Measuring the Higher Education System

Kexin Luan, Qing Sun, Zhenxin Song

Yantai Research Institute of China Agricultural University, Shandong Yantai 264670

Keywords: Entropy weight method, Evaluation system, BP neural network.

Abstract: On the basis of collecting abundant indicators of the health of the higher education system, 14 indicators were selected using principal component analysis (PCA) to evaluate the health of the system of higher education. The entropy weight method (EWM) is used to obtain the weight vector of the system, and the indicators are standardized to calculate the score [1]. At the same time, the standard of system health is obtained through the K-means algorithm. In the end, an evaluation system including higher education background, higher education scale, higher education quality, higher education investment and higher education contribution is established. Considering the differences in the national conditions of various countries, and also in order to test the adaptability of the evaluation model, construct BP neural network for comprehensive evaluation, define the level scale division of different indicators, and finally obtain the ranking of the 8 sample countries through training samples.

1. Introduction

Higher education refers to all post-secondary education, including public and private universities, colleges, technical training colleges and vocational schools. Higher education not only benefits individuals, but also benefits society. It has played a role in promoting economic growth, reducing poverty, and improving social equity. However, Challenges still remain. Despite the large number of graduates from higher education, many still do not have the relevant skills needed to successfully integrate into the labor market. At the same time, many countries with limited resources are striving to meet the growing needs of larger student groups without compromising the quality of their educational products. Moreover, different regions have different problems, or the severity of the problems is also different. Therefore, it is necessary to construct the same evaluation system for different regions in order for government to make targeted policies.

2. Model principle

a. Collected a large amount of data to construct a comprehensive selection of higher education system indicators from five aspects: HEQ, HES, HEC, HEI, and HEB. The entropy method is used to calculate the weight of the indicator, and the result obtained is more objective and fair.

b. Use the BP neural network to test the effectiveness of the evaluation model, define the level scale division of different indicators, and finally obtain the ranking of the 8 sample countries through the training samples.

c. In essence, BP neural network is particularly suitable for solving problems with too many elements and complex mechanisms. BP network has a certain ability to promote. Once the learning is completed, the BP network will deal with the problem with higher accuracy.

d. The prediction accuracy of the BP neural network is affected by the number of hidden layer nodes. In order to control the training time and number of times and prevent the model from overfitting, we will use the hidden layer. The number of nodes is controlled at 5. This reduces the prediction error and enhances the robustness of the model.

e. Since the background, investment, scale, quality and contribution of higher education will dynamically change with the implementation of policies, there is a sustainable cycle. After the policy is implemented, the health status of higher education will be evaluated. Therefore, the establishment

  DOI: 10.25236/iemb.2021.006

of three-level indicators in the establishment of the model has enhanced the authenticity and accuracy of the model.

3. Model establishment

3.1 Data Indicators

All data in this study come from statistical reports or public databases of international authoritative organizations. The indicators includes GDP per capita, Higher education expenditure as a percentage of government expenditure, Higher education enrollment rate, Number of R&D technicians per 1 million people, Workforce with higher education.

3.2 Higher Education System Model

In order to make all indicators equal, it needs to normalize the data because these indicators have a wide range of values. If we do not deal with their value, then the results of our model may be dominated by one of these indicators [2]. When analyzing all indicators, we found that they can be divided into three types. After the data normalization, the value range is between -1 and 2.

For the positive index (the larger the value, the better), the calculation formula for data normalization is:

$$X_i = \frac{X_i - X_{min}}{X_{max} - X_{min}}$$

For negative indicators (the smaller the value, the better), the calculation formula for dimensionless processing is:

$$X_i = \frac{X_{max} - X_i}{X_{max} - X_{min}}$$

For intermediate indicators (the closer the value is to the fixed value, the better), the calculation formula for data normalization is:

$$x_{ij} = \begin{cases} 1 - \dfrac{s_1 - d_{ij}}{\max\left\{s_1 - d_j^{\min}, d_j^{\max} - s_2\right\}} \\ 1 - \dfrac{d_{ij} - s_2}{\max\left\{s_2 - d_j^{\min}, d_j^{\min} - s_2\right\}} \end{cases}$$

