The Application of Cloud Computing and Big Data Correlation Analysis to the Intelligent Evaluation Modeling

Yunxuan Xiao, Leize Yang*

Computer Science Beijing JiaoTong University Beijing, China

*Corresponding author: 19726072@bjtu.edu.cn

Abstract. This passage gave an analyse big data collected from UNESCO and report from governments then give health index to countries and focus on Brazil to give some possible solutions to the problem. First, we aim to develop EHSIS model to assess the health of any national higher education system. We define the weight of the indexes by subjective weighting and objective weighting by using Analytic Hierarchy Process and Principal Component Analysis (AHP) to compare the importance of different weight and Entropy Method to get the weight from comparing the degree of dispersion of indicators. Then combine the weight and the normalized data to get the model to give scores to sample countries relying on the eight indexes. Then, we screened out several indicators with greater influence through Principal Component Analysis and correlation analysis of factors and extracted the recommended adjustable indicators to 3. We found that the proportion of foreign students has the greatest impact on the health and sustainability of higher education in the country, and it has a high correlation with QS excellent schools and UN News excellent schools, which are difficult to control. Other indicators that have a greater impact include the teacher-student ratio, employment rate, etc. Last, we forecast the changes of indexes without the policy by Grey prediction and use the memory-less of Markov Chain to predict the change after introducing new policies. In order to evaluate the effectiveness of our policy, we utilize EHSIS model to give the score and compare the score with and without the policy and find a significant increase in score. What’s more, we discuss the real-world influence and the hardness of achieving the goal.

1. Introduction

Nowadays, higher education in the world becoming more and more important. Different countries have different higher education systems due to their separate national systems and their own conditions. However, the level of higher Education Level varies. Our paper starts from 8 factors analyze the health index from different countries and proposes improving policies to countries with low levels of higher education.

2. Assumptions and symbol representation

2.1. Model Assumptions
Only consider that the higher education is affected only by the proportion of foreign students, the proportion of top 500 universities in QS, USNEWS rankings, the employment rate, and the teacher-student ratio, number of journals, education expenditure as a proportion of GDP. Assume that the policy we introduce in Brazil only get effect from the main factors the proportion of foreign students, the proportion of top 500 universities in QS, USNEWS rankings, the employment rate, and the teacher-student ratio. Assuming that the policy implemented in Brazil makes the probability of indicator growth greater than the probability of reducing the indicator, and the probability of growth predicted by the Markov chain is used to approximate the growth rate. Maintaining the Integrity of the Specifications

2.2. Explanation of the mathematical symbols

\( x_i, i = 1, \ldots, 8 \), each of the number represents the 8 indexes.
\( \sigma_i, i = 1, \ldots, 8 \), each of them represents the weights of the index.

- \( i=1 \): proportion of foreign students among all the students,
- \( i=2 \): the proportion of top 500 universities in QS Rankings,
- \( i=3 \): the proportion of top 500 universities in QS Rankings,
- \( i=4 \): the employment rate among all the college students
- \( i=5 \): the teacher-student ratio in colleges and universities,
- \( i=6 \): amount of journals from colleges and universities,
- \( i=7 \): Education expenditure as a proportion of GDP.
- \( i=8 \): University density: per 1000 square in each country.

\( CI = \frac{\text{max}(A)-n}{n-1} \), Consistency index used in consistency test.
\( CR = \frac{CI}{R^2} \), Consistency Ratio which is used in consistency test.
\( Z = \frac{Z_i-\mu}{\sigma} \), Z is the standardized data. \( \mu \) represents average number. \( \sigma \) Represents the standard deviation.

3. Education Health and Sustainability Index System

3.1. Purpose of model

Develop an assessing evaluation system that can be used in evaluating any country’s higher education health and sustainability indexes. Model can be applied to different various countries. Final income index can represent the health and sustainability of the country’s higher education system.

3.2. Data Collection and Processing

Define we collect data from the database and we have confirmed their accuracy. In order to compare the indicators more fairly, we processed these data like USNEWS ranking top 500 universities ratio of all the university. For example, considering the different sizes of different countries, we use percentage data, but not amount or number. In different application scenarios, we standardize and normalize the data to eliminate the dimensional influence between different indicators. We use box plots in SPSS to remove abnormal data and supplement them with interpolation.

