Research on the Evaluation Model of Higher Education Health Level

Yiming Zhu

Shanghai University of Engineering Science, Shanghai, 201620, China

Keywords: Higher education, BP neural network, Genetic algorithm.

Abstract: In the past four decades, the expansion of higher education has become a mainstream trend in the development of higher education in the world and has had a profound impact on all dimensions of society and the higher education system. From a global perspective, the proportion of people receiving higher education is getting higher and higher. Whether the health of the higher education system affects the future development of a country, so it is very important to establish a higher education health level evaluation system. This article builds evaluation models for the health status and sustainability of higher education in various countries by BP neural network and Genetic algorithm. Then use the model to evaluate higher education in all countries. Based on the evaluation results, this article provides targeted suggestions for the better development of the higher education system. The health evaluation model and sustainability evaluation model established in this article cannot only be evaluated in higher education, but also can be evaluated in various industries, such as politics, economy, military, and agriculture. The short-term and long-term evaluation systems are similar Combined, the evaluation results obtained are more accurate and objective, and have more reference value.

1. Introduction

In the past four decades, the expansion of higher education has become a mainstream trend in the development of higher education in the world and has had a profound impact on all dimensions of society and the higher education system. From a global perspective, the proportion of people receiving higher education is getting higher and higher. By 2014, the gross enrollment rate of higher education in 64 countries reached 50%. Scholars are presented by Martin Trow have conducted research on education in the popularization stage of higher education and believe that higher education in the popularization stage will shape a new relationship between the country, education, and society [1]. Higher education plays a vital role in the development of the country, so it is necessary to determine an accurate evaluation system for the health level of higher education.

2. Overall problem analysis

This article is a standard evaluation topic. We build evaluation models for the health status and sustainability of higher education in various countries and use the model to evaluate higher education in all countries, and based on the evaluation results, we provide Higher education put forward corresponding opinions and suggestions [2]. First, this article makes the following assumptions: 1. It is assumed that the education data collected from all countries are true and reliable. 2.If only the 12 factors mentioned in the article are used to represent all factors that may affect higher education. 3.Assume that only the 6 countries mentioned in the article are used to represent all the countries to be analyzed. 4.Higher education institutions are intricate and diverse. It is assumed that the analysis process is not affected by extreme data.

3. Model establishment and solution

3.1 National Higher Education Measurement Index

In the data collection of this article, we divide the indicator set into the basis of national higher education, now and in the future three directions [3], a total of 12 indicators, as shown in table 1.

Table 1. The main factors affecting the level of education

Indicator classification	Indicator set
National Higher Education Foundation	Number of papers
	Degree Value (QS Rankings)
	Number of cumulative Nobel laureates
	Cumulative Fields Award winners
	Education index
State higher education now	Number of students
	Proportion of national education input
	Number of universities
	Average tuition fees for higher education
The Future of National Higher Education	Students' prospects (employment rate)
	Access to education (enrollment)
	Equity (gender ratio)

In order to eliminate the influence of size and order of magnitude between various indicators, and to ensure the progress of the neural network, we perform data normalization. The normalization formula is as follows:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{1}$$

Where normalized data are the maximum and minimum values of the original data.

For the 12 indicators of a sample quantified above, if they are directly analyzed for data, they will be affected by noise and extreme data. At the same time, the number of input layers and too many data will seriously affect the convergence speed. Therefore, it is necessary to adjust and reduce the dimensions of all indicators. Since the extraction of each indicator set has nothing to do with the year, it is directly related to 6 countries for 5 years the data, a total of 30 samples are subjected to principal component analysis.