The basic idea of the entropy weight method is to determine the objective weight according to the variability of the index. Generally speaking, if the information entropy of a certain index is smaller, it indicates that the degree of variation of the index value is greater, the more information provided, the greater the role it can play in the comprehensive evaluation, and the greater its weight. After normalizing the indicators, calculate the proportion of the j-th country in the i-th indicator:

$$P_{ij} = \frac{X_{ij}}{\sum_{i=1}^{n} X_{ij}}$$

Calculate the entropy value of the i-th index:

$$e_i = -k \sum_{i=1}^{n} p_{ij} \ln\left(p_{ij}\right)$$

Calculate information entropy redundancy:

$$d_j = 1 - e_j$$

Calculate the weight of each indicator:

$$w_j = \frac{d_j}{\sum_{j=1}^{m} d_j}$$

According to the above steps, we can get the final weights. The weights of specific indicators based on the data are shown in Table 1.

Table.1. Second-class indicator System

Second-class indicator	Primary indicator and weight
Higher Education Contribution (26.22%)	Workforce with high education (0.23%)
	High-tech export (27.31%)
	Published in scientific (72.46%)
Higher Education Investment (2.87%)	Ratio of Internet users (88.15%)
	Ratio of higher education (11.85%)
Higher Education Scale (3.60%)	Enrollment rate (86.67%)
	Teacher student ratio (13.33%)
Higher Education Background (7.82%)	GDP per capita (83.25%)
	Proportion of urban population (16.75%)
Higher Education Quality (59.48%)	Unemployment ratio (7.80%)
	Proportion of academic woman (6.43%)
	Number of Nobel Prize winners (49.45%)
	Number of R&D technicians (25.17%)
	Exchange rate (11.15%)

We can see that the weights of Higher Education Quality (HEQ) and Higher Education Contribution (HEC) are very high, which is very close to our expectations. HEQ and HEC are important indicators to measure the health of higher education, so the weighted results are reliable. On the basis of the above indicator system and weight method, a comprehensive indicator to measure the health of the higher education system is proposed, and the weight method is used to evaluate the weight w of HEM [3]. Let HEM represent the health of the higher education system. The function of HEM can be expressed as:

$$HEM = \omega_1 * HEC + \omega_2 * HEI + \omega_3 * HES + \omega_4 * HEB + \omega_5 * HEQ$$

The specific indicator weights are as follows:

$$w = [0.2622\ 0.0287\ 0.0360\ 0.0782\ 0.5948]$$

Through the above analysis, a three-level higher education system health evaluation model has been established. The three levels of evaluation indicators are shown in Figure 1.

The sustainability of higher education changes to a certain extent according to the change of initial indicators, so the time series model is disturbed to a certain extent. Set a deviation of 0.4% and 0.6% for per capita GDP. After forecasting through the time series model, the final scores are 95.6% and 88.3% similar to the original results. Therefore, it can be believe that the model which has very strong stability can assess the health status of higher education in the country.

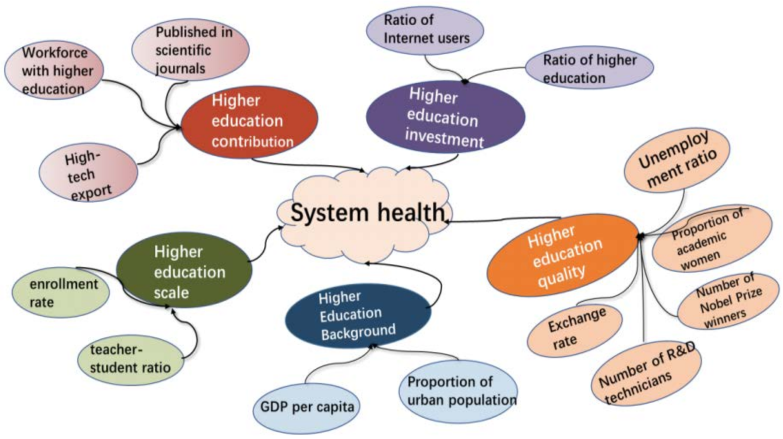


Figure 1. Three-level indicator system diagram

3.3 BP neural network model

The BP neural network itself has non-linear characteristics. It only needs to be trained with samples of special problems. It can reflect the characteristics of this type of problem in the weights of the interconnections between neurons, and the output can give the result of the to-be-solved problem. If the output layer cannot get the desired result, the error is back-propagated, so that the output result of the BP neural network is constantly approaching the expected output.

