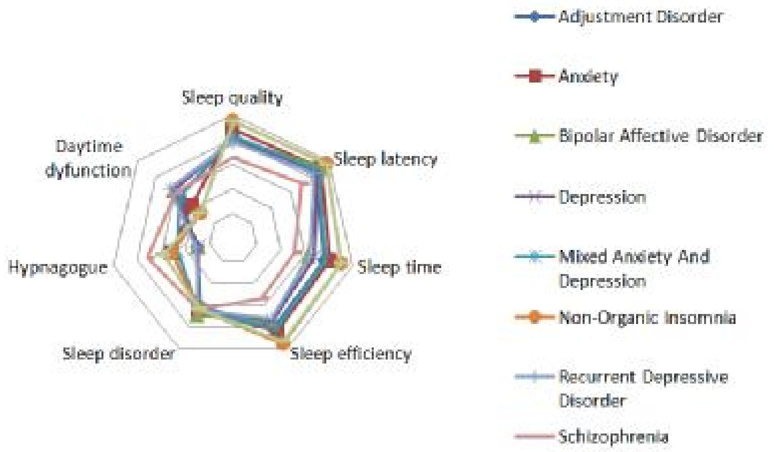

Figure 1. The Radar Chart of Diagnostic Results and Sleep Indicator

As can be seen from the figure above, each disease that has been diagnosed must have one or more sleep indicators, which have highly relevant scores. In the case of Schizophrenia, patient has four indicators that can be judged good, but his indicator of Hypnagogue is very poor. In addition, we can see that the scores of various sleep indicators of patients who have different diseases are also different. That is, according to the different relevant scores of each indicator, it is possible to determine what kind of diagnosis the patient may obtain.

Based on the MATLAB software, the variance of each index in each diagnosis result is listed to reflect the stability of the above indicators.

Table 2. 15 Types of Major Diseases Table

Diagnosis	Sleep quality	Sleep latency	Sleep time	Sleep efficiency	Sleep disorder	Hypnagogue	Daytime dyfunction
A. Disorder	0.490	0.975	1.0	1.0865	0.429	1.4	0.9764
Anxiety	0.636	0.742	1.0	1.0752	0.451	1.8	0.9057
B.A. Disorder	0.677	0.777	1.0	1.1512	0.588	1.4	0.7520
Depression	0.715	0.909	1.1	1.1963	0.520	1.4	0.9196
Mixed Anxiety And Depression	0.741	0.926	1.0	1.1540	0.521	1.8	1.0995
N.O. Insomnia	0.468	0.611	0.9	0.8417	0.511	1.9	0.8471
R.D. Disorder	0.614	0.799	1.1	1.1855	0.652	1.5	0.8103
Schizophrenia	0.949	1.151	1.4	1.2986	0.825	1.7	0.9662
Sleep Disorder	0.504	0.686	0.9	0.9113	0.427	1.9	0.8644

(Remark: A. Disorder=Adjustment Disorder; B.A. Disorder= Bipolar Affective Disorder; N.O. Insomnia= Non-Organic Insomnia; R.D. Disorder= Recurrent Depressive Disorder)

Table 2 is the variance value of various sleep indicators for certain diagnostic result. The value is smaller, the indicators are more stable and the mean of the data are more reliable.

To sum up, the different symptom can be estimated by the patients' relevant scores on different sleep indicators.

3. BP Neural Network Model of Diagnostic Classification

In this question, we establish a classification model based on BP neural network. The data constructed by the neural network come from the data of Annex II. We use 70% of the data in Annex II as the training set, 15% as the verification set and 15% as the test set.

3.1. Data Processing

We used MATLAB to classify the data in Annex II and analyze them, excluding the number of patients less than 0.5% of the total surveyed, compressing 289 symptoms to 15. The specific data as shown in the following table (Codes are shown in Annex 2).[1]