If the correlation coefficient between indicators and indicators is set, then

$$r(X_i, X_j) = \frac{\mathrm{Cov}(X_i, X_j)}{\sqrt{\mathrm{Var}[X_i]\mathrm{Var}[X_j]}} \tag{2}$$

Where $r(X_i, X_j)$ is correlation coefficient, $\mathrm{Cov}(X_i, X_j)$ is Covariance of X_i and X_j,$\mathrm{Var}[X_i]$ is variance of X_i, $\mathrm{Var}[X_j]$is variance of X_j

We can judge whether the principal component analysis method is feasible according to the size of the correlation coefficients among all indicator variables. If most of the $r(x, X)$ is greater than or equal to 0.3, it indicates that the principal component.

This article establishes the evaluation system of BP neural network based on the three first-level indicators extracted in the previous part. Due to lack of data and training set, we try to use standard data for training and evaluate the health status of higher education in 5 countries. In the evaluation at that time, we regarded the health status of higher education in the United States as level 5, and the health status of higher education in South Africa as level 1 and classified by level.

We can judge whether it is feasible as input data based on the size of the correlation coefficients among all indicator variables. If most of the $r(x, X)$ are greater than or equal to 0.3, it indicates that the use of three first-level indicators is initially feasible. It can be used as the input layer of the neural network to jointly determine the output result. The level of correlation of inputs are shown as table 2.

Table 2. Level of correlation of inputs

Components	National Higher Education Foundation	State higher education now	The Future
National Higher Education Foundation	1.000	0.212	-0.719
State higher education now	0.212	1.000	-0.216
The Future of National Higher Education	-0.719	-0.216	1.000

3.2 Overview of BP neural network

Multi-layer feed forward neural network includes BP neural network. Signal forward transmission and error back propagation are the key features of this type of neural network. When it is in the forward pass, it starts from the input signal and reaches the input layer, then passes through the hidden layer, and finally reaches the output layer. The processing is terminated layer by layer. During this period, the neuron states of different layers are being transmitted up and down, which will only affect the neuron state of the next layer. If the output layer does not have the expected output, it will change to back propagation, thus relying on the prediction error to adjust the network weight and threshold. In this way, the BP neural network can predict more and more accurately and keep approaching the expected output [4]. We can see the topological structure of the BP neural network through Figure 1 below.

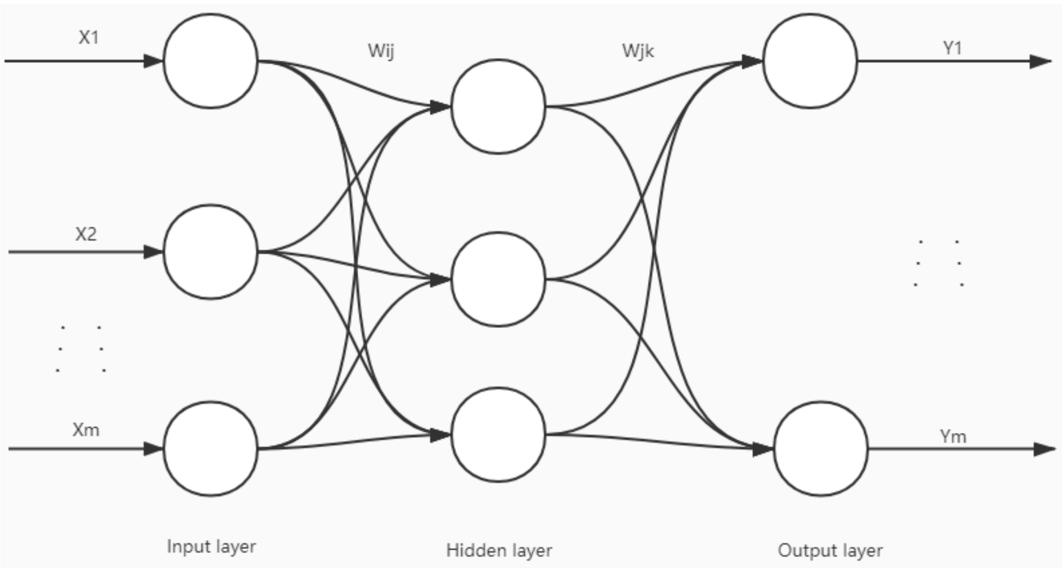


Figure 1. BP topology of neural network

Before using the BP neural network to make predictions, the most important step is to train the network. Only in this way can the neural network have the predictive ability through associative memory. In the BP neural network training process, there are mainly the following seven steps.

