Comprehensive Quality Evaluation for Secondary School Students Based on Big Data

Yulong Chen[1], Yanyan Li[1*], Wenzheng Li[1], Xinyou Li[1], Weiqi Zhou[1], Xiaoye Wang[1], Yiyi Long[1] and Yi Wu[1]

[1]School of Computer and Information Engineering, Hubei Normal University, Huangshi, China

{ Yulong Chen, Yanyan Li, Wenzheng Li, Xinyou Li, Weiqi Zhou, Xiaoye Wang, Yiyi Long, Yi Wu }

ylchen0424@163.com, 3601892961@qq.com, gaia233yue@gmail.com, wekek198@gmail.com, 2387516978@qq.com, 1084165590@qq.com, 1572469804@qq.com, 3334123658@qq.com

Abstract: The present study proposes a kind of comprehensive quality evaluation model based on big data with an attempt to rank and grade secondary school students' comprehensive quality ability in Shandong, in order to optimize the traditional comprehensive quality evaluation model. The model is a fusion of TOPSIS method, RSR method, and entropy weight method. The Pandas package in the Python programming language is used for data extraction and data processing, the Matplotlib package is used to draw line graphs, and the Numpy package is used for model implementation and computation. The results of the ranking reveal that 2 students have imbalanced comprehensive quality development and that 5 students with balanced comprehensive quality development achieve higher rankings. The grading results reveal that 12% of the students receive an "excellent" rating, 68% receive a "good" rating, and 20% receive a "poor" rating. The results prove that the model optimizes the traditional comprehensive quality evaluation method and focuses more on the comprehensive development of students in the process of ranking and grading them. The proposed model is significant for the development of comprehensive quality evaluation model.

Keywords: comprehensive quality evaluation model; big data; rank and grade; student's comprehensive abilities

1 Introduction

With the continuous advancement of education in China, comprehensive quality assessment of secondary school students has become a hot direction to improve the education ecology, and comprehensive quality assessment of secondary school students is an important initiative to comprehensively deepen the reform of evaluation methods and deeply promote quality education for students. Change the practice of assigning grades to students, innovate process evaluation methods for moral, intellectual, physical, social, and aesthetic work, and improve the comprehensive quality evaluation system[1]. The scientific evaluation method is an important guarantee that the comprehensive quality evaluation can give full play to the function of educating people.

We discusses educational changes in comprehensive student evaluation in response to educational changes. In education, evaluation methods and results are crucial, not only to help students learn but also to have a certain impact on the selection of talents in society[2]. Therefore, evaluation methods are worthy to be taken seriously [3].

The importance of students' total development is becoming more apparent in this day and age, which inevitably means that curriculum and methods of evaluating student performance will change. Therefore, it is necessary to reconsider educational assessment, Robert L. Linn proposed in 1991 a criterion to solve the problems related to this area, but his solution is still rather vague[4]. Robert J. Mislevy proposed the use of a sparse matrix of SAT scores to address student assessment, which is close to the scientific method but does not take into account enough factors[5]. In terms of promoting the reform of assessment methods, the establishment of student digital files, the promotion of big data-based assessment methods, and the support of the development of the whole process of longitudinal assessment of students' learning for all grades and the horizontal assessment of all elements of moral, intellectual, physical, social and aesthetic development, Richard J Patz considered the application of big data technology in the examination process to consider the proficiency of students[6], which greatly proves the impact of big data information science on educational reform. The use of big data technology to fully analyze and excavate the comprehensive quality information of secondary school students enables the process of comprehensive quality evaluation of secondary school students to be precise and intelligent[7], thus improving the quality of comprehensive quality evaluation. Ao et al. proposed using RSR to deal with this problem for the comprehensive quality evaluation of college students[8], and compared the results to the traditional method.Based on the understanding of comprehensive quality assessment, TOPSIS[9] with Entroy method and RSR for comprehensive quality ranking and grading of secondary school students were proposed by us and have been proven to be effective.

