ISSN 2663-8169 Vol. 3, Issue 3: 27-32, DOI: 10.25236/IJNDE.2021.030307

Based on BP neural network method to establish higher education health evaluation system

Menghui Chu, Zekai Wang, Zirui Zhao

Electrical Information Department, Shandong University of Science and Technology, Jinan, Shandong, 250031, China

Abstract: *This paper is here to develop a model that can assess the health status of the higher education system in any country. Then select a country and propose a set of policies that will move that country from its current state to its target state. Five indicators for the evaluation of the education system are formulated. Then, Python crawler is used to capture the data of different countries under these five indicators from the Internet. Then, K-means++ clustering method is used to classify countries into three classes. The significance and specific results of classification are shown in 4.2.1 and Table3. Then, according to different types of countries, TOPSIS based on entropy weight method is used to establish a system that can evaluate the health status of higher education in any country.*

Keywords: *Python crawler, K-means++ clustering, TOPSIS, PCA, BP neural network, Fuzzy comprehensive evaluation*

1. Background

Higher education is of great importance to every country. From Germany to the United States, from Japan to Australia, we see different ways of higher education in different countries. Each country not only trains its own students but also attracts a large number of international students every year. The diversity of students and higher education means the exchange and integration of different cultures around the world. In such an ecological environment, it is particularly important to establish a health assessment standard for higher education. Suggestions for countries with low health level of higher education and implementation are of great help to the improvement of a country's comprehensive strength.

There are two main tasks that are required of us in this question. Task 1: Develop a model that can assess the health of the higher education system in any country. Task 2: Select a country based on a series of analyses, determine a healthy and sustainable state (target state), and come up with a set of policies that can move the country from its current state to its ideal state. It will need to collect the data of several indicators of the evaluation system in different countries by using crawler and other methods, and then pre-process the data. Then it can classify countries by K-means clustering according to the collected data, and develop a system that can evaluate the health status of higher education in any country by using TOPSIS evaluation model based on entropy weight method according to different types of countries. In this way, the country types will then be judged and can then be evaluated under the respective categories.

2. Case study of Vietnam

The third and fourth question is analyzed from the actual perspective of Vietnam

2.1 Dimensional reduction of indicators by principal component analysis

Higher education is often affected by multiple indicators, which are often accompanied by noise and unimportant indicators, thus affecting the data analysis process. In order to make effective recommendations for the selected countries, the data dimension should be reduced here to present a better vision

2.1.1 Correlation test

KMO and Bartlett test were used to test the correlation degree of the original 5 indexes with SPSS

ISSN 2663-8169 Vol. 3, Issue 3: 27-32, DOI: 10.25236/IJNDE.2021.030307

Table 1: KMO and Bartlett tests

KMO sampling fitness measure		.628
Bartlett's test for sphericity	The approximate chi-square	140.09
	Degrees of freedom	10
	significant	.000

According to the KMO sampling fitness measure 0.628>0.5, it can be seen that the original 5 indicators have a high degree of correlation, which meets the premise of using principal component analysis. According to the 96 samples obtained by Python crawler and 5 indicators, a sample original matrix with the size of 96×5 can be constituted. Due to the large number, it will not be listed here . Using the sample correlation coefficient matrix formula:

$$R = \frac{\sum_{k=1}^{n}(x_{ki}-\bar{x}_i)(x_{kj}-\bar{x}_j)}{\sqrt{\sum_{k=1}^{n}(x_{ki}-\bar{x}_i)^2(x_{kj}-\bar{x}_j)^2}}$$

Table 2: Correlation coefficient matrix

	x1	x2	x3	x4	x5
X1	1.000	0.403	-0.270	-0.128	0.822
X2	0.403	1.000	-0.267	-0.294	0.359
X3	-0.270	-0.267	1.000	0.034	-0.263
X4	-0.128	-0.294	0.034	1.000	-0.103
X5	0.822	0.359	-0.263	-0.103	1.000

2.1.2 Dimensional reduction was performed on the index

The eigenvalues and eigenvectors of the correlation coefficient matrix R were calculated using MATLAB

$$\text{contribution} = \frac{\lambda_i}{\sum_{k=1}^{p}\lambda_k} \quad (i = 1,2 \ldots p)$$

$$\text{Cumulative contribution rate} = \frac{\sum_{k=1}^{i}\lambda_k}{\sum_{k=1}^{p}\lambda_k} \quad (i = 1,2 \ldots p)$$