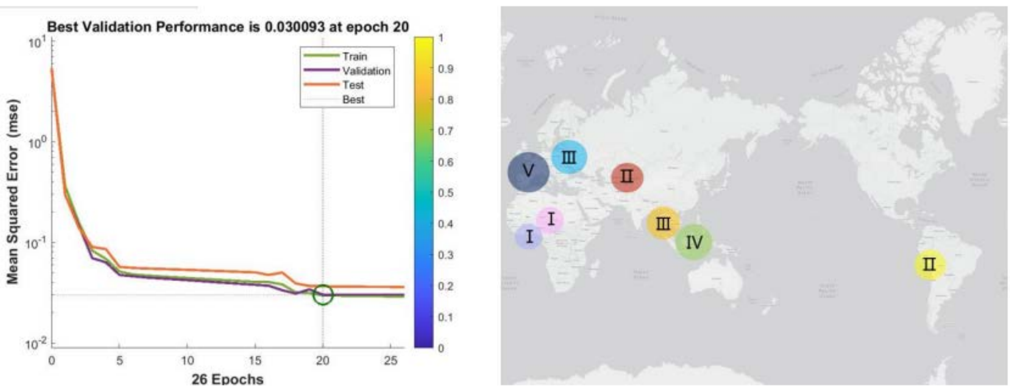
Considering that countries have huge differences in historical evolution, human geography, economic development, political conditions, and social customs. This part uses BP neural network model to test the applicability of the constructed model and its index system. Higher education health evaluation has 5 first-level evaluation indicators. The 5 first-level indicators in the model are divided into 5 grades to make a standard set, and the level scale of different indicators is defined [4].

The prediction accuracy of the BP neural network is affected by the number of hidden layer nodes. In order to control the training time and number of times and prevent the model from overfitting, it will use the hidden layer. The number of nodes is controlled at 5. This reduces the prediction error and enhances the robustness of the model.

There are 5 nodes in the input layer. Among them, C1 represents Higher education contribution; C2 represents Higher education investment; C3 represents Higher education scale; C4 represents Higher education back-ground; C5 represents Higher education quality. The output of the network, which is the result of higher education health evaluation, is a single output. Therefore, the number of output layer nodes is one. On the basis of summarizing a large number of network structure designs, through own experimental research, the number of hidden layer neurons is 5. It can be seen from the figure that the network converges after 26 iterations and reaches the specified error accuracy. Substituting German data into the test, the health status of the German higher education system is rated as V, and the result is consistent with the preset conclusion, which shows that the model has strong applicability. After substituting the data of eight countries into the neural network model, a more accurate distribution map of the health status of higher education was obtained.

The neural network construction of this model is a multi-layer feedforward neural network. In order to verify the sensitivity of the model, the data of the national higher education health indicators are disturbed to a certain extent, and the output results are analyzed again through the neural network model for verification. Deviations of 0.3% and 0.5% are set for the enrollment rate of higher education in various countries. After the initial data is brought into the neural network model, the final grade is the same as the original grade. It has been verified that under the disturbance of a certain range of indicator data, the higher education health status evaluation model has strong stability.

In essence, BP neural network is particularly suitable for solving problems with too many elements and complex mechanisms. BP network has a certain ability to promote. Once the learning is completed, the BP network will deal with the problem with higher accuracy.



