IJHSR

ijhighschoolresearch.org

RESEARCH ARTICLE

The Rationality Analysis and Prediction of THE World University Rankings

Yingao Gao

Mount Douglas Secondary School, 3970 Gordon Head Rd, Victoria, BC, Canada; gcliclig@gmail.com

ABSTRACT: In the perspective of the wide population of high-school teens, choosing a suitable university can help students take advantage of their strengths better, and ultimately affect their life journey. In general, the world university rankings analyze each post-secondary college through the aspects of their teaching, research, and the future employment of their students, which provides important information for students and guides their choices. However, different rankings have distinct focuses and different weights on each indicator, which may mislead students, and possibly cause confusion. Therefore, this paper takes THE Rankings as an example, utilizes statistical theories to analyze the rationality of performance indicators, and uses principal component analysis (PCA) to optimize the weights of each indicator, so the scores and rankings will accurately reflect the characteristics, strengths and weaknesses of each university. Also, the data of the top 200 universities in the past years is used to train a neural network model to predict the future rankings, providing a reference for senior high-school students.

KEYWORDS: Mathematics; Analysis; THE World University Rankings; Principal Component Analysis; Neural Network.

Introduction

THE Rankings is a world university ranking published by Times Higher Education, which judges the universities on 5 indicators, including Teaching, Research, Citations, International Outlook, and Industry Income. These 5 indicators are then split into 13 more precise indicators, as shown in Table 1. Until 2021, THE Rankings covered more than 1500 universities in 93 countries and territories.

Table 1: Indicators and weights of THE Rankings.

Indicators	Weights	Sub Indicators	Weights
Teaching	30%	Reputation survey	15%
		Staff-to-student ratio	4.5%
		Doctorate-to-bachelor's ratio	2.25%
		Doctorates-awarded-to-academic-staff ratio	6%
		Institutional income	2.25%
Research	30%	Reputation survey	18%
		Research income	6%
		Research productivity	6%
Citations	30%	Research Influence	30%
International Outlook	7.5%	Proportion of international students	2.5%
		Proportion of international staff	2.5%
		International collaboration	2.5%
Industry Income	2.5%	Knowledge Transfer	2.5%

Table 1 shows that reputation impacts the scores of indicators heavily, which affects Teaching and Research the most. Although reputation can reflect the level of universities, it is hard to measure accurately and fairly, and is mainly subjective. The weights of indicators reflect the focus of the ranking, but not all of the indicators are fairly weighted. Bowman *et al.* propose that the rankings are affected by the anchoring effect which means that they may be swayed by other institutions or rankings, and the authors also provide plans to optimize the reputations investigating method.[1] Soh points out that in some rankings, there is a negative or no correlation between indicators, and it also questions the rationality of indicators and the grading system.[2] Authors in the Sapienza University of Rome explored the consistency of different rankings, found out there are major discrepancies, and point out that indicators and their weights may be unreasonable.[3] Others believe that among the perfectly seeming grading system, there are only a small number of indicators that contribute heavily to the final ranking, which can cause fraud.[4] Therefore, our goal is to optimize the indicators so that they can accurately and truthfully evaluate universities to provide students with relatively reasonable references.

The world university rankings have strong hysteresis. When universities implement new policies on teaching or research, it often takes years to take effect. The Universities' reputations get affected by many different factors, and it is very hard to change the stereotype of people. Therefore, universities cannot maintain the current or future standard as the same as the one used in the ranking. Ranking predictions based on machine learning can utilize a massive amount of data from past years and seek out patterns, coming up with relatively reliable results, which solves the hysteresis problem to some extent, and helps high-school students to find suitable post-secondary education opportunities.

This paper uses data from the official website of THE rankings and adopts a data fitting method and a correlation method to analyze the data from both statistical and graphical perspectives, to optimize the existing ranking system. Also, this paper trains a back propagation neural network using past years' data to predict upcoming ranking.

Methods

Indicator Analysis:

To optimize the weights of indicators, a deep understanding on the rationality of the current indicators is needed. This paper determines the rationality of indicators in two aspects: First, the score of an indicator should correlate closely with the rank. That is, the higher a university scores on an indicator, in

DOI: 10.36838/v4i3.3

general, the higher its rank is. Besides, the score distribution of the indicators should be reasonable, so it can differentiate the different levels of universities.

Least square fitting[5] is used to determine whether there is a correlation between each indicator and ranking. Assume $A=\{(x_i,y_i)\}_{i=1}^{N}$ is a set with *N* data points, where *x* is the ranking, and *y* is the score. We find a straight-line *l*: *y=kx + b*, where the sum of the distance of every point $(x_i,y_i)_{i=1}^{N}$ in *A* to the corresponding point $(x_i,\hat{y}_i)_{i=1}^{N}$ on line *l* is the smallest, which can be expressed as Equation (1).

$$L=\sum_{i=1}^{N}\left|y_i-kx_i+b\right| \tag{1}$$

The distance L is determined by the parameters k and b, which can be solved using the polyfit () function in MATLAB.