After calculation, eigenvalues, corresponding eigenvectors and contribution rates of the correlation coefficient matrix are listed in the following table:

Table 3: Results after dimensionality reduction

	X1	X2	X3	X4	X5
X1	0.577	-0.229	0.297	0.111	-0.717
X2	0.444	0.358	-0.179	-0.800	0.045
X3	-0.328	0.159	0.879	-0.307	0.002
X4	-0.209	-0.849	-0.091	-0.476	-0.009
X5	0.564	-0.270	0.314	0.162	0.696
The eigenvalue	2.314	1.050	0.866	0.594	0.176
contribution	0.463	0.210	0.173	0.119	0.035
Cumulative contribution rate	0.463	0.673	0.846	0.965	1.000

As can be seen from the table, the cumulative contribution rate of the first three principal components is 84.6%. Therefore, it can be considered to take only the first three principal components to effectively summarize the original variables. The three main levels are divided by reference [6].

The first principal component F1 has a moderate positive load on X1, X2, and X5, and an equal negative load on X3 and X4. And mainly focused on the social support for education, so it is called social services. The second principal component, F2, has a large positive charge at X1 and a very small negative charge at X4, hence the so-called brain drains. The third principal component, F3, has a large positive load in X3 and a large positive load in X5. X1, X2 and X4 are all small, which can be called talent training.

ISSN 2663-8169 Vol. 3, Issue 3: 27-32, DOI: 10.25236/IJNDE.2021.030307

2.2 Establishment of fuzzy comprehensive evaluation blueprint

Based on the three indexes that have the greatest influence on national higher education obtained by 5.1 master hierarchy analysis, the paper reevaluates the national economy of class B according to the fuzzy comprehensive evaluation. This is no longer A part of the score because the education system in A country has improved. Based on the final results, the country ranked first in the evaluation system was selected as the vision for the future development of Vietnam's higher education system

2.2.1 Establishment of fuzzy comprehensive evaluation model

The relationship between the three principal components and the original index obtained by dimensionality reduction according to principal component analysis,

$$F1=0.577X1+0.444X2-0.328X3-0.209X4+0.564X5 \quad (1)$$

$$F2=-0.229X1+0.358X2+0.159X3-0.849X4-0.270X5 \quad (2)$$

$$F3=0.297X1-0.179X2+0.879X3-0.091X4+0.314X5 \quad (3)$$

According to the above relation, we processed the national data of category B in the fourth question, and obtained the data sets with the three principal components as indicators. Then the index weight was determined by the coefficient of variation method i Calculate the mean of the i index

$$\bar{x}_i = \frac{1}{n}\sum_{j=1}^{n} a_{ij}$$

$$s_i^2 = \frac{1}{n-1}\sum_{j=1}^{n}(a_{ij} - \bar{x}_i)^2$$

ii $vi=si/|\bar{x}i|$,normalized weight coefficientv_i

$$w_{i=} v_i / \sum v_i$$

The fuzzy synthesis matrix is constructed

$$R = (r_{ij})_{m*n} = \begin{pmatrix} r_{11} & \cdots & r_{1n} \\ \vdots & \ddots & \vdots \\ r_{m1} & \cdots & r_{mn} \end{pmatrix}$$

The quantitative value of higher education level can be obtained by choosing appropriate fuzzy synthesis operator and using weighted average algorithm to carry out fuzzy synthesis operation. Fuzzy synthesis operator:

$$M(\cdot,+): b_j = \sum_{i=1}^{m}(a_i \cdot r_{ij})$$

Sort the obtained quantized values:

Table 4: Ranking of quantitative indicators

	Social services	The brain drains	Develop talent	Fuzzy comprehensive evaluation	ranking
China	14.6584548	40.7450381	5.336175058	0.032925922	1
Vietnam	16.4359726	6.653362043	8.407507644	0.013128127	40

According to the obtained data, China ranks the first in the fuzzy comprehensive evaluation, while Vietnam ranks the 40th. Combined with the three indicators, this paper will give relevant suggestions based on China as the development vision of Vietnam's higher education system.