2 Methodology

In this study, the comprehensive quality evaluation profile of secondary school students provided by the Second Data Application Innovation and Entrepreneurship Competition in Shandong Province This study takes the comprehensive quality evaluation files of secondary school students provided by the Second Data Application Innovation and Entrepreneurship Competition in Shandong Province as the research object. The five dimensions of ideological and moral, academic, physical, extended learning, and practice were selected from the student files to research the comprehensive quality evaluation of secondary school students. The data was extracted using Python programming and imported into an Excel spreadsheet. The data for the academic and physical fitness dimensions are represented by the average academic credits and the average physical fitness score in the student's profile. The data for the three dimensions of ideology, extended learning, and practice is represented by the number of typical examples of ideology and morality, the number of artistic experiences, and the number of social practices in the student profile. Then, We consider implementing comprehensive evaluation of education from these methods:

The comprehensive quality evaluation model for secondary school students incl udes 3 parts: data processing, comprehensive quality assessment, and ranking and grading. In terms of data processing, the quantitative values of the five indicators of students were obtained through indicator screening and statistical analysis and were recorded $X_1 \sim X_5$ in order. Let the number of students be n, each student h as m indicators, and the original student data matrix $X \in R^{n \times m}$,where X_{ij} denotes the j_{th} indicator value of the i_{th} student. The comprehensive quality evaluation in cludes 3 parts: calculation of indicator weights, ranking of fit, and grading test, T he detailed process is described below. The comprehensive quality assessment mo del based on big data is shown in Fig 1.

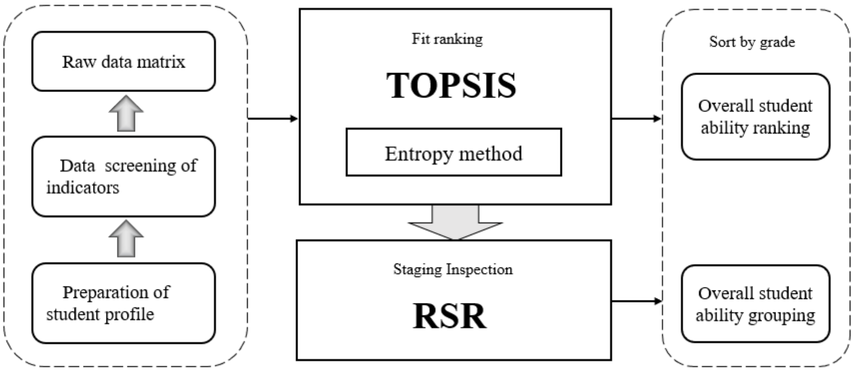

Fig. 1. Comprehensive quality evaluation model based on big data

2.1 Calculation of indicator weights

Indicator weights are calculated using the entropy weight method. The entropy weighting method is an objective assignment method that determines the weight of an indicator based on the variability of the orderliness of the information contained in each indicator and relies only on the dispersion of the data itself. The size of the indicator weight indicates the importance of the indicator. The greater the weight of the indicator, the greater the weight in the comprehensive evaluation.

First, the original student data matrix X is ratio-normalized using the ratio-normalization:$p_{ij} = x_{ij} / \sum_{i=1}^{n} x_{ij}$.

where x_{ij} denotes the value of the ith parameter for the jth studentp_{ij}, denotes the ratio normalized index value. Then, the entropy value corresponding to each index of all students is calculated:$e_j = -k_{ij} \sum_{i=1}^{n} p_{ij} \, ln \, p_{ij}$, $(j = 1,2,\cdots,m)$.

where k_{ij} is related to the sample index value, $k_{ij} = 1/ ln\left(p_{ij}\right)$. Finally, the weights corresponding to each indicator were calculated for all students:$\omega_j = (1 - e_j) / \sum_{k=1}^{m}(1 - e_k)$, $(j = 1,2,\cdots,m)$.

2.2 Fit ranking

The fit ranking is performed by combining the TOPSIS method (Technique for Order Preference by Similarity to Ideal Solution) and the weights obtained from the index weights to calculate the fit and ranking. The TOPSIS method is a ranking method based on the proximity of a limited number of evaluation objects to the idealized target.

method. In this study, this method was used to establish a comprehensive evaluation of secondary school students to provide a quantitative explanation of the comprehensive rationality of student evaluations.