Table 3. Statistical Table of Diagnostic Results

Value	Count	Percent
Schizophrenia	46	0.85%

Depression	1472	27.25%
Sleep disorder, Depression	33	0.61%
Sleep disorder	1699	31.45%
Anxiety disorder	861	15.94%
Anxiety	402	7.44%
Bipolar Affective Disorder	132	2.44%
Adjustment Disorder	58	1.07%
Recurrent Depressive Disorder	74	1.37%
Anxiety, Sleep disorder	50	0.93%
Consultation	36	0.67%
Mixed Anxiety And Depression	368	6.81%
Mixed Anxiety And Depression Disorder	32	0.59%
Non-Organic Insomnia	106	1.96%
Sleep disorder, Mixed Anxiety And Depression	33	0.61%

As shown in Table 3, the findings of the population in the survey are mainly focused on Schizophrenia, Depression, Sleep disorder, Depression, Sleep disorder, Anxiety disorder, Anxiety, Bipolar Affective Disorder, Adjustment Disorder, Recurrent Depressive Disorder, Anxiety, Sleep disorder, Consultation, Mixed Anxiety And Depression, Mixed Anxiety And Depression Disorder, Non-Organic Insomnia, Sleep disorder, Mixed Anxiety And Depression. However, the percentage of other diagnoses is too small, so the reliability of the data is too low to be considered. To make sure that the model is reliable when we build the model, 70% of the data is used to construct the model and 30% of the data is used to test the model.

Through the data processing, we found that there is no linear relationship between the diagnostic results and sleep quality indicators. In order to get a clear and accurate correlation between sleep indicators and diagnostic results, we set up a classification model of BP neural network to study how the sleep indicators affect the diagnosis results.

BP neural network is composed of non-linear transform unit. The information processing process is determined by neurons, activation functions, network topology, connection weights and neurons threshold. Through error back propagation to constantly adjust the network weights and thresholds, so that the sum of the square error of the network will be minimum. [2]We build BP neural network model based on MATLAB (2014a) software to identify the classification of diagnosis results. The basic flow chart is as follows:

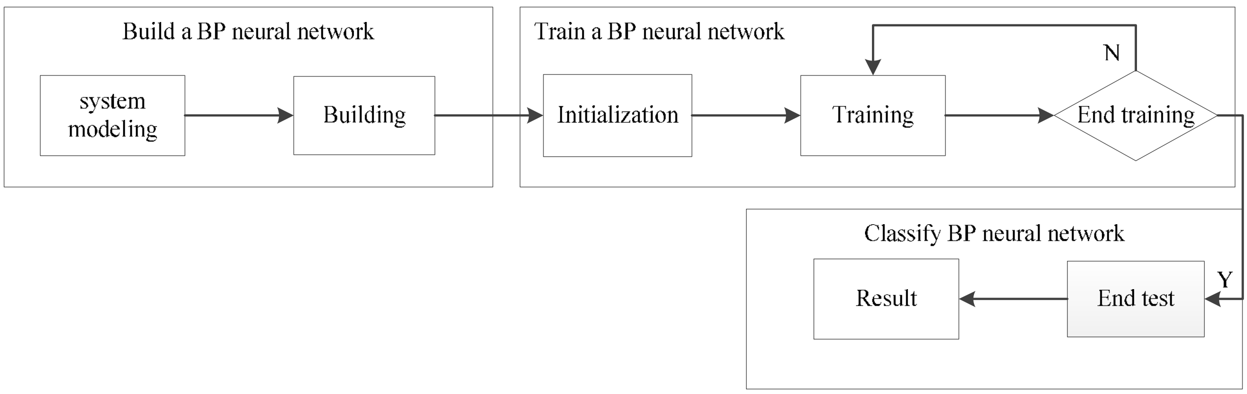

Figure 2. Operation Flow Chart

The topology is shown in Figure 3.

The input layer is the indicators of sleep. The output layer is the diagnosis of the disease. And the hidden nodes (the number of hidden neurons) need to be given subjectively.

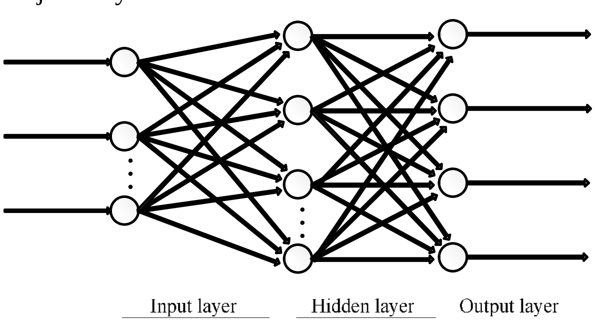

Figure 3. Neural Network Topology

Before BP neural network prediction, it is necessary to train the network. Through training, the network has the ability of associative memory and predictive power. The steps are as follows:

Step 1: Network initialization. According to the input and output sequences we can obtain the following: network input layer node n, the number of hidden layer nodes l, the number of output layer nodes m, initialize the input layer, the hidden layer and output layer neurons connection weights initialize the hidden layer threshold a, the output layer thresholdb, given learning rate and neuron excitation function.

