



An Ideological and Political Education Evaluation Method of University Students Based on Data Mining

Liyan Tu^(✉) and Lan Wu

Inner Mongolia University for the Nationalities, Tongliao 028000, China
tlyimun@163.com

Abstract. The development of big data technology and data mining technology has brought new opportunities for the scientific and innovative development of ideological and political education in colleges and universities. The evaluation of ideological and political education in colleges and universities in the context of big data was studied in this paper. An evaluation method of college students' ideological and political education based on data mining was proposed. The proposed method uses K-means clustering method to analyze the data of the "worker's assessment scale" of the counselor, and can achieve the evaluation of the ideological and political management effect of the counselor. The experimental results show that compared with traditional evaluation methods, the evaluation results of this method are more accurate and objective.

Keywords: Big data · Data mining · Educational evaluation · Objectivity

1 Introduction

Ideological and political education of college students is an important part of higher education. The ideological and political education of college students relates to the realization of the fundamental task of high moral values establishment and people cultivation [1, 2]. With the rapid development of China's economy and society and the further advancement of education reform, the ideological and political education of college students is facing new problems and tests. How to adapt to the new situation is an urgent problem to be solved in the current ideological and political education in colleges and universities. Evaluating the ideological and political education of college students and establishing a scientific evaluation system and evaluation method are effective ways to promote ideological and political education in colleges and universities.

The emergence of big data technology signals that the society has entered the era of data. With the rapid development of modern information technology, data mining plays an increasingly widespread and convenient role in our lives. In universities where network technology is most fully applied, all data left by teachers and students can be collected and analyzed through big data technology [3, 4]. Using data mining techniques to find useful information for education can better innovate teaching methods and methods, and achieve zero-distance personalized education for college students. Therefore, it is of great theoretical and practical significance to explore the optimization

of ideological and political education in universities based on the perspective of big data and data mining. How to effectively apply big data technology and data mining technology to the evaluation of ideological and political education in colleges and universities is a question that needs to be solved.

Since 2013, research on the use of big data technology and data mining technology to promote education development has gradually emerged [5]. Liang et al. discussed how to improve the ideological work of colleges and universities in terms of strengthening data awareness and improving the ability to use big data [6]. Zhang et al. proposed that universities must continuously strengthen the collection, excavation and analysis of student data to ensure the normal implementation of ideological education in universities [7]. Zhou et al. advocates that the use of big data technology can uncover the hidden rules behind university data, and then change the way and decision of education managers [8]. Gu et al. advocates the establishment of big data systems to analyze and process big data, so that the ideological and political education in universities can be transformed from empirical teaching to scientific teaching [9]. These studies show that there has been preliminary exploration and research on the application of big data in the evaluation of ideological and political education in colleges and universities. However, the level of research is not deep enough and there is no specific evaluation method, which requires further study and improvement.

This paper studies the management evaluation of university ideological and political education in the context of big data. At present, the traditional analysis method is based on the evaluation of absolute scores. This method has some defects and deficiencies in the objectivity and accuracy of the evaluation results. The traditional analysis method to evaluate the counselor cannot effectively and properly evaluate the counselor's management effectiveness. For this reason, this paper firstly uses the cluster analysis in data mining technology to establish the evaluation model of university ideological and political education management. Then, the k-means clustering algorithm is used to analyze the counselor's "work evaluation scale" data, which can effectively overcome the defects and shortcomings of traditional analysis methods.

2 K-Means Clustering Algorithm