Although least square fitting can show the correlation between scores and ranking intuitively, it cannot quantitatively analyze the correlation. Calculating the correlation coefficient of the ranking and score is one solution to evaluate the correlation. Pearson coefficient is used here and the larger absolute value of it indicates the higher correlation.[6] Positive or negative indicates whether the two are positively correlated or negatively correlated. The Pearson coefficient can be calculated by Equation (2).

$$P=\frac{\mathrm{cov}(X,Y)}{\sigma X\sigma Y} \tag{2}$$

where cov() is the covariance, σ is the standard deviation, *X* and *Y* represent the score of an indicator and the ranking respectively. The covariance is used to measure the variation trend of *X* and *Y*. If the covariance is negative, it indicates that the overall trend of the two is opposite and negatively correlated; if the covariance is positive, the two are positively correlated. The standard deviation in the denominator is to normalize the covariance to ensure that it can accurately reflect the correlation between *X* and *Y*.

Indicator Optimization:

The previous analysis shows that some indicators have problems in weight allocation or scoring standards. Some indicators have no contribution to the results or have the negative effect on the evaluation of universities, so that the ranking results cannot fully represent the actual level of universities. Since the scores of each indicator cannot be modified, the strategy of weight optimization is adopted here to modify the contribution of each indicator to the ranking, so as to reduce the influence of unreasonable indicators and scores on the ranking. Principal component analysis (PCA) is used here to find a better weight for each indicator.[7]

PCA is a statistical method that converts a group of variables that may be correlated into linearly independent variables (principal components) through orthogonal transformation, and utilizes the original information contained in the principal components to conduct subject analysis. To determine the weight, the principal components should be expressed as the linear combination of the original indicators, following Equation (3).

$$F_i=a_{i1}X_1+a_{i2}X_2+a_{i3}X_3+a_{i4}X_4+a_{i5}X_5 \tag{3}$$

where X is the standardized score of each indicator, and a is the eigenvector corresponding to the eigenvalue of the covariance matrix of X.

Ranking Prediction:

The world university ranking is usually hysteretic, with changes in teaching policies often not reflected in the rankings until several years later. The ranking of a certain university may change greatly over time. For example, Peking University ranked 42nd in 2016, 30th in 2019, and 24th in 2020.[8] In order to better reflect the current level of universities and provide some references for students who are about to choose universities, this paper tries to predict the rankings.

It is difficult to predict the ranking directly since it is influenced by scores of various indicators and other external factors. Therefore, this paper only predicts the total score (the score of an indicator is also possible) of a certain university. In the later stage, the total score can be used to calculate the rankings. The mean square error (MSE) is used to measure the difference between the predicted scores and the ground truth to verify the feasibility of the proposed method, which is given by Equation (5).

$$MSE=\frac{1}{n}\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2 \tag{5}$$

where *Y* is the predicted scores and $\hat{Y}$ is the ground truth. It is generally believed that the method is feasible when the difference between the predicted score and the ground truth is within 5 points, that is, the MSE is less than 25.

In this paper, a back propagation neural network is proposed to predict scores.[9] The neural network imitates the connection between nerve cells of the human brain to transmit information. The main idea is to use a large amount of data to train the model, so as to complete corresponding tasks under similar conditions.[10] In the training phase, the loss function is utilized to judge the error between the output and the ground truth, and model parameters are updated by a back propagation algorithm to make the output of the next epoch closer to the ground truth. By comparing the performance of models with different layers, this paper chooses to build a neural network model with five hidden layers, as shown in Figure 1.

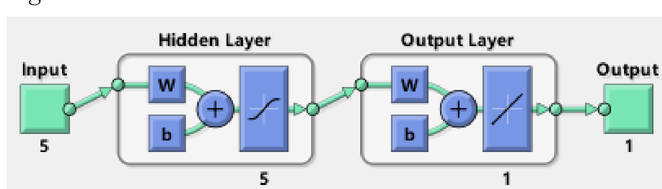

Figure 1: The score prediction neural network.

Results and Discussion

Indicator Analysis:

This paper uses polyfit() function to fit and analyze the relationship between the scores of each indicators and the ranking. Figure 2 shows the least square fitting results of the top 200 universities in 2020 THE Rankings, which shows the correlation between ranking and the score of each indicator intuitively.

