Hybrid binary quantization and BP network on action mechanism

Chenhui Jiang¹, Zhiyang Shangguan¹, Huan He¹, Jinglou Zhang², Jing Ping¹ and Naiwei Li*¹

¹Fuzhou Medical College of Nanchang University, Fuzhou Jiangxi China 344000
²Shandong Anran Nanometer Industry Development Co., Ltd, Weihai Shandong China 264200

*Corresponding author and e-mail: Naiwei Li, 201307020102@hnu.edu.cn

Abstract. BP neural network is a kind of artificial neural network, which is mainly used in function approximation, model recognition, classification and data compression. In the case of limited analysis ability of traditional research methods, BP neural network can overcome the problems of nonlinear relationship. In this paper, the improved BP neural network was used to predict the properties. The data processing of the properties, tastes and was carried out by binary quantization. The prediction results' correct rate was 87.5%.

1. Introduction

Artificial neural network has the characteristics of self-learning and self-adaptive, and it can deal with nonlinear data well, which just corresponds to the nonlinear and high-dimensional characteristics of traditional Chinese medicine data. In recent years, it has been widely used in the field of traditional Chinese medicine [1]. Liu Xing et al. [2] took the near infrared spectral data of job's tears from different producing areas and varieties as the research object, and established the classification and recognition model of job's tears by using the learning vector neural network, with the prediction accuracy of more than 90%; Xiang Cuiyu et al. [3] near infrared diffuse reflectance spectroscopy combined with artificial neural network was used to identify Rhubarb from different areas, and the recognition accuracy was 95%; Tang Yanfeng et al. [4] used Fourier transform infrared spectroscopy to scan 42 kinds of Viola samples, and used radial basis function neural network to establish a classification model to identify wild Viola and cultivated Viola, and the recognition accuracy reached 95.24%; Li Fang et al. [5] identified and predicted the origin of Radix Astragali by FTIR (Fourier transform infrared absorption spectrometer), and the prediction accuracy was more than 83%. In this paper, the improved BP neural network is used to predict the tonic category of traditional Chinese medicine, and it is realized by Python language.

Traditional Chinese medicine industry is a strategic industry with broad prospects and development advantages in China's current social and economic development. The property theory of traditional Chinese medicine is the theory of traditional Chinese medicine summarized and refined by the predecessors in the long-term practice of medicine. The efficacy is a high generalization of the therapeutic effect of drugs according to the theory of traditional Chinese medicine, and both are important parts of the theory of traditional Chinese medicine [6]. Jingui Shenqi Pill was discovered by
Zhang Zhongjing, a medical sage. Through the continuous clinical practice of doctors in past dynasties, the application of this prescription has been expanded on the original basis. It is still used today because of its remarkable efficacy.

2. Principle

BP neural network is a concept put forward by scientists headed by Rumelhart and McClelland in 1986. It is a kind of multilayer feedforward neural network trained according to the error back propagation algorithm and is the most widely used neural network. BP network is composed of input layer, hidden layer and output layer. Given the training set \( D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\} \), where \( x_n \in \mathbb{R}^d \), \( y_n \in \mathbb{R}^l \), the input sample is composed of \( D \) attributes, and the output is \( L \)-dimensional real valued variable. The BP neural network model is shown in Figure 1.

![Figure 1. BP neural network model.](image)

Neuron is a topological network based on biological research and brain response mechanism, which simulates the process of neural conflict. The terminals of multiple dendrites receive external signals and transmit them to neurons for fusion. Finally, they transmit them to other neurons or effectors through axons. The topological structure of neurons is shown in Figure 2.

![Figure 2. Topology structure of neurons.](image)

For the \( i \)-th neuron, \( x_1, x_2, \ldots, x_j \) is the input of the neuron, the input is often the independent variable which has a key impact on the system model, and \( w_1, w_2, \ldots, w_j \) is the connection weight to adjust the weight ratio of each input. There are many ways to input signals to neurons. The most convenient linear weighted summation method is selected to obtain the net input of \( N_i \) neurons:

\[
Net_{in} = \sum_{i=1}^{n} w_i \times x_i
\]  

(1)
$\theta$ is the threshold of the neuron. According to the knowledge of biology, the neuron will be activated only when the information received reaches the threshold. Therefore, $Net_{in}$ and $\theta$ are compared and then processed by the activation function to produce the output of the neuron.