3.3. Model Develop

In order to analyze the health and sustainability of higher education of different countries, we build Education Health and Sustainability Index System (EHSIS). We collect 8 indicators for each of 24 separate countries. Through the Analytical Hierarchy Process (AHP), we define 2 decision layers (Figure-1) that contain eight sets of weight when evaluating the health level of national education. Starting from the first criteria, we structure comparison matrix by comparison method, we get the weight of each factor. \( CI = \frac{\text{max}(A)-n}{n-1} \), \( CR = \frac{CI}{R^2} \). For every layer, CR<0.1, so each layer passed the consistency test.
Figure 1. The basic architecture of EHSIS.

But here is the question that AHP is subjective and may not give a precise weight for EHSIS. To solve this problem, we introduce Entropy Method. First, we calculate the proportion of the i-th element under the j-th index.

$$P_{ij} = \frac{x_{ij}}{\sum_{i=1}^{n} x_{ij}}, \quad i = 1, \ldots, n, j = 1, \ldots, m$$

(1)

Second, we calculate the entropy of the j-th indicator

$$e_j = -k \sum_{i=1}^{n} p_{ij} \ln(p_{ij}), \quad j = 1, \ldots, m$$

(2)

Using (1) and (2), we obtain that \( k = 1/\ln(n) > 0 \), satisfy \( e_j \geq 0 \), then, we calculate the information entropy redundancy

$$d_j = 1 - e_j, \quad j = 1, \ldots, m$$

(3)

Finally, we calculate the weight of each indicator

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

(4)

By the statistics we collected, we get another weight by Entropy Method and we combine the two types of weights collected from two methods and multiply it with the normalized data. Then we can get the EHSIS, the factor weights are as following graph:
4. Evaluation of Health and Sustainability of Sample

Apply the theoretical assessing system to sample countries. According to the EHSIS, get the health and sustainability indexes of sample countries. Choose a country with a higher education system that has room for improvement.

4.1. Health and Sustainability Indexes of Sample Countries

We collect the data and visualize it as shown in Figure-3. First, we get all the data normalized by

\[ y_i = \frac{x_i - \min_{1 \leq j \leq n} [x_j]}{\max_{1 \leq j \leq n} [x_j] - \min_{1 \leq j \leq n} [x_j]} \]  

Then we put these data in EHSIS model and get the scores for each country as shown in Figure-4. And we analyze the law of data and divide these countries into three levels. According to the final indexes, we classify all countries into 3 levels:
Figure 3. Data regularities of the three levels of sample countries.

When the index $< 0.25$, its C level, when $0.25 \leq \text{index} \leq 0.4$, its B level, index$>0.4$, it’s A level.

Good: Switzerland, New Zealand, Canada, Australia, Singapore, US, UK
Medium: China, Russia, Japan, Germany, France, Italy, Spain
Weak: India, Kenya, South Africa, South Korea, Thailand, Brazil, Malaysia, Mexico

Figure 4. The country scores calculated by EHSIS.

4.2. The Reason We Choose Brazil
First, in weak-level countries, the higher education system has a higher level of improvement. Brazil’s foreign students’ percentage is the lowest. According to EHSIS, for several items that have a relatively large impact, Brazil is far behind the world average level. Also, Brazil’s education statistics are open and transparent. So, we choose Brazil as the reform county.
5. Policies for Brazil to Adjust Existing Higher Education System

Analyze Brazil’s weakness in 8 indicators. Specify a targeted improvement plan.

5.1. Brazil’s National Education Conditions

From 2000 to 2010, education investment of the Brazilian Federal Government in the past ten years has not reached the average level of the member countries of the Organization for Economic Cooperation and Development. In Brazil, the standard of public education is far lower than private education, and most people choose private schools, which make it difficult to implement national education reforms. At the basic education stage, only about half (51.8%) of Brazilian citizens have completed high school education.

In addition, from a historical perspective, Brazil’s primary education is behind

5.2. Analyse indicator weights

By comparing Brazil’s data with the average of 24 countries to reflect the need to improve Brazilian higher education. Standardized $z = \frac{x - \mu}{\sigma}$.

| Indicator | Weight |
|-----------|--------|
| ESP/GDP   | 0.654548 |
| Foreign Students Percentage | -1.015603 |
| University density | -0.43323 |
| QS500 Percentage | -0.637952 |
| USNEWS500 Percentage | -0.688506 |
| Academic articles number | -0.35858 |
| TSR        | -1.165489 |
| Employment rate | 0.235617 |

Table 1. Indicator weights of Brazil's data.