Step 1: Initialize the neural network. Relying on the input and output sequence in the neural network system, the number of input layer nodes, hidden layer nodes and output layer nodes is determined in sequence. At the same time, the connection weights between the input layer, hidden layer and output layer neurons are all reset to zero. And need to reset the hidden layer threshold and the output layer threshold to zero, and finally determine a certain learning speed and neuron activation function.

Step 2: The hidden layer output calculation. According to the input vector, the connection weight between the input layer and the hidden layer, and the hidden layer threshold, the hidden layer output is calculated.

$$H_j = f\left(\sum_{i=1}^{n} \omega_{ii} x_i - a_j\right) \quad j = 1,2,\dots,l \tag{3}$$

$$f(x) = \frac{1}{1 + c^x} \tag{4}$$

Where H is implicit layer output, f is implicit layer excitation function, ω is input layer and hidden layer ask, a is implicit threshold, l is the number of hidden layer nodes.

Step 3: Output calculation of output layer. According to the output of the hidden layer, connect the weight and the threshold to calculate the prediction output of the BP neural network.

$$O_k = \sum_{j=1}^{l} H_j \omega_{jk} - b_k \quad k = 1,2,\dots,m \tag{5}$$

Where O is network Forecast Output.

Step 4: Error calculation. According to the predicted output of the network and the expected output. the network predicted error is calculated.

$$e_k = Y_k - O_k \quad k = 1,2,\dots,m \tag{6}$$

Where e is network prediction error, Y is expected output.

Step 5: The weight is updated. Update network connection weights according to network prediction errors e.

$$\omega_{ij} = \omega_{ij} + \eta H_j\left(1 - H_j\right) x(i) \sum_{k=1}^{m} \omega_{jk} e_k \quad i = 1,2,\dots,n; j = 1,2,\dots,l \tag{7}$$

$$\omega_{jk} = \omega_{jk} + \eta H_j e_k \quad j = 1,2,\dots,l; k = 1,2,\dots,m \tag{8}$$

Where η is the learning rate.

Step 6: Threshold update. Update the network node threshold a and b according to the network prediction error e.

$$a_j = a_j + \eta H_l\left(1 - H_j\right) \sum_{k=1}^{m} \omega_{jk} e_k \quad j = 1,2,\dots,l \tag{9}$$

$$b_k = b_k + e_k \quad k = 1,2,\dots,m \tag{10}$$

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

3.3 Model establishment

In the BP neural network algorithm, the modeling of the speech feature signal classification algorithm mainly includes the following steps: firstly, BP neural network construction, secondly BP neural network training, and finally BP neural network evaluation.

The construction of BP neural network has the nature of system input and output data, so the structure of BP neural network is determined accordingly. In this model, the input signal of the first-level indicators is 3-dimensional, and there are five types of evaluation indicators for the health status of national higher education, so the structure of the BP neural network used in this model is 3-5-1. It is expressed as: the input layer has 3 nodes, the hidden layer has 5 nodes, and the output layer has 1 node. When constructing the evaluation model of BP neural network, it is necessary to have a certain number of learning samples to establish an evaluation system. In order to minimize the evaluation error, this article considers the use of behavioral anchor quantification methods to define different horizontal scales (the following table), and construct a data set that meets the conditions as training samples.

Subsequently, the data from the United States and South Africa are substituted into the neural network for testing to verify whether it is consistent with the preset classification.