If the output value has a certain range constraint, such as for classification, the most commonly used function is sigmoid function, which can transform the input signal from negative infinity to positive infinity into output between 0 and 1. If there are no constraints, you can use a linear activation function (that is, the sum of weights multiplied). In this way, the output is as follows:

$$y_j = f(Net_{in} - \theta_i)$$  \hspace{1cm} (2)

Through the simplification of this formula, set up the first input value is a constant for $\theta$, weight of $-1$, then we can get the formula:

$$y_j = f(\sum_{i=0}^{g} w_i * x_i)$$  \hspace{1cm} (3)

Where $w_0 = -1, x_0 = \theta_j$, and $f$ is the activation of the selected function.

By above knowable, there’re a total of three layers BP neural network structure, namely: the input layer, hidden layer and output layer, and the weights of input layer to hidden layer can be set to $V_{ih}$, the first $h$ a hidden layer neurons of the threshold is set to $\gamma_h$. And the hidden layer to output layer weights set to $w_{jh}$, the first $j$ output layer neuron threshold value is expressed with $\theta_j$. In Figure 3, a total of $d$ input neurons, $q$ hidden neurons and $l$ output neurons.

$$y_i^k = f(\beta_i - \theta_i)$$  \hspace{1cm} (4)
The error of the prediction result is expressed by the least square method. And adjust the value of \((d + l + 1) \times q + l\) parameters according to the error, and reduce the error step by step. After that, the parameters are updated by gradient descent method. Since the function always changes the fastest along the direction of the gradient, the partial derivative is calculated for each parameter to be adjusted. If the partial derivative is greater than 0, it changes in the opposite direction of the partial derivative, otherwise it changes in the current direction. Here, we can use \(-1\) partial derivative, that is, we can get the required value of parameter change. At the same time, set a learning rate, and the learning efficiency \(\eta\) can not be too fast or too slow. (too fast will easily lead to over optimal solution, if too slow will reduce the efficiency of the algorithm). A parameter adjustment formula can be obtained:

\[
Param^+ = -\eta \frac{\partial E_k}{\partial Param} \tag{5}
\]

The following is the weight adjustment value from hidden layer to output layer:

\[
\Delta w_{hj} = -\eta \frac{\partial E_k}{\partial w_{hj}} \tag{6}
\]

Calculation process:

a. Input layer to hidden layer:

\[
\alpha_h = \sum_{i=1}^{d} v_{ih} * x_i \tag{7}
\]

Matrix is available:

\[
\begin{bmatrix}
  v_{11} & v_{12} & v_{13} & \cdots & v_{1q} \\
  v_{21} & v_{22} & v_{23} & \cdots & v_{2q} \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  v_{d1} & v_{d2} & v_{d3} & \cdots & v_{dq}
\end{bmatrix}
\]

b. Through the hidden layer activation function:

\[
b_h = f(\alpha_h - \gamma_h) \tag{8}
\]

c. Hidden layer to output layer:

\[
\beta_h = \sum_{h=1}^{g} w_{hj} * b_h \tag{9}
\]

Matrix is available:

\[
\begin{bmatrix}
  w_{11} & w_{12} & w_{13} & \cdots & w_{1l} \\
  w_{21} & w_{22} & w_{23} & \cdots & w_{2l} \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  w_{q1} & w_{q2} & w_{q3} & \cdots & w_{ql}
\end{bmatrix}
\]

d. After activation function of the output layer:
\[ y^k_j = f(\beta_j - \theta_j) \] (10)

e. Error calculation.