To make efficient policy for higher education system, determining the indicators with greater impact is necessary. Firstly, standardized real data were used to obtain the component matrix and total variance with principal component analysis (PCA) method, 8 indicators are compressed to 3 principal components. Then calculate the coefficient of linear combination of each index in each principal component by using factor loading and principal component characteristic root, and comprehensively obtain the model coefficient of the score. Finally, the comprehensive model coefficients are calculated and normalized to determine the weight.
Table 2. Initial Eigenvalue and Extract the Sum of Squares of the Load.

| Ingredient | Total | Variance | Grand Total | Total | Variance | Grand Total |
|------------|-------|----------|-------------|-------|----------|-------------|
| 1          | 3.029 | 37.863   | 37.863      | 3.029 | 37.863   | 37.863      |
| 2          | 1.757 | 21.960   | 59.823      | 1.757 | 21.960   | 59.823      |
| 3          | 1.310 | 16.379   | 76.202      | 1.310 | 16.379   | 76.202      |
| 4          | 0.638 | 7.981    | 84.183      |       |          |             |
| 5          | 0.547 | 6.843    | 91.026      |       |          |             |
| 6          | 0.388 | 4.847    | 95.873      |       |          |             |
| 7          | 0.256 | 3.196    | 99.069      |       |          |             |
| 8          | 0.075 | 0.931    | 100.000     |       |          |             |

Table 3. Brazil's 8 indicators and their linear combination coefficients.

| Indicator               | Coefficient | Linear Combination Coefficient |
|-------------------------|-------------|--------------------------------|
| USNEWS500 percentage   | 0.260       | 18.148%                        |
| Foreign Students Percentage | 0.333     | 23.272%                        |
| QS500 Percentage        | 0.221       | 15.446%                        |
| Education Spending/GDP  | 0.063       | 4.427%                         |
| Employment Rate         | 0.209       | 14.609%                        |
| STR                     | 0.309       | 21.576%                        |
| University density      | 0.066       | 4.583%                         |
| Academic Articles Number| -0.03       | -2.062%                        |

5.3. Evaluation of Reformed Higher Education System

We have mentioned evaluation of Brazil's existing Higher Education in 5.2.

To simplify the process, we use GM (1, 1) to predict the data changes in the each of 5 indicators in the next ten years.

Figure 6. Future higher education trends in Brazil.
Use the memory lessness of Markov chains to predict the growth rate of each year over the previous year. Combine Markov chain model and Grey Prediction to predict the annual indicator data for the next 10 years. Fit the closely related QS500 College Percentage and USNEWS500 College percentage.

\[
Pr(X_{n+1} = x \mid X_n = y) = Pr(X_n = x \mid X_{n-1} = y)
\]

(7)

\[
Pr(X_n = x_n \mid X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \ldots, X_{n-m} = x_{n-m})
\]

(8)

According the increase rate of the foreign student’s number percentage, predict USNEWS500 college percentage and QS500 college percentage. Then take what Markov predicted as the expect, and give the year table of policies complementing.

6. Policies for Brazil to Adjust Existing Higher Education System
Propose a series of suitable policies based on the extracted indicators and the established model Customize a relatively feasible schedule

6.1. Specific Policy
Brazil's per capita income is weak and higher education in Brazil is mostly private. The state financially supports private universities. Providing funding for poor students to enter private universities. For
example, set up funds to offer scholarships for outstanding students. Improve the level of national public education and increase domestic public investment. Building more public universities to support students from low-income family. In addition, strengthen primary education. The Brazilian government provides preferential policies to introduce high-quality foreign universities and cooperate in running schools. The cooperation of Brazilian school and advanced foreign school can make it more attractive for international students. What’s more, introduction of foreign teachers may also lead to the growth of teacher-pupil ratio. Brazil has stronger skills vocational system. The government should keep developing vocational education as a main factor to attract international students. Give preferential policies for local companies and multinational companies such as tax reduction, lower prices for the use of land and investment, which may bring up the employment rate of college students as well as facilitate national development. Government should provide better treatment to school teachers. The low teacher-students ratio leads to low quality of higher education.