(a)Network converges (b)Map of the health status

Figure 2. Error analysis and result graph

4. Model results

It is necessary to choose a suitable standard to evaluate the health of the higher education system based on the model results. Therefore, use KA (K-means clustering algorithm) to develop a reasonable standard. In this algorithm, the number of clusters is determined to be 2 and the maximum number of iterations is set to 10. SPSS is used to continuously loop the cluster centers, and finally the clustering results are output.

This section first clusters the data of 8 cities and calculates the two cluster centers of each indicator. Then use the mean value of the indicator center as the standard boundary. Use histograms to describe the standards of each indicator, and qualitatively evaluate the two types of indicators through intervals. Use "normal", "good", "excellent" and other qualitative indicators to measure the health of the higher education system. "Normal" means that the higher education system is unhealthy. "Good" means that the higher education system is healthier. "Excellent" is used to describe the health of the higher education system. Combining five secondary indicators, a radar chart is used to describe the overall health of the higher education system. The standard of system health is shown in Figure3.

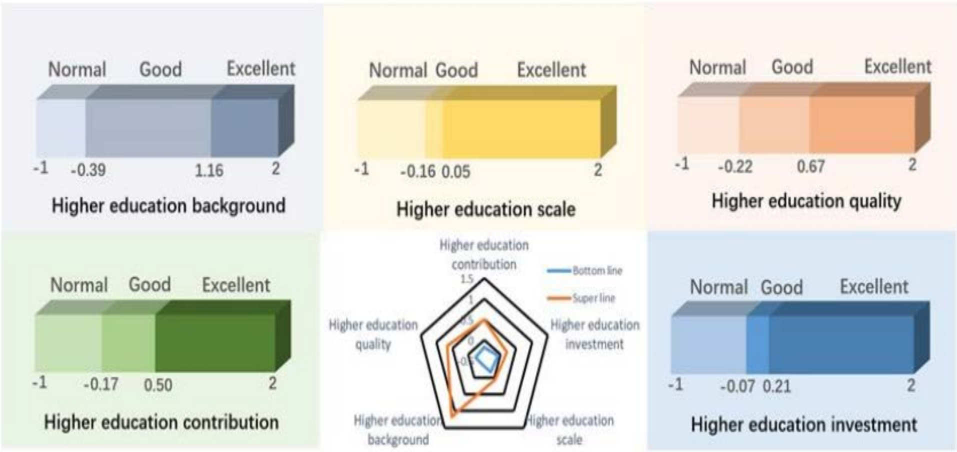


Figure 3. The Standard of System Health

According to the standards of the national higher education system established in 3.4, HEB is unhealthy in the range (-1, -0.39), and Togo is -0.95, so the country has more room for improvement in terms of higher education background. HES was rated as unhealthy in the interval (-1.-0.16), while Togo was rated as - 0.9234. HEQ is unhealthy in the range (-1, -0.22), and Togo is -0.57. Compared with other aspects, Togo is more optimistic in this aspect. HEC is rated as un- healthy in the interval (-1, -0.17), and Togo is -0.61. HEI is unhealthy in the range (-1, -0.07), while Togo is -1.29. Togo's investment in higher education is seriously insufficient, which restricts the development of higher education in this country. Since HEC and HEQ account for the largest proportion in the evaluation model have established, should focus on improvement in the next policy formulation.

References

[1] Ye Fugui, Ma Ying. From a world-class university to a world-class higher education system [J]. China Higher Education, 2013 (02): 61-63.

[2] Gu Yuanyuan Can China’s education by studying abroad reduce the poverty of the source countries? An empirical study based on the countries along the "Belt and Road" [J]. Educational Research, 2019, 40 (11): 124-142.

[3] Wang Zhengqing, Wang Yin, Sun Xinyan. Research on the Measurement of the Competitiveness of Higher Education and Related Factors in Countries along the "Belt and Road" [J]. Journal of Southwest University (Social Science Edition), 2021, 47 (01): 112 - 123 + 227.

[4] Bound John, Braga Breno, Khanna Gaurav, Turner Sarah. The Globalization of Post- secondary Education: The Role of International Students in the US Higher Education System [J]. Journal of Economic Perspectives, 2021, 35 (1).