Research on College Class Style Evaluation Based on Big Data

Yanfei Chen
School of Electronic Information Engineering
Foshan University
Foshan, China

Chunying Huang
School of Electronic Information Engineering
Foshan University
Foshan, China

Xinhong Chen
School of Electronic Information Engineering
Foshan University
Foshan, China

Guitian Liu
School of Electronic Information Engineering
Foshan University
Foshan, China

Zhiwei Zheng
School of Electronic Information Engineering
Foshan University
Foshan, China

Ruijia Liu
School of Electronic Information Engineering
Foshan University
Foshan, China

Xiaohua Li *
School of Electronic Information Engineering
Foshan University
Foshan, China

Abstract—**Class style is an important indicator affecting the quality of talent training in colleges and universities. How to establish an effective class style evaluation model is of great significance to promote the formation of excellent class style and provide decision support for university personnel training programs. Compared with previous studies, this paper comprehensively considers the factors affecting the evaluation based on the principle of combining subjectivity with objectivity. With the changes of higher education reform, we have prospectively incorporated the indicators that may reflect the status of the work style into the evaluation system. On the basis of these innovative principles, this paper realized low-cost data collection with digital campus, and combined entropy weight method, high-order factor analysis method and similarity cosine calculation method to build a class evaluation analysis model based on big data.**

Keywords—*big data; class style evaluation; entropy weight method; factor analysis*

I. INTRODUCTION

The construction of class style is one of the important parts of the development of school education, which is influenced by the content of teaching and the behavior of students in school. At present, most of the research perspectives are based on certain topics, and the sample data is collected directly through questionnaires for analysis [1]. This type of analysis has limited dimensions. It lacks of professionalism and model to support, and does not make full use of the advantages in the process of college construction. Therefore, most of the current class evaluation system has not integrated a more comprehensive class evaluation object and has not yet established an effective mathematical model. Most universities have carried out digital campus construction, which provides big data conditions from different angles and in-depth analysis for the study of the construction of class style. Based on the digital portrait of students and classes, this paper establishes the class style evaluation data object of each class in the department, and constructs the class style evaluation model with multi-dimensional data analysis. The output of the model can make people intuitively get the class with outstanding achievements in the construction of the class. Through the data mining of the excellent class, the hidden performances of many excellent classes in the history are obtained, which provides a reference for the construction of colleges and universities, refined management, and the innovation of educational models.

II. CURRENT SITUATION AND DEMAND ANALYSIS OF CLASS STYLE EVALUATION IN COLLEGES AND UNIVERSITIES

A. Existing problem

The main problems existing in the current evaluation of class atmosphere in Colleges and universities are as follows: (1) Class activities have an impact on the construction of class style, but some of its indicators, such as the potential value of the theme of class activities, are not easy to quantify, which makes it difficult for such indicators to be included in the

evaluation system, (2) Although the construction of class style is rich in content, the performance of students' professional learning is the main factor, while the construction of campus culture and investment of teaching resources are paid less attention to[2], (3) in the evaluation principle of class style, we should distinguish between faculties and departments. Different institutes have different needs in infrastructure and faculty level, which makes the evaluation of class style have to be analyzed in the case of distinguishing different faculty and department, (4) the data describing students and classes are severely fragmented. Although data within a school can be shared at low cost, due to the lack of unified planning of campus information, various functional departments have built many systems based on their own business requirements. These systems are not interoperable, so no one department can come up with data that reflects the overall picture of students and classes. Those made the evaluation of students and classes by departments and teachers at all levels more subjective and one-sided. Therefore, different teachers have different images of the same student or class. Images of the same student or class, (5) the data has insufficient driving force for student and class management. On the one hand, the fragmentation of data describing students and classes makes the data only drive part applications. On the other hand, most of the existing data are based on traditional goal management, which only describes students and classes in rough lines, coarse granularity and stages. It is hard to reflect the real face of students and classes. This results in that unless students and classes are particularly prominent, it is difficult to reflect the real-time differences between ordinary students and classes. Thus the drive for accurate management and evaluation of students and classes is weak, (6) the inadequate investment in information technology leads to difficulties in data acquisition. To make an overall portrait of the students and classes, it is necessary for schools to plan a complete data sensing network, data transmission network coverage and data acquisition platform. But now these aspects of construction investment are insufficient, (7) The evaluation process of students and classes is rough, the scope of evaluation is not comprehensive, and the evaluation method is subjective[1].Firstly, existing data describes students and classes in rough lines, coarse granularity and stages, which makes it impossible to refine real-time evaluation based on current data. Secondly, most of them adopt simple statistical comparative evaluation. Some of the applications that use mathematical models to support student and class management decisions are based on the evaluation indicators or grades given subjectively by individuals (experts or decision makers).This decision is not comprehensive and subjective, which based on the analysis about cluster, classification or random process analysis of some aspect of student or class management.