2.2.2 Vision for higher education development

Higher education is an important position to cultivate high-quality talents and promote technological development and innovation. In recent years, China, as a new generation country, has developed to different degrees in various fields. According to the reference [7], it is not difficult to see that behind the rapid development of China, one of the main reasons lies in the development of higher education. Compared with countries with A-level education in Europe and the United States, that is, countries with developed education, it is difficult for Vietnam to follow suit because such countries have A relatively complete education system and corresponding supporting construction of educational facilities. However, China, as a country with B-level education, is also a developing country. Due to the imperfection of various indicators, Vietnam can learn more from the development of higher education. Therefore, China is regarded as an ideal system for Vietnam's current development.

ISSN 2663-8169 Vol. 3, Issue 3: 27-32, DOI: 10.25236/IJNDE.2021.030307

As can be seen from the figure, there is a gap between Vietnam and China in various indicators, especially in the aspect of brain drain. With the development of market economy and the fierce competition for talents, Vietnam lacks a complete industrial system and higher education, leading to a large number of talents flowing overseas. Combined with the news reports in recent years, it is not difficult to see that Vietnam's response measures to the largescale brain drain need to be improved. Combined with the reference [8], it is not difficult to see that China has taken corresponding measures to deal with large-scale talents, such as strengthening the development of human resources, promoting the improvement of the education system, and improving the construction of the environment for talent development. There are many similarities between Vietnam and China's political system, so Vietnam has a lot of space to learn from China's corresponding measures to reduce the number of brains drain. In terms of personnel training and social services, the gap between Vietnam and China is not very big. Vietnam should learn from China's excellent places and continue to develop its own advantages to promote the better development of higher education. The best way to deal with Vietnam's current higher education is to learn from China's advanced higher education level. On this basis, when the external factors reach a high level, we can learn from the advanced higher education experience of the United States, Britain, and Japan and so on to promote our own development.

2.3 A measure of the current and desired system health

And model, based on the results of the three blueprint is presented based on the problems and literature collection, sorting and five indicators of national higher education popularity, overseas students to share, the proportion of the total, total national education and higher education enrollment, the proportion of people receiving higher education, we put the

Vietnam's five index data as the health of the current system, based on five indicators data in China, we put forward the healthy and sustainable health system.

First of all, in order to look at the health status of the current system and the ideal system from a macro perspective, we put the data of Category B countries into the model of Question 1 to solve the problem and make a comprehensive score. In order to objectively reflect the difference between the current state and the ideal state, we processed the scores of Category B countries and found that the comprehensive score of the ideal system (China) ranked first with a score of 0.036, indicating that the health status of the ideal system was relatively excellent. The current system (Vietnam) ranked in the lower middle with a score of 0.011, indicating that the current system is in average health and there is still much room for improvement. After the evaluation, in order to observe the differences between the current system and the ideal system in all aspects and make plans of different strength, we compared the data of the current system and the ideal system under the five indicators, as shown in the figure below:

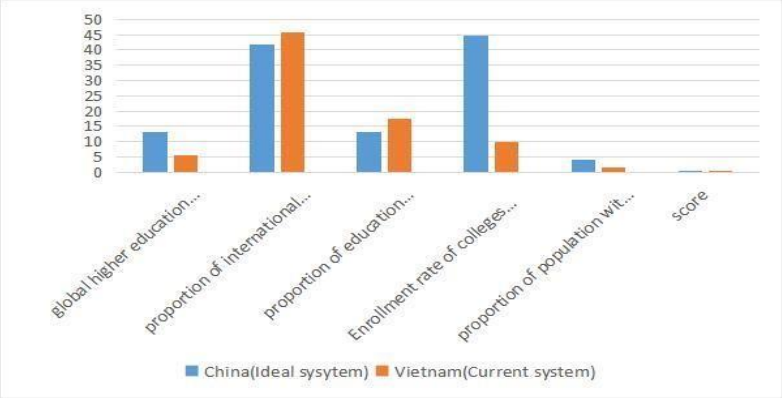

Figure 1: Comparison of five indicators between China and Vietnam

3. Sensitivity Analysis

In the health evaluation model of higher education, there are four indicators: the penetration rate of higher education in each country, the share of overseas students, the proportion of education expenditure in each country, the enrollment rate of higher education, the proportion of the population receiving higher education and other variables. Through sensitivity analysis of these variables, their initial values (taking Vietnam as an example) are the penetration rate of higher education in each country = 5.37118%, the

ISSN 2663-8169 Vol. 3, Issue 3: 27-32, DOI: 10.25236/IJNDE.2021.030307

share of overseas students =45.83744%, the proportion of education expenditure in each country =17.58658%, the enrollment rate of higher education =9.961576%, and the proportion of the population receiving higher education =1.48111%. According to the principal component analysis in the third question, the first index contributes the most, so only the first index is changed here.