3.4 Improvement of genetic algorithm to neural network

In the repeated tests of the BP neural network, we found that the established BP neural network is not stable, and the convergence time is often too long, and the weights and thresholds of each test differ greatly. This shows that under the data, it is difficult to achieve good evaluation results with pure BP neural network. In fact, the determination of network weights and thresholds has always been the difficulty of network training. Randomly selected parameters will seriously affect the accuracy of the network. Therefore, consider adding genetic algorithms to the neural network, and use the optimal individuals trained by the genetic algorithm to optimize the weights and thresholds of the network. Specific steps are as follows:

Step 1: Individual coding and population initialization. The individual contains the ownership value and threshold of the entire neural network. This article uses real number coding to encode individuals. The encoding length is:

$$S = n \times m + m \times l + m + l \tag{11}$$

Among them, *m* is the number of hidden layer nodes; *n* is the number of input layer nodes; *l* is the number of output layer nodes.

The size of the population has a great impact on the global search performance of the genetic algorithm. The initial population size in this experiment is 20.

Step 2: Setting the fitness function. The fitness function is set as the reciprocal of the sum of squares of the neural network error:

$$f = \frac{1}{SE} \tag{12}$$

Among them, *SE* is the sum of squared errors between the predicted output of the neural network and the expected output.

Step 3: Selection of individuals. The selection of individuals can be based on probability values. The formula is as follows:

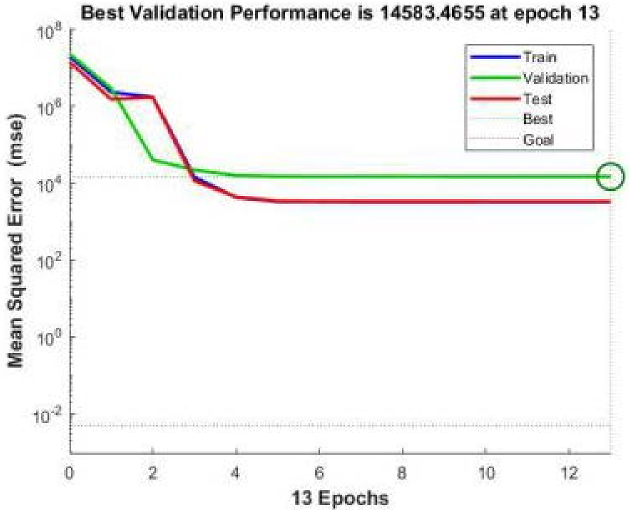
$$P_i = \frac{f_i}{\sum_{i=1}^{k} f_i} \tag{13}$$

Among them: *f* is the fitness value of the individual i; *k* is the number of individuals in the population.

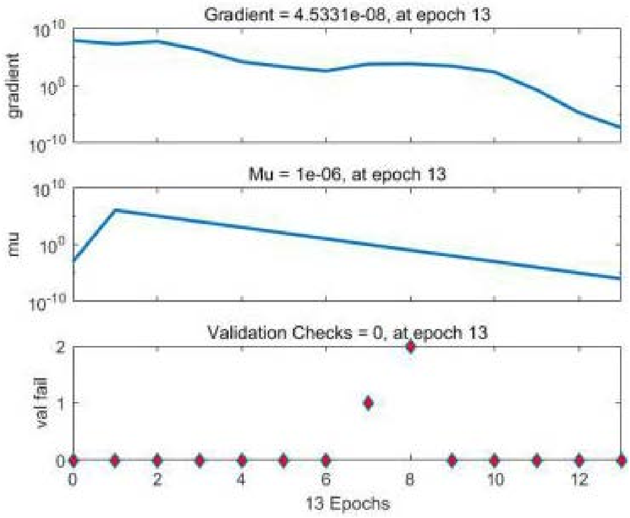
Step 4: Crossover operation and mutation operation. The optimal individual does not perform crossover operation, but directly copied to the next generation. For other individuals, mutation operations are performed with mutation probability Pm to produce another new individual. In this experiment, *Pc*=0.4，*Pm*=0.07, the evolutionary algebra is 100.