Step 2: output calculation of implicit function. The hidden layer output H is calculated based on the input vector X, the input layer, the hidden layer connection weight and the hidden layer threshold a.

$$H_J = f(\sum_{i=1}^{n} w_{ij} x_i - a_j)\ j = 1, 2, \ldots,\ l$$

In this formula, l is the number of hidden layer nodes and f is excitation function in the hidden layer. The function has many forms of expression and the function be selected here is:

$$f(x) = \frac{1}{1+e^{-x}}$$

Step 3: output calculation in the input layer. According to hidden layer, H can be output, then connect weight W_{jk} and threshold b, so that BP neural network prediction output O can be calculated.

$$O = \sum_{j=1}^{l} H_j w_{jk} - b_k \quad k = 1, 2, \ldots, m$$

Step 4: Error Calculation. According to network prediction output O and expectation output Y, we can calculate network prediction errore.

$$e_k = Y_k - O_k \quad k = 1, 2, \ldots,\ m$$

Step 5: Weight update. According to the network prediction error e update network connection weight.

$$w_{ij} = w_{ij} + hH_j(1-H_j)x(i)\sum_{k=1}^{m} w_{jk} e_k ? = \quad \ldots\ n\ j = \quad \ldots\ l$$

$$w_{jk} = w_{jk} + hH_j e_k\ j = 1, 2, \ldots, l; k = 1, 2, \ldots, \mathrm{m}$$

$$(\sum_{i=1}^{n} w_{ij} x_i - a_j)\ j = 1, 2, \ldots,\ l$$

In this formula, η is the learning rate.

Step 6: Threshold update. The network node threshold a, b are updated based on the network prediction error e.

$$a_j = a_j + hH_j(1-H_j)\sum_{k=1}^{m} w_{jk} e_k\ j = 1, 2, \ldots, l$$

$$b_k = b_k + e_k\ k = 1, 2, \ldots, m$$

Step 7: Determine whether the algorithm iteration is over. If not, return to Step 2.

3.2. Model solving and testing

3.2.1 Data preparation

By using the MATLAB pattern recognition toolbox, the input data is automatically normalized and the output data is quantized. [3] Since the output of the diagnostic results can be summarized as 15 types, so it can be represented to 16 categories by the use of four-digit binary representation （0001,0010,0011,0100， ⋯ ， 1111）. For example,1 is equal with 0001. After a simple process, the input data and the newly added binary input data as shown:

Table 4. The indicators of sleep

Diagnosis	Input							Original output	Binary output			
Schizophrenia	1	0	0	0	2	2	2	8	1	0	0	0
Schizophrenia	2	3	2	1	2	2	0	8	1	0	0	0
Depression	1	1	1	1	0	2	0	4	0	1	0	0
Depression	3	3	3	3	3	3	2	4	0	1	0	0
Sleep disorder	2	3	0	3	2	0	2	9	1	0	0	1
Depression	2	0	1	2	2	0	2	4	0	1	0	0
Depression	3	3	3	3	2	0	1	4	0	1	0	0
Sleep disorder	3	3	3	3	1	0	1	9	1	0	0	1
Sleep disorder	3	3	2	2	2	3	1	9	1	0	0	1
Depression	1	1	1	0	2	3	1	4	0	1	0	0
Sleep disorder	3	2	3	3	3	3	1	9	1	0	0	1
Sleep disorder	3	2	3	3	1	3	2	9	1	0	0	1
Depression	3	3	3	3	2	0	2	4	0	1	0	0
Sleep disorder	2	2	3	1	1	0	0	9	1	0	0	1
……	…	…	…	…	…	…	…	…	…	…	…	…

The table above transforms the data given in Annex II into a matrix of four rows for calculating the neural network.