B. Demand Analysis

In order to solve the above problems, class style evaluation needs to be based on certain principles. First we used a formal processing method, and then used the matrix composed of class multidimensional data as the input of the model. Finally, the class evaluation object was modeled by the algorithm to achieve the following analysis objectives.

1) Comprehensive analysis. There is much information describing a student, such as basic student information, student

family status, etc. [3]. Some of the information is static, and some of the information is constantly changing, such as consumption and competition conditions, etc. Those dynamically changing information are actually of high value, but they are often overlooked by traditional evaluation systems. Therefore, when setting the evaluation indicators, we must fully consider the factors that can reflect the basic situation of the class, and incorporate these factors into the evaluation system to make the evaluation angle more comprehensive [4].

2) Forward-looking analysis. The education methods of higher education are constantly reformed, and the connotation of class style will change with the change of education methods. Therefore, when setting the evaluation index of class style, we should fully consider the direction and trend of future education mode reform, and prospectively incorporate the indicators that may reflect the status of the class in the future [5].

3) Subjective and objective analysis. According to the principle of combining subjectivity and objectivity, indicators such as students' perceptions of the occupational environment and learning pressures are subjectively evaluated through questionnaires. At the same time, the indicators such as test scores, awards, and punishments can be directly evaluated objectively to compare the advantages and disadvantages. Finally, the results of the evaluation after comprehensive decision-making will be more scientific.

4) Departmental analysis. Formulating scientific and reasonable evaluation standards is an important part of the evaluation of class style. In the process of formulating evaluation criteria, the formulators should pay full attention to the differences between different departments and establish a multi-dimensional dynamic evaluation standard system according to different departments to realize real-time evaluation of dynamic standard class styles with objective and subjective supplements.

III. CONSTRUCTION OF CLASS STYLE EVALUATION IN COLLEGES AND UNIVERSITIES

A. Analysis on the affecting factors of class style evaluation in Colleges and Universities

Through analysis and the definition of experts, the main factors affecting the evaluation of class style in colleges and universities are classes and students, as shown in Table 1:

B. Colleges and universities class style appraise the process

Colleges and universities class style appraise process includes four sub-processes: questionnaire evaluation, comprehensive evaluation, and model output and revision validation.

1) Questionnaire evaluation: Questionnaire evaluation collects data by using questionnaires, and uses high-order factor analysis technology of structural equation model to get Factor Load and Fitting Index of questionnaire, so as to weight allocation: add all loads to the total load, divide a load of each index by total load. Assuming the number of all classes in a department is N, the weights are used to get the questionnaire

scores of each class, that is to say, the questionnaire evaluation scores of N classes are obtained.

Typeface	Attribute Set
Classes	teacher's evaluation of class discipline, number of class awards, rank of class score, participation rate of class activities, participation rate of class meeting
Students	sex, score of psychological test, student's evaluation of teachers, monthly flow of campus network, monthly consumption of one-card, number of consultation to school doctor, time of library lending, frequency to library, number of apply for laboratory, time of using laboratory, number of advanced events, number of academic achievement, internship duration, whether to participate in the competition, whether a league member, whether a party member, whether to be punished, whether to get a scholarship, whether to go abroad, whether to be an instructor, whether success in postgraduate entrance examination, whether to apply for postgraduate entrance examination, whether passed TEM4,whether passed TEM6, whether to get a degree ,whether employment within two years after graduation, whether a single parent[6], whether to participate to only one society, whether to donate blood. the degree of cognition about occupational environment, the clear-cut degree of career, whether the family supports it, the clear-cut degree of learning objective, the confidence of learning, the efficiency of learning ,the degree of learning stress, the frequency of preview, the frequency of review, the length of time in a class, whether to pay attention to current events, initiative to expend professional knowledge, frequency of dormitory cleaning, the atmosphere of dormitory learning, dormitory membership, awareness of the major, the satisfaction with the majors studied, the satisfaction with the class, the evaluation of the reasonableness of the school's schedule, the evaluation of the scientific degree of the difficulty-and-easy connection in the curriculum.