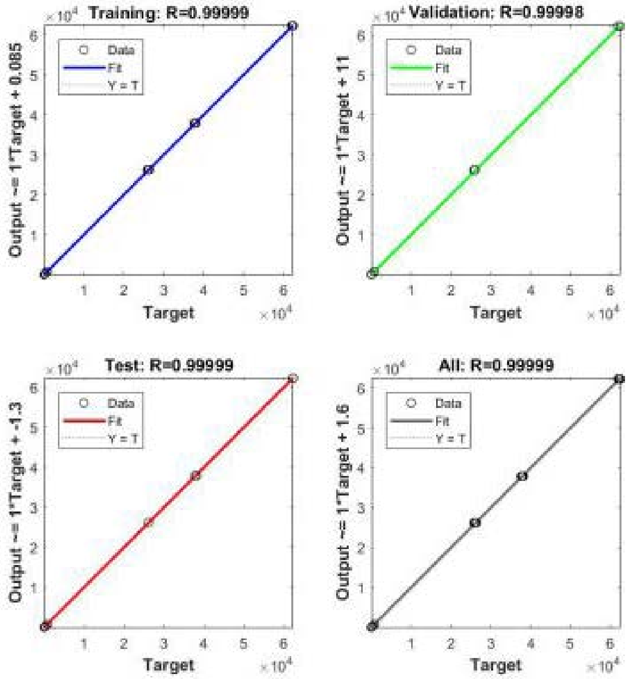
Step 5: Cycle operation.

GA-BP neural network training error is shown as figure 2.



Training Performance





Training State Regression

Figure 2. GA-BP neural network training error map

4. Result of the problem

According to our national higher education health evaluation model, we can get the specific ratings of six countries as table 3.

Table 3. Level of correlation of inputs

Country	United States	Germany	Japan	Australia	South Africa	India
Ratings	5	3	3	2	1	1

5. Model evaluation

5.1 Advantages of the model

1.The models and data in this article look at the global higher education situation from both horizontal and vertical dimensions.

2. In terms of empirical models and empirical methods, this article has absorbed advanced models and methods in related research fields around the world.

3.The combination of genetic algorithm and neural network has high classification accuracy, strong robustness, and fault tolerance to noise.

5.2 Insufficiency of the model

1.Since the past year, the world has been affected by the new crown epidemic, and the economies of all countries are in a haze [5]. The first promotion of online higher education and immature response methods have caused a lot of noise in last year's data.

2.Due to the lack of data on higher education in specific countries and specific majors, analysis and empirical research cannot be conducted for specific majors.

3.The error and convergence speed are very slow during network training.

6. Conclusion

Because the sustainability evaluation of national higher education is based on prediction and uses health evaluation model to evaluate, in fact, we only need to disturb the health evaluation model to analyze the sensitivity of the model.

We choose to use the Monte Carlo algorithm to disturb the initial data of each country, set the deviation of 1/3/5% respectively, and still use our previous health evaluation model to evaluate it. The results are 96.7%, 90% and 83% respectively.

References

[1] Jinyuan Ma, Yuzhuo Cai. Innovations in a higher education system: the role of embedded agency [J]. Higher Education, 2021.

[2] YunTao Pan. Research on The Quality of Graduate Education from The Perspective of Whole Process Quality Management[J]. International Journal of Social Sciences in Universities, 2020, 3 (4).

[3] Blerim Saqipi, Xhavit Rexhaj. Moving from survival to development: Experiences in Kosovo from developing academic identity in a transitional context[J]. European Journal of Education, 2020, 55 (4).

[4] Song Gao, Juan Huang, Yaru Li, Guiyan Liu, Fan Bi, Zhipeng Bai. A forecasting model for wave heights based on a long short-term memory neural network [J]. Acta Oceanologica Sinica, 2021.

[5] Nuha Hassan Elm basher Eltayib, Nizar Khalil Abubaker Suliman. The Impact of COVID-19 on the Sudanese Higher Education System [J]. Journal of Educational Research and Policies, 2020, 2 (11).