First, the original student data matrix is vector normalized to obtain the student normalization matrix $Z \in R^{n\times m}$, where: $Z_{ij} = x_{ij}/\sqrt{\sum_{i=1}^{n} x_{ij}^2}$.

The maximum and minimum values of the elements in each column Z are counted separately to obtain the optimal solution Z^+ and the worst solution Z^-:

$$Z^+ = (max\{z_{11}, \dots, z_{n1}\}, \dots, max\{z_{1m}, \dots, z_{nm}\})$$

$$Z^- = (min\{z_{11}, \dots, z_{n1}\}, \dots, min\{z_{1m}, \dots, z_{nm}\})$$

Then, the degree of proximity D_i^+,D_i^- between each student's indicators and the optimal and inferior solutions was calculated by the optimal and inferior solutions and the weight values ω_j of each student's indicators were obtained by introducing the entropy weighting method:$D_i^+ = \sqrt{\sum_{j=1}^{m} \omega_j \left(Z_j^{\pm} z_{ij}\right)^2}$, $D_i^- = \sqrt{\sum_{j=1}^{m} \omega_j \left(Z_j^- - z_{ij}\right)^2}$. Finally, the fit between each student and the optimal solution was calculated $C_i = D_i^-/(D_i^+ + D_i^-)$, and the fit was used as the total score for each student, and all students were ranked according to the size of the total score.

2.3 Grading test

Firstly, the TOPSIS method was used to calculate C_i for each student, and C_i was substituted to determine the distribution; secondly, they RSR were sorted from smallest to largest, the frequency f of each group was listed, the cumulative frequency $\sum_{i=1}^{n} f$ of each group was RSR calculated, the rank R and average rank $\bar{R}$ of each group was determined, and the downward cumulative frequency $\frac{\bar{R}}{n}$ was calculated; finally, the corresponding probability unit $Probit$ values were found based on the cumulative frequency query percentages and the probability unit comparison table.

After determining the RSR distribution the resulting probability unit$Probit$ was used as the independent variable and the RSR value was used as the dependent variable to form the regression equation $RSR = a + bProbit$. The equation was tested for statistical significance using least squares estimation. The RSR method uses information from multiple evaluation indicators to measure the sum level of multiple indicators and introduces a fit to all subjects in addition to eliminating the effect of the scale of the indicators. The comprehensive process is shown in Fig 2.

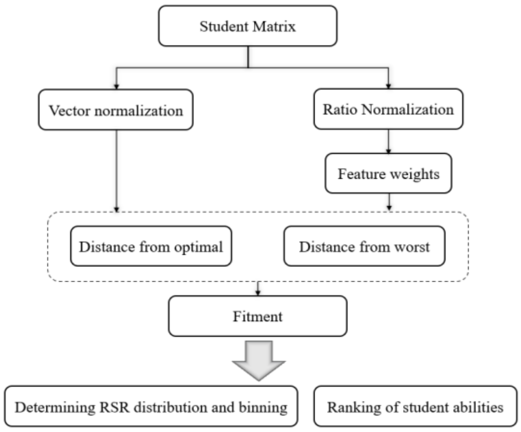

Fig. 2. Calculation flow

3 Experiment and results

3.1 Analysis of students' comprehensive ability ranking results

In this study, the positive ideal solution, the negative ideal solution, and indicator weight values corresponding to each indicator are obtained according to the calculation steps of the TOPSIS method and entropy weight method[10]. The positive solution represents the optimal solution in the fit ranking method and consists of the maximum value of each indicator parameter. The negative solution represents the worst solution in the fit ranking method and is composed of the minimum value of each indicator parameter. The indicator weight values are obtained using the entropy weight method and are used to calculate the importance of the information contained in each indicator of the student. The statistics of positive ideal solutions, negative ideal solutions, and indicator weights of students are shown in Table 1.