Let the higher education penetration rate fluctuate continuously by 5%, and use MATLAB to calculate the corresponding comprehensive score (the score here is not normalized), and draw the relationship between the two as follows:

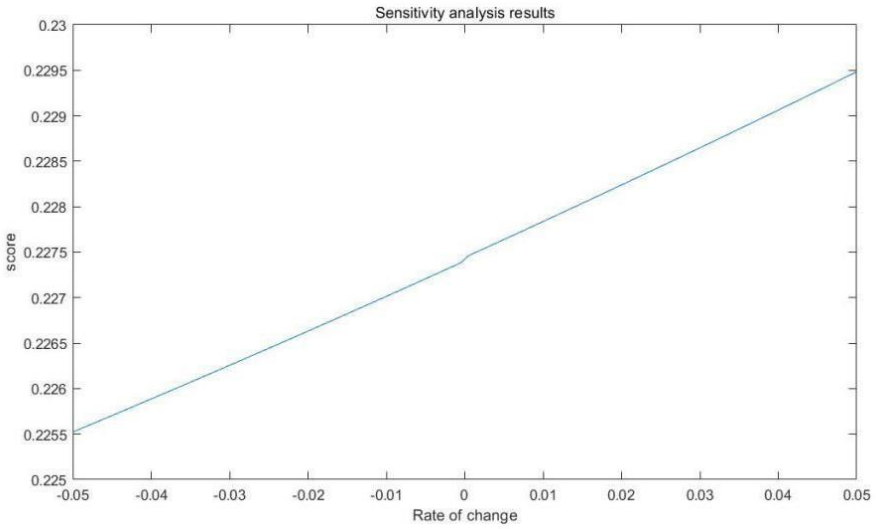

Figure 2: Sensitivity analysis results

From figure 1 it can be seen that with the fluctuation of the higher education penetration rate, the score basically tends to increase in a straight line, so the model is more sensitive to the higher education penetration rate of different countries. According to this analysis, we should pay special attention to this index when implementing policies, and try our best to ensure the continuous growth of the penetration rate of higher education in all countries, and take corresponding measures to prevent its decline.

4. Conclusion

The comprehensive evaluation of each country according to the health evaluation system of higher education found that the comprehensive strength of a country can not completely replace its health level of higher education. After the dimensionality reduction of the influence index, it can be seen that the penetration rate of higher education has a great impact on the health status of a country's higher education. The formulation of policies requires the prediction of the results to ensure the high efficiency of the implementation of policies. Policy implementation also needs a long time, adhere to the implementation of the corresponding policies can be relatively accelerated to reach the target state. Higher education is closely related to each of us, and we should actively respond to the policy of improving the health level of higher education formulated by the country.

References

[1] Liang Qian. Three challenges of quality assurance in higher education and their solutions [J]. Educational Research, Tsinghua University, 2020, 41(01): 142-148.
[2] Li Liguo, Yi Peng, Xue Xinlong. Overcoming the middle income trap re-quires increasing education investment-----Characteristics and enlightenment of national education investment in different stages of economic development [J]. Higher Education Research in China, 2016, (09): 55-62. DOI: 10.16298/ j.cnki.1004-3667.2016.09.10
[3] NGUYEN DANG THUY THUY. Research on Financial Management of Basic Education in Quang Ngai Province, Vietnam [D]. Central China Nor-mal University, 2018. [7].
[4]Haisheng Pan, Xing Weng. Research on the Coupling Relationship be-tween Higher Vocational Education and Economic and Social Development in China——Panel data of 31 provinces from 2006 to 2018 [J/OL]. Management of HigheEucation: 112 [2021-02-07 16:30].

ISSN 2663-8169 Vol. 3, Issue 3: 27-32, DOI: 10.25236/IJNDE.2021.030307

https://doi.org/10.13316/j.cnki.jhem.20210206.002.DOI:10.13316/j.cnki.jhem.20210206. 002.
[5] Sun Yanbo. How to Effectively Avoid the Risk of Brain Drench in China under the Background of Globalization [J]. Business, 2012, (06): 150-151.