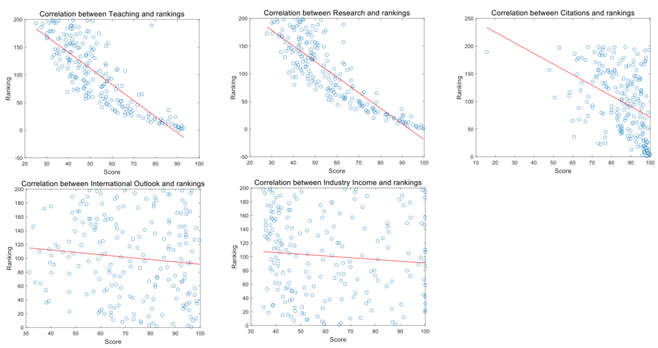

Figure 2: Results of the least square fitting of the top 200 universities for each indicator.

The distribution of data is scattered, and the slope of the fitted line is small, indicating that the correlation is weak. Data are distributed near the line with a large slope, indicating that there is a strong correlation between indicators and rankings. As shown in Figure 2, Teaching, Research, and Citations are strongly correlated with ranking, while International Outlook and Industry Income are weak. The analysis shows that the scores of International Outlook and Industry Income have smaller weights and have less impact on ranking than the other three indicators. In addition, it also reflects the problems in the correlation among indicators, that is, Teaching, Research, and Citations are more correlated. The above conclusions are consistent with the weight distribution in Table 1. However, some other different conclusions can be drawn in Figure 2. The weights of teaching, research, and the citations are the same, but the sum of the squared residuals of the citations is significantly larger than the first two, which means there may be some problems with the scoring of the citations. Besides, although the weights of the international outlook and industry income are relatively low, it can be seen from the data points that the scores of the two indicators have almost no effect on the ranking, whether the two indicators can be reflected in the final ranking is worth considering.

Then, I made a quantitative analysis of the correlation between each indicator and the ranking. In MATLAB, we can directly calculate Pearson correlation coefficients by using corr() function. The results of coefficients between scores and rankings are shown in Table 2.

Table 2: Pearson coefficient between indicators and rankings. The larger the absolute value of the coefficient, the greater the correlation with the ranking.

Indicators	Teaching	Research	Citation	International Outlook	Industry Income
P	-0.8096	-0.8560	-0.4096	-0.1082	-0.0907

Table 2 quantitatively shows that the score of Teaching and Research has the highest correlation with ranking, while International Outlook and Industry Income have a low correlation with ranking, which is consistent with the conclusion of least square fitting.

Next, I evaluated the rationality of scoring. Rational distribution of scores should be able to separate the good and poor schools, namely the scores should not be too concentrated in distribution in a certain range, and the number of universities that fall in the interval of high or low scores should be less. Also, the number of the moderate level universities should be relatively greater, which is in the common cognition of the evaluation of things. The data of the top 200 universities in THE Rankings in 2020 is used here to draw a score distribution histogram to analyze the rationality of indicators, as shown in Figure 3.

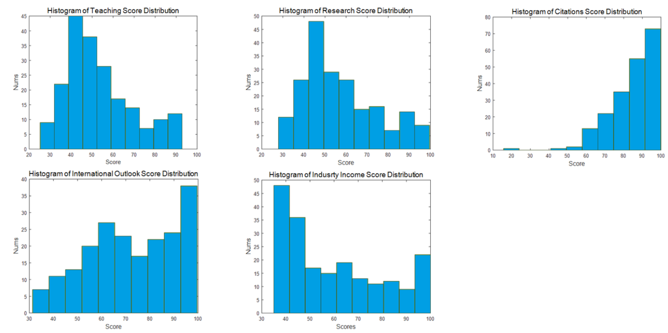

Figure 3: Histogram of score distribution of five indicators. There may be some problems with the Citations score distribution.

The score distribution of Teaching and Research is relatively reasonable, while there are many extreme scores in Citations and Industry Income. Taking Citations as an example, more than half of the universities scored more than 80 points, which is equivalent to adding roughly the same score to most of the universities. Besides, the citations of a paper depend on a variety of factors, including academic level, publication date, and field of study. Because of all these factors, it has little impact on the overall rankings, and it is difficult to distinguish the quality of the published paper. Theoretically, the weight of this kind of indicator should be reduced accordingly.

By analyzing the indicators and weights of THE Rankings, we can see that there are indeed some problems in the evaluation system that affect the reliability of rankings. The weak correlation between some indicators and the ranking or no correlation between indicators reflect the unreasonable selection of indicators and weight allocation in the ranking, which need to be specially adjusted for these problems.

Indicator Optimization:

The pca() function in MATLAB can be used to obtain the information content E in the five principal components. The results are shown in Table 3. The first four principal components already contain more than 98% of the information, so only the first four are used for subsequent analyses.

Table 3: The information content of each principal component.

Principal Component	F1	F2	F3	F4	F5
Information E (%)	39.3681	28.7645	17.9556	12.1994	1.7125

The relative weights W of indicators in Equation (4) can be obtained by evaluating the importance of each principal component with the amount of information it contained. The final weight is the normalized weight vector W, as shown in Table 4.