3.2.1. Network training will import the training sample monitoring point data into MATLAB

Network training will import the training sample monitoring point data into MATLAB, and using MATLAB's own pattern recognition toolbox to adjust the parameters to get the result.[4] Because network training can be different every time, so we use the mean square error and confusion matrix to evaluate the effectiveness of the network. If the accuracy of the result is too low, continue to adjust it and train several times to correct the neural network until the training results are satisfactory. The results shown in Figure:

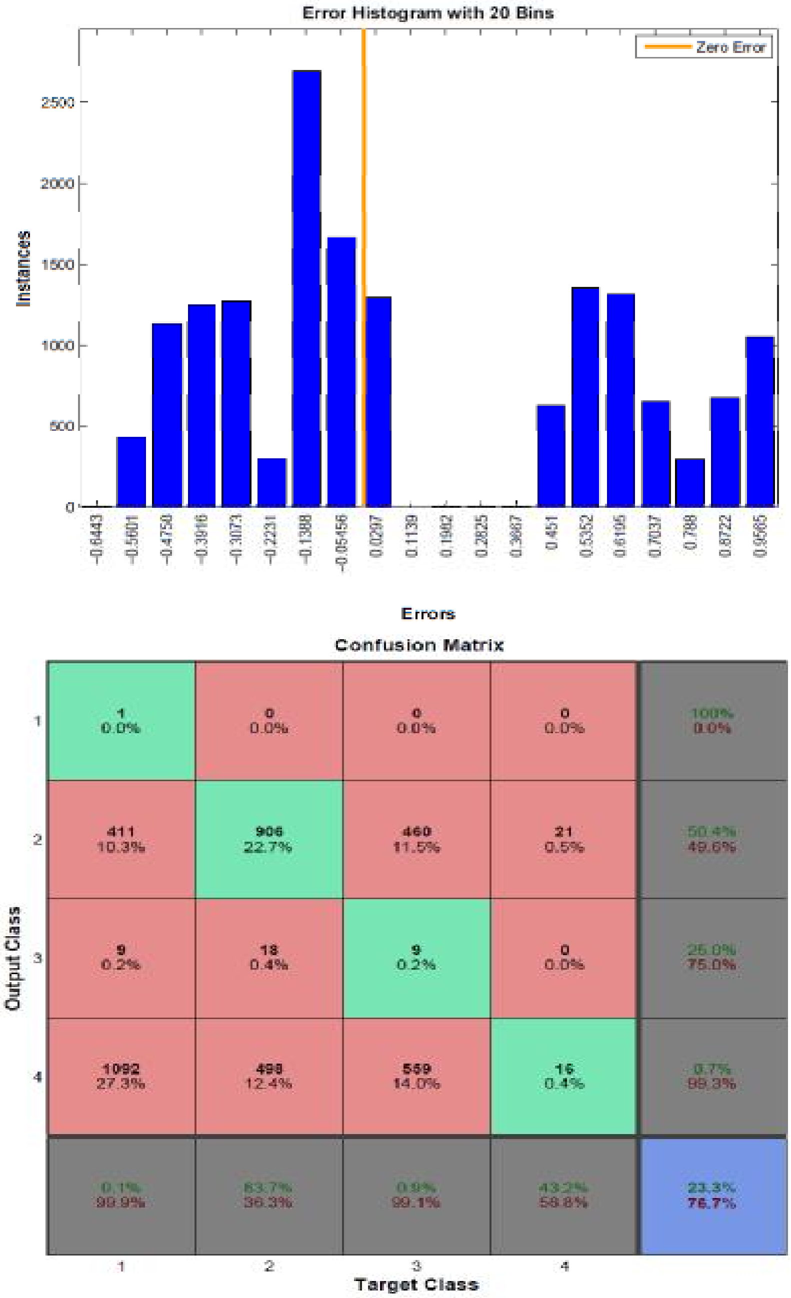

Figure 4. Error Histogram with 20 bins and Confusion Matrix

The final error of BP neural network is shown in the figure above. So we can see that most of the training results are close to 0 error, so the neural network is feasible to a great extent. The results of network training can be different each time, so we use the mean square error and confusion matrix to evaluate the effect of the network. If the result is not satisfactory (the accuracy is less than 80 percent), continue to adjust the relevant parameters and train many times until the satisfactory result is achieved.