2) Comprehensive evaluation: In the data management module of the system, the data is formatted and imported into the database. The system extracts N data objects from the database. Assuming that there are X evaluation attributes for each object, a matrix A of N times X can be formed. The weights of X evaluation attributes in A are calculated by using the method of entropy weight. Then the weights of these attributes are weighted and calculated. Finally, the comprehensive evaluation scores of N data objects are obtained.

3) Model output: In a certain proportion, the weighted sum of the comprehensive evaluation score and the questionnaire evaluation score are used to get the final scores of all classes, and then the classes are sorted according to the scores to get the optimal class.

Revision validation: Verify whether the excellent class in the historical data agrees with the excellent class in the output of the model. There are two kinds of excellent classes from historical data, one is the complete ranking of excellent classes and the other is only one optimal class. According to different situations, different verification methods are used: (1) if the validation basis of historical data is complete ranking, the historical verification data ranking is calculated with the comprehensive evaluation ranking and the questionnaire

evaluation ranking respectively to obtain two similarity cosine values, and then compare the two and adjust the evaluation method with poor similarity. The revision of the questionnaire evaluation can use the new version of the questionnaire from the depth and angle of different problems [7]. The revision of the comprehensive evaluation can be adjusted by modifying the attributes of data acquisition, optimizing mathematical analysis methods and so on, so as to get a new round of model output. The process of calculating the similarity mentioned above is recursive until it is consistent with the results of historical data. (2) If the validation basis obtained from historical data is only the best class A, we can compare A with the best class output from the model. If the two are inconsistent, we can find out the ranking of A in the comprehensive evaluation ranking and questionnaire evaluation ranking, adjust the lower ranking evaluation method, reduce the proportion allocated to it, until with the results of historical data tend to be consistent.

IV. ANALYSIS ALGORITHM

A. Application of Entropy Weight Method

The application process of entropy weight method [8] in college class style evaluation is as follows: (1) Constructing the original matrix according to the data source, we suppose there are N students, each corresponding to M indicators, then I is the value of the j indicator of the ith student (i = 1,...,N; j = 1,..., M); (2) The unit of measurement of each indicator is not uniform, so these indicators should be standardized before calculating the information entropy. It is to convert the absolute value of the indicator into a relative value, so that the heterogeneous indicator is homogenized. In addition, since the positive and negative indicators represent different meanings, different algorithms are needed for normalization of the forward and negative indicators:

Positive indicator:

$$x_{ij} = \frac{x_{ij} - \min\{x_{1j},...,x_{nj}\}}{\max\{x_{1j},...,x_{nj}\} - \min\{x_{1j},...,x_{nj}\}} \quad (1)$$

Negative indicators:

$$x_{ij} = \frac{\max\{x_{1j},...,x_{nj}\} - x_{ij}}{\max\{x_{1j},...,x_{nj}\} - \min\{x_{1j},...,x_{nj}\}} \quad (2)$$

Note: For the sake of convenience, the normalized data X_{ij} is still recorded as X_{ij}.

(3) Calculate the proportion of the ith sample value under the j indicator to the indicator:

$$p_{ij} = \frac{x_{ij}}{\sum_{i=1}^{n} x_{ij}}, i = 1,...,n, j = 1,...,m \quad (3)$$

(4) Calculate the information entropy of the jth indicator:

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

$$k = 1/\ln(n) > 0 \tag{4}$$

$$e_j \geq 0$$

(5) Calculate information entropy redundancy:

$$d_j = 1 - e_j, j = 1,...,m \tag{5}$$

(6) Calculate the weight of each indicator:

$$\omega_j = \frac{d_j}{\sum_{j=1}^{m} d_j}, j = 1,...,m \tag{6}$$

(7) Calculate the composite score for each sample:

$$s_i = \sum_{j=1}^{m} \omega_j x_{ij}, i = 1,...,n \tag{7}$$

B. Application of Factor Analysis

On the basis of factor analysis [8], the processing of questionnaires in colleges and universities is as follows: (1) set up the questionnaire. All indicators are in the form of Likert five-level scale in the questionnaire, and each problem is classified [9]. (2) Data collection. Let all users complete the questionnaire and we get the data. (3) The high-order factor analysis technique of structural equation model is used to evaluate and analyze the data, and the Factor Load and Fitting Index of each index are obtained. (4) Weight assignment. According to the load magnitude of the high-order factor analysis of the model obtained in step (3), the weights are assigned to the indicators. The load of the index system is added, and then the load of the index is divided by the total load as the weight. Similarly, other weights can be obtained.

V. CONCLUSION

In the research process of this model, we found that the traditional class style evaluation system has incomplete factors, the data collection method is too single, and the subjective degree of the evaluation process is too strong and so on, which led to the system cannot provide valuable value for the construction of college class style. The emergence of big data technology makes it possible to objectively evaluate the class style of colleges and universities. Through real-time transmission and analysis of big data, this research realized real-time monitoring, as well as constructed a reasonable, scientific and comprehensively opening evaluation model of colleges, which solved the defects of the traditional class style evaluation model.

On the basis of previous studies, this paper has carried out the improvement of subjective combine with objective combination and departmental distinction. Firstly, relying on big data and using formal processing methods, the matrix composed of class multidimensional data is used as model input. Then, the model algorithm is used to model and analyze the class evaluation object, and the specific values of each index are processed by the algorithm to obtain the weight of each indicator. Finally, the total score of each class is obtained, and the class style of each class is evaluated. The results of the evaluation help to promote the management and development of colleges and universities, and also help to improve the teaching methods.

Whether the mathematical methods used in this paper are most suitable for this project still needs to be verified and we have not yet conducted in-depth research on the degree of association between attributes. We hope that through verification and later maintenance and upgrading, more suitable mathematical methods can be found to solve the problems leftover, so as to make the evaluation results of the evaluation model of class style in Colleges more scientific and reasonable.

REFERENCES

[1] S.L. Li, "Main Factors Affecting the Current Situation of Undergraduate Study Style and Empirical Analysis," Heilongjiang Researches on Higher Education, Vol.28 No.04, April 2009, pp.128-130. (In Chinese)

[2] X.Y. Fu, "Research on key influencing factors of undergraduate science in K University," Kunming: Kunming University of Science and Technology, 2015, p.58. (In Chinese)

[3] H.Q. Guo, "Investigation and analysis of the current situation of psychological survey in colleges and universities in Zhejiang Province," The psychological counseling committee of college students of China Association for Mental Health, "The exploration of "heart" in the past 20 years - Research on the theory and method of mental health education in colleges and universities," The psychological counseling committee of college students of China Association for Mental Health :The psychological counseling committee of college students of China Association for Mental Health, 2010, pp.128-130. (In Chinese)

[4] F. Leng, Z.H. Liu and H.X. Zheng, "Analysis of College Students' Evaluation Model Based on Big Data," China Market, vol.23 No.51, December 2016, pp.216+224. (In Chinese)

[5] W. Liang, "Campus Network User Behavior Analysis System Research and Implementation," Beijing: Beijing Jiaotong University, 2009, p.71. (In Chinese)

[6] C.N. Tang, W. Dong, "Research on the Correlation between the Construction Factors of High-level University Study Style," Journal of Hunan University of Technology (Social Science Edition), vol.13 No.01, February 2008, pp.106-108. (In Chinese)

[7] E.K. Xie, X. Xie and Y.D. Zhang, "Study on the Evaluation Index System of College Students' Study Style Construction Based on Entropy Right," Economic Research Guide, vol.8 No.05, February 2012, pp.227-229. (In Chinese)

[8] F. Tian, W.L. Cao, "Empirical Study on the Academic Index System," Higher Education Development and Evaluation, vol.15 No.02, March 2008, pp.15-21+121. (In Chinese)

[9] J. Liu, T. Wang, "Research on the Evaluation of College Academic Discipline Construction Based on Principal Component Analysis-Taking University A as an Example," Economic Research Guide, vol.8 No.20, July 2012, pp.236-239. (In Chinese)