Table 4: The comparison of weights before and after optimization.

Indicators	Teaching	Research	Citations	International Outlook	Industry Income
Original Weights	0.3	0.3	0.3	0.075	0.025
Optimized Weights	0.2236	0.2714	0.2136	0.2011	0.0903

DOI: 10.36838/v4i3.3

The optimization results show that the weights of Teaching, Research, and Citations have decreased to varying degrees, while the weights of International Outlook and Industry Income have increased. As previously analyzed, the correlation between rankings and International Outlook or Industry Income is weak, because the original weights of the two indexes are too small to be reflected in the final rankings. The weights of Teaching and Research decrease since both of them are affected by reputation, and the decrease in the weight is conducive to improving the objectivity of the ranking. The decrease in Citations may be due to the unreasonable scoring range of this indicator, which needs to reduce its influence on ranking.

The results of the optimization and analyses methods in this paper verify each other, which proves that the analyses of indicator and weight optimization are reasonable to a certain extent. The method proposed in this paper is not the only and not necessarily the optimal optimization method. The purpose is to provide a new way to analyze the reliability of world university rankings by using the principles of mathematics and statistics.

Ranking Prediction :

I collected the scores of the top 200 universities in THE Rankings from 2011 to 2016 and selected 147 of them that have been in the top 200 for six years. The 147 universities were divided into three parts: a training set, validation set, and test set, with a data ratio of 80:20:47. These three were used to learn data features, update parameters, and test performance respectively. With the scores from 2011 to 2015 as input, the neural network model learned the features and changing trend of scores in the five years and predicts the scores in 2016. Finally, the MSE of the neural network reaches 12.21, which means that the average error between the ground truth and the predicted value is only about 3.5 points, which proves that it is relatively reliable to use neural network to predict scores.

■ Conclusion

This paper analyzes the rationality of THE World University Rankings. The correlation between indicator scores and rankings are used to measure the rationality of the indicators, using the least square fitting and the Pearson correlation coefficient to filter out the reasonable indicators from both statistical and graphical perspectives. The Principal Component Analysis is used to optimize the weights of indicators and combines the optimized results with theoretical analysis for validation. In addition, this paper designs a back propagation neural network model, which contains 5 hidden layers, using data from years 2011-2016 to train the model, and data from 2017 to check the model. It has an error margin of around 3.5, so it can predict future rankings relatively accurately.

The purpose of this paper is to provide new ideas and methods for the analysis of the world university rankings, hoping to offer an improved reference for many high school students to choose their own university. Due to my deficiencies in the breadth and depth of relevant knowledge, this paper still has deficiencies in the rationality evaluation criteria of weights and the robustness optimization of neural network performance, and I still need further study and research in this.

■ Acknowledgments

I would first like to express my highest gratitude to my parents for their endless help in my research. In addition, I would like to thank Dr. Chang for providing me with the ranking data of the THE World University Rankings. Finally, I am thankful to the IJHSR team for providing the platform to show my research.

■ References

1. Bowman N A, Bastedo M N. Anchoring effects in world university rankings: exploring biases in reputation scores. Higher Education, 2011, 61(4): 431-444.
2. Soh K. What the Overall doesn't tell about world university rankings: examples from ARWU, QSWUR, and THEWUR in 2013. Journal of Higher Education Policy and Management, 2015, 37(3): 295-307.
3. Moed H F. A critical comparative analysis of five world university rankings. Scientometrics, 2017, 110(2): 967-990.
4. Hou Y W, Jacob W J. What contributes more to the ranking of higher education institutions? A comparison of three world university rankings. International Education Journal: Comparative Perspectives, 2017, 16(4): 29-46.
5. Chen G, Ren Z, Sun H. Curve fitting in least-square method and its realization with Matlab. Ordnance industry automation, 2005, 3: 063.
6. Benesty J, Chen J, Huang Y, Cohen I. Pearson correlation coefficient[M]//Noise reduction in speech processing. Springer, 2009: 1-4.
7. Wold S, Esbensen K, Geladi P. Principal component analysis. Chemometrics and intelligent laboratory systems, 1987, 2(1-3): 37-52.
8. The official website of THE World University Rankings. https://www.timeshighereducation.com
9. Hecht-Nielsen R. Theory of the backpropagation neural network. Neural networks for perception. Academic Press, 1992: 65-93.
10. Li J, Cheng J, Shi J, H Fei. Brief introduction of back propagation (BP) neural network algorithm and its improvement. Advances in computer science and information engineering. Springer, 2012: 553-558.

■ Author

Yingao Gao is currently a junior at Mount Douglas Secondary in Victoria, BC. He gained interest in math and computer science by participating in competitions and hopes to continue studying these in post-secondary school.