Table 1. Solution and weighting statistics

	Academics	Physique	Moral	Practices	Extended
Negative	0.088	0.078	0	0	0
Positive	0.121	0.138	0.297	0.262	0.34
Weight	0.001	0	0.301	0.085	0.31

As can be seen from Table 1, the positive ideal solution has more stable data, and the negative ideal solution has the lowest values of ideological character, practical and extended learning as 0. This indicates that there is an uneven development of student's abilities in the three indicators. The ranking of the indicator weight values shows that the other four indicators have more influence on the overall evaluation than academics, further indicating that this method improves the shortcomings of the traditional evaluation system that relies only on academic scores to consider students unilaterally. By using the positive ideal solution, negative ideal solution, and index weights, we can calculate the proximity between each evaluation index and the best and worst solutions for each student, and finally use the proximity calculation to get the fit of each evaluation object and rank the students according to the fit, and finally get the rank of each student. The students' fitting scores are shown in Fig 3.

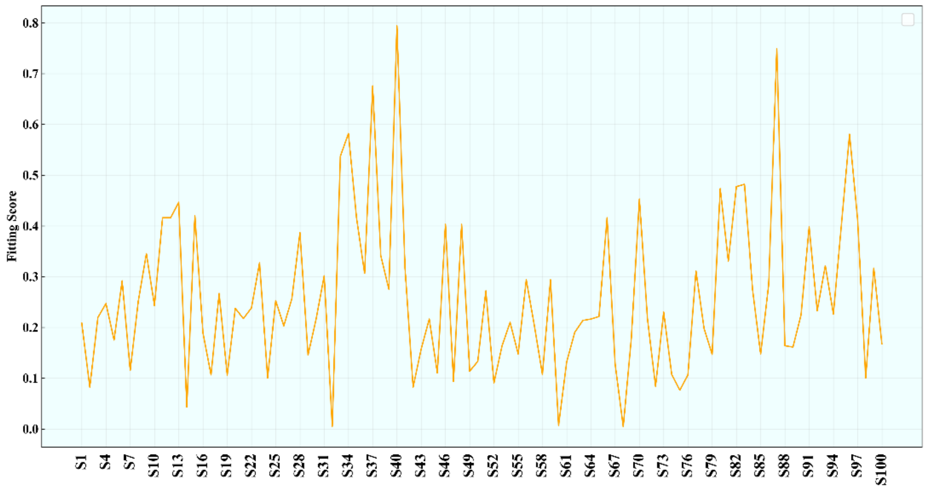

Fig. 3. Comprehensive student fit score

From Fig 3, it can be seen that there is a bifurcation in the comprehensive ability among students, for example, the 32nd student's comprehensive ability fit score is less than 0.1, while the 40th student's comprehensive ability fit is greater than 0.7, which is a better indication of the uneven development of student's abilities. In addition, the top 5 students in the comprehensive ranking of the fit were students 40, 87, 37, 34, and 96, respectively, and by looking at the index scores of these 5 students, we found that none of these 5 students had the most outstanding index values, but the comprehensive ranking was high because of the balanced ability values of the indexes. This result proves that this method can accurately evaluate students' abilities from multiple perspectives in a comprehensive manner.

3.2 Analysis of students' comprehensive ability grading results

A series of statistical results were obtained by using C_i instead of RSR to determine theRSR distribution, which is shown in Table 2. With the probability single Probit as the independent variable and C_i as the dependent variable, a linear regression equation $C_i = 0.161Probit - 0.548$ was established in SPSS software, and the obtained equation was: significance was 0.0, indicating that the linear regression equation is statistically significant and can be carried out the next step of the binning operation.

Table 2. RSR distribution statistics

C_i	f	$\sum f$	R	$\bar{R}$	Downward frequency	$Probit$
0.01	3	3	1,2,3	2	2	2.95
0.04	1	4	4	4	4	3.
…	…	…	…	…	…	…
0.75	1	99	99	99	99	7.33
0.79	1	100	100	100	99.75	7.815

The comprehensive quality evaluation results of 100 secondary school students were divided into three grades: excellent, good, and poor, and the percentages P for dividing the grades were set by themselves, and the probability units corresponding to Pthem were brought into the above regression equation to obtain the graded critical C_i values. The obtained C_i values were compared in size with the actual C_i values obtained by the TOPSIS method for each student, and the students were graded operationally, and the results are shown in Table 3.