6.2. Year Time
2021, the Brazilian federal government issued a plan to strengthen foreign cooperation in running schools, and proposed reforms on teacher welfare benefits. Select the University of Sao Paulo and the Catholic University of Sao Paulo as the pilot universities (a public university, a private university) to conduct a pilot program of foreign cooperation in running schools, introduce foreign teachers, and begin qualification screening of applicants from the two universities, and college students with excellent academic performance can achieve scholarships. Providing subsidies to students with family difficulties. 2021, School of Elementary Education was established in the State of Sao Paulo.
2022, the Brazilian government made a policy of introducing foreign-funded enterprises and cooperation between enterprises and universities, with the University of Phnom Penh as a pilot university. 2022, the Brazilian government promulgated a series of tax reductions and investment recommendations for companies.
2023, scholarship and subsidy reform for two universities (University of Sao Paulo and Catholic University of Sao Paulo) will be completed.
2023, the Brazilian government initially resolved the reform of teacher welfare benefits in São Paulo State, Rio de Janeiro, etc., and began to implement it throughout the country.
2024, the Brazilian government successfully introduced companies in Rio de Janeiro and set up long-term cooperation with the Federal University of Rio de Janeiro.
2024, complete the introduction of foreign cooperation to pilot universities (University of Sao Paulo and Catholic University of Sao Paulo).
2024, the state of Sao Paulo will be deployed as a pilot state to reform the primary education system.
2025, Establish cooperation between many companies and universities in Rio de Janeiro.
2025, at the same time, starting from Rio de Janeiro, is piloting the establishment of strategic cooperation between vocational and technical colleges and enterprises.

7. Evaluation of Effectiveness, Influence on the World By Policy Complement
Use EHSIS give the health and sustainability index of the adjusted higher education system. Acknowledge and discuss the impact on the world during the policy transition period.

7.1. Evaluate the Effective of Policy Produced By EHSIS
According to the data of the ten years after the Markov chain forecast due to the policy change, the score is then based on the data of the gray prediction ten years. By EHSIS, and the data forecasted, we give the score of the health and sustainability index of the higher education system in ten years. Then we tested the validity of our model.

7.2. Influence on the World When Complementing Policies
It must be admitted that many indicators are difficult to achieve noticeable results through direct means, such as the two world rankings, but through other control methods, they can have a greater impact on
the national teaching environment. After the opening of cooperation in running schools, new teaching projects and school construction will make the distribution of universities more intensive, and the composition of Brazilian students will also change significantly. International students will gradually occupy a certain proportion. In the case of strengthening international academic exchanges, Brazil’s higher education the enthusiasm for educating students will gradually increase. At the same time, due to the participation of international students, excellent schools are discovered and promoted, and the ranking of some universities may be improved. As for teachers, because teachers’ salaries have been improved, more talents are willing to invest in the higher education industry. The impact of this is that a teacher will face fewer students, reducing the burden and increasing the quality of teaching. Due to the new school-enterprise cooperation policy, under the continuous increase in demand for higher-level talents, the employment rate of undergraduates is likely to reverse the downward trend during the transition period and eventually reach a sustainable level. At the same time, due to the indirect influence of policies, other difficult indexes (such as the proportion of QS excellent schools, the proportion of US news excellent schools) will also increase accordingly. Nevertheless, we cannot ignore the changes in the global situation. According to forecasts, other countries are equally focusing on the development of the education industry. Therefore, the health status of higher education in all countries after many years is difficult to predict.

References

[1] Bureau of Higher Education Statistics, https://www.hesa.ac.uk/data-and-analysis/students/whos-in-he
[2] Global Business Data Platform, https://www.statista.com/
[3] National Center for Educational Statistics, https://nces.ed.gov/programs/digest/current_tables.asp
[4] Oecd database, https://www.oecd-ilibrary.org/education/education-at-a-glance-2020_3c442221-en
[5] Parliament of Australia website, https://www.aph.gov.au/About_Parliament/Parliamentary_Departments
[6] QS University Ranking Official Website, https://www.qschina.cn/university-rankings/world-university-rankings/2021
[7] Simon. S, Equity, quality and relevance in higher education in Brazil, Anais da Academia Brasileira de Ciências, 2004, vol. 76, pp. 15-16.
[8] Unesco website, http://uis.unesco.org/country/DE