3.3. Model Prediction Test

Using 30% of the data that not involved in the model to test, and import the data into the model. The form of the data is the same with the training, each input data also has seven indicators, so that we can obtain the mean square test figure and error test figure:

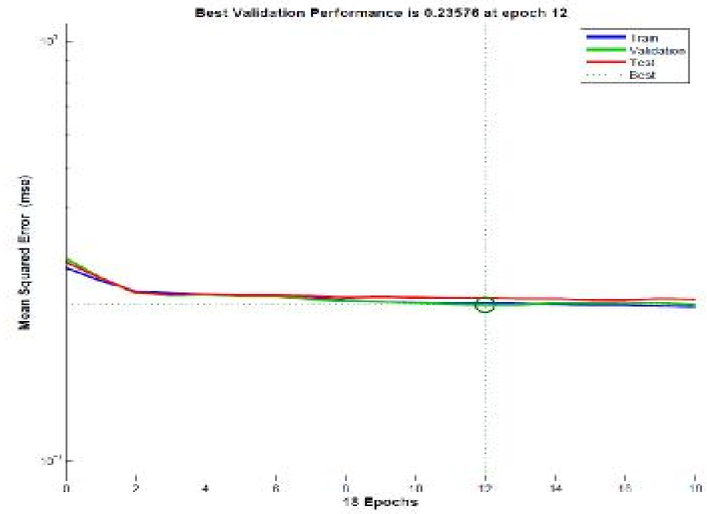

Figure 5. Best Validation Performance

It can be seen from Figure 6 that when the 40th iteration is reached, the effect of the neural network training is optimal. Therefore, we choose this network as the neural network model of this question.

3.4. The diagnosis of the Patient

Importing the sleep indicators in Annex III into the neural network model, the diagnosis of each patient can be obtained as follows:

Table 5. Diagnostic Results

Number	Sleep quality	Sleep latency	Sleep time	Sleep efficiency	Sleep disorder	Hypnagogue	Daytime dysfunction	Diagnosis
1	1	0	1	0	1	1	1	Depression
2	2	1	2	1	2	1	0	Sleep disorder
3	3	2	3	2	3	0	0	Sleep disorder
4	3	3	3	0	2	2	2	Sleep disorder
5	2	2	2	3	0	3	0	Sleep disorder
6	3	2	2	0	1	3	1	Depression
7	1	1	1	1	0	2	2	Sleep disorder
8	1	3	3	2	1	1	1	Sleep disorder
9	3	2	2	3	3	2	0	Sleep disorder
10	1	2	3	2	1	2	1	Sleep disorder

Table 5 shows the diagnosed results of each patient in Annex III. The diagnosis of patient No. 1 and No. 5 are Depression and patient No. 2, No. 3, No. 4, No. 6, No. 7, No. 8 and No. 9 were diagnosed as Sleep Disorder.

4 . Conclusions

In this paper, it can diagnose results of each patient by using the BP natural network model. Unlike traditional classification model, BP Neural Network Classification Model has the characteristics of stable results, large diagnosis range and low requirements (which can be discrete variables or continuous variables). But the establishment of the model requires high technology and parameters, which are more difficult to set.

Acknowledge

Tianjin Agricultural University student innovation and entrepreneurship training project (No. 201808018).

References

[1] Wang Yifu, Sun Yujun, Guo Xiaoyi. Study on Masson Pine Stand Biomass Model Based on BP Neural Network [J]. Journal of Beijing Forestry University,2013,35(02):17-21.

[2] Yu Xiuli. Study on Prediction of Environmental Pollution Degree Based on Pattern Recognition in [J]. Computer and Information Technology,2017,25(05):16-18.

[3] Huang Li. BP neural network algorithm improvement and Application Research [D]. Chongqing Normal University, 2008.

[4] Shen Jianbo, Lei x D, Li Yutang, Lan Ying. Bayesian regularization BP neural network prediction of average height of Larix olgensis plantation [J/OL]. Journal of Nanjing Forestry University (NATURAL SCIENCE EDITION), based on 2018, 1-11 (2017-09-25).