Table 3. Staging Results Statistics

Grade	P	$Probit$	C_i	Grading student id
Excellent	$\geq 85\%$	> 6	> 0.42	12,15,33, … , 87,96
Good	$\geq 50\%$	$4.16 \sim 6$	$0.12 \sim 0.42$	1,3,4, … ,97,99,100
Poor	$<50\%$	< 4.16	< 0.12	2,7,14, … ,75,76,98

As shown in Table 3, a total of 12% of the students received excellent evaluations, 68% received good evaluations, and 20% received poor evaluations. Comprehensive, the number of students who received excellent and poor ratings was relatively small, and the number of students who received good ratings was relatively large. The results of the comprehensive quality evaluation were normally distributed, the differentiation of students' evaluations was obvious, and the students were ranked reasonably, which improved the quality of the comprehensive quality evaluation. Finally, the results of the grading results were tested using the variance consistency test, and the results showed that the values of each grade C_i met the variance chi-square (F=96.156, p<0.001), that is, the differences between the obtained grades were statistically significant, which further illustrated the accuracy and reasonableness of the grading scheme and conclusions.

4 Conclusion

The deepening of education evaluation reform is a key period for the gradual improvement and enrichment of the comprehensive quality evaluation of students at the secondary school level. Comprehensive quality assessment reform is the inevitable result of deepening reform of education itself, and this reform course of groping forward is also a colorful stroke in the history of China's education development, showing the historical vein of fairness and scientific nation of China's education. In the context of big data, the comprehensive quality evaluation method of secondary school students based on big data deeply integrate the data of all dimensions of students' comprehensive quality and realize the flexibility, scientificity, and accuracy of comprehensive quality evaluation.

5 Renferences

1. Feng, Qin, Cheng Yao, and Zheng Xiao. "Construction of Students' Comprehensive Quality Evaluation Model Based on Improved AHP." In *2018 13th International Conference on Computer Science & Education (ICCSE)*, pp. 1-5. IEEE, 2018.
2. Martone, Andrea, and Stephen G. Sireci. "Evaluating alignment between curriculum, assessment, and instruction." *Review of educational research* 79, no. 4 (2009): 1332-136
3. Jonsson, Anders, and Gunilla Svingby. "The use of scoring rubrics: Reliability, validity and educational consequences." *Educational research review* 2, no. 2 (2007): 130-144.
4. Linn, Robert L., Eva L. Baker, and Stephen B. Dunbar. "Complex, performance-based assessment: Expectations and validation criteria." *Educational researcher* 20, no. 8 (1991): 15-21.
5. Mislevy, Robert J., Albert E. Beaton, Bruce Kaplan, and Kathleen M. Sheehan. "Estimating population characteristics from sparse matrix samples of item responses." *Journal of Educational Measurement* 29, no. 2 (1992): 133-161.
6. Patz, Richard J., Brian W. Junker, Matthew S. Johnson, and Louis T. Mariano. "The hierarchical rater model for rated test items and its application to large-scale educational assessment data." *Journal of Educational and Behavioral Statistics* 27, no. 4 (2002): 341-384.
7. Baradwaj, Brijesh Kumar, and Saurabh Pal. "Mining educational data to analyze students' performance." *arXiv preprint arXiv:1201.3417* (2012).
8. Ao, Xiqin, and Aixiang He. "Research on the comprehensive evaluation method of college students' achievements based on multi-view RSR." In *Journal of Physics: Conference Series*, vol. 1616, no. 1, p. 012072. IOP Publishing, 2020.
9. Yoon, Kwangsun, and Ch L. Hwang. "TOPSIS (technique for order preference by similarity to ideal solution)–a multiple attribute decision making, w: Multiple attribute decision making–methods and applications, a state-of-the-at survey." *Berlin: Springer Verlag* 128 (1981): 140.
10. Chen, Faan, Jianjun Wang, and Yajuan Deng. "Road safety risk evaluation by means of improved entropy TOPSIS–RSR." *Safety science* 79 (2015): 39-54.