3. Experiment and discussion

3.1. Medicinal quantitative indicators

The main components of Jingui Shenqi Pill are Rehmannia glutinosa, Chinese yam, Cornus officinalis, Rhizoma alismatis, Cortex Moutan, Poria cocos, Ramulus Cinnamomi and processed aconite. This model considers the relationship between the pharmacodynamics of traditional Chinese medicine and 22 attributes of five flavors, four Qi and meridian tropism. In this chapter, we use the method of binary quantization to carry out numerical processing. For example, the five flavors of traditional Chinese medicine are sour, bitter, sweet, pungent and salty. The characteristic of each herb is marked as 1, otherwise it is 0. In the same way, the four Qi, meridian tropism and tonic category are also treated by binary quantization. The following table shows some quantitative results:

| Traditional Chinese medicine name | Cold | Hot | Lukewarm | Cool | Salty | Sour | Bitter |
|----------------------------------|------|-----|----------|------|-------|------|--------|
| Glutinous Rehmannia              | 0    | 0   | 1        | 0    | 0     | 0    | 0      |
| Yam                              | 0    | 0   | 0        | 1    | 0     | 0    | 0      |
| Cornus Officinalis               | 0    | 0   | 1        | 0    | 0     | 1    | 0      |
| Alisma Orientalis                | 1    | 0   | 0        | 0    | 0     | 0    | 0      |
| Cortex Moutan                    | 1    | 0   | 0        | 0    | 0     | 0    | 1      |
| Poria Cocos                      | 0    | 0   | 0        | 1    | 0     | 0    | 0      |
| Ramulus Cinnamomi                | 0    | 0   | 1        | 0    | 0     | 0    | 0      |
| Fried Aconitum Carmichaeli       | 0    | 1   | 0        | 0    | 0     | 0    | 0      |

| Traditional Chinese medicine name | Sweet | Bladder | Qi | Yang | Blood | Yin | Pungent |
|-----------------------------------|-------|---------|----|------|-------|-----|---------|
| Glutinous Rehmannia               | 1     | 0       | 0  | 0    | 1     | 0   | 0       |
| Yam                               | 1     | 0       | 1  | 0    | 0     | 0   | 0       |
| Cornus Officinalis                | 0     | 0       | 0  | 0    | 1     | 0   | 0       |
| Alisma Orientalis                 | 1     | 1       | 0  | 0    | 0     | 1   | 0       |
| Cortex Moutan                     | 0     | 0       | 0  | 0    | 0     | 1   | 1       |
| Poria Cocos                       | 1     | 0       | 0  | 1    | 0     | 0   | 0       |
| Ramulus Cinnamomi                 | 1     | 1       | 0  | 1    | 0     | 0   | 1       |
| Fried Aconitum Carmichaeli        | 1     | 0       | 0  | 1    | 0     | 0   | 1       |

Python is widely used in the field of data processing because it is easy to learn, fast and has rich extended class library. In this paper, the BP neural network model is constructed by using Python numpy and other libraries. According to the flow chart of BP neural network, first of all, combined with the experimental object of this paper, that is, the five flavors, four Qi and meridian tropism of traditional Chinese medicine, 22 parameters are selected as the input unit, and the four parameters of Tonifying Qi, Yang, blood and Yin are selected as the output unit, and the output range is set to 0 ~ 1. After comprehensive consideration, the hidden layer of the model is determined as one layer, and the number of neurons in the hidden layer is determined to be 44, so as to establish the BP neural network.
model of traditional Chinese medicine efficacy, and test and train it. In this paper, a three-layer BP neural network with 22 input layer nodes, 44 hidden layer nodes and 4 output layer nodes is established. Because sigmoid can better reflect the input and output characteristics of neural network, this paper uses the S-type function. Eight traditional Chinese medicines of Jingui Shenqi pills were included in this model as training data.

3.2. Output result and discussion of BP neural network
After 800 iterations, the weight error reaches the optimal value. The actual output results calculated by Python are shown in Table 2 and Table 3, and the accuracy rate is as high as 87.5%, which is less different from the actual results, so the model is more successful.

5. Conclusions.

Table 2. Expected output and actual output.

| Traditional Chinese medicine name | Expected output | Actual output |
|----------------------------------|-----------------|--------------|
|                                  | Tonify Qi | Tonify Yang | Tonify Blood | Tonify Yin | Tonify Qi | Tonify Yang | Tonify Blood | Tonify Yin |
| Glutinous Rehmannia              | 0          | 0           | 1            | 0          | 0.149     | 0.13508     | 0.941        | 0.10893    |
| Yam                              | 1          | 0           | 0            | 0          | 0.94781   | 0.15189     | 0.14028      | 0.16513    |
| Cornus Officinalis               | 0          | 0           | 1            | 0          | 0.14877   | 0.13325     | 0.92992      | 0.17754    |
| Alisma Orientalis                | 0          | 0           | 0            | 1          | 0.13708   | 0.17199     | 0.17647      | 0.93572    |
| Cortex Moutan                    | 0          | 0           | 0            | 1          | 0.11176   | 0.10238     | 0.10318      | 0.93268    |
| Poria Cocos                      | 0          | 1           | 0            | 0          | 0.17433   | 0.9133      | 0.12408      | 0.13554    |
| Ramulus Cinnamomoi              | 0          | 1           | 0            | 0          | 0.38451   | 0.19065     | 0.11884      | 0.15115    |
| Fried Aconitum Carmichaeli      | 0          | 1           | 0            | 0          | 0.14739   | 0.92768     | 0.1892       | 0.18341    |
Table 3. Comparison of actual and expected values of samples.

| Traditional Chinese medicine name       | Tonify Qi | Tonify Yang | Tonify Blood | Tonify Yin |
|----------------------------------------|-----------|-------------|--------------|------------|
| Glutinous Rehmannia                    | correct   | correct     | correct      | correct    |
| Yam                                    | correct   | correct     | correct      | correct    |
| Cornus Officinalis                     | correct   | correct     | correct      | correct    |
| Alisma Orientalis                      | correct   | correct     | correct      | correct    |
| Cortex Moutan                          | correct   | correct     | correct      | correct    |
| Poria Cocos                            | correct   | correct     | correct      | correct    |
| Ramulus Cinnamomii                     | correct   | error       | correct      | correct    |
| Fried Aconitum Carmichaeli             | correct   | correct     | correct      | correct    |

Note: $7/8 = 87.5\%$

4. Conclusions
The experimental results show that only Guizhi Yiwei is wrong in the eight ingredients of Jingui Shenqi pills tested in this paper, and the accuracy rate is as high as 87.5, which proves that the model has a certain reference value in the analysis and prediction of traditional Chinese medicine properties. Based on the basic data of four Qi and five flavors and the relationship between meridian tropism and pharmacodynamics, the model can predict the tonic category. Because most of the data used in this paper are sweet and belong to the kidney meridian, the data have strong similarity. However, the relationship between the properties of traditional Chinese medicine and each component is nonlinear, and the mechanism of action is complex. It is difficult to achieve a comprehensive analysis with traditional research methods. As the most widely used neural network, BP neural network has strong self-learning ability, adaptability and nonlinear relationship processing ability. Therefore, it is feasible to combine BP neural network with traditional Chinese medicine property to establish a model to conduct a systematic research.

Acknowledgement
Science and Technology Research Project of Jiangxi Provincial Department of Education in 2016. No: GJJ160184; Project Name: Study on the action mechanism of Jinkui Shenqi Pill "tonifying kidney" based on the key gene of embryonic stem cells - Nanog gene.

References
[1] Zhou Rongrong, Yan Runhong, Wang liping, et al. The data mining technology in the application of prescription of traditional Chinese medicine scientific research [J]. Chinese journal of traditional Chinese medicine, 2018 (9): 4016-4020.
[2] Liu Xing, Mao Danzhuo, Wang Zhengwu, et al. Yi kinds of near infrared spectral technology quickly identify [J]. Spectroscopy and spectral analysis, 2014 (5): 1259-1263.
[3] Xiang Cuiyu, Song Xiaoyong. Chinese herbal medicine rhubarb of near infrared spectral analysis and artificial neural network to identify [J]. When Jane national physician national medicine, 2012, 12 (3): 3168-3169.
[4] Tang Yanfeng, Hou zhong, Wang Zhibao, et al. The infrared spectrum of herba violae and RBF neural network identification [J]. Journal of hubei agriculture, does 2014 (1): 132-134.
[5] Li fang, Li Qin Gu Zhirong et al. Based on FTIR huangqi herbs origin identification and prediction research [J]. Journal of natural products research and development, 2016, 28(11): 1712-1717.
[6] Li Yang, Feng Xv, Song Ge, et al. The research method of the traditional Chinese medicine medicinal progress [J]. Journal of medical information, 2018, 31 (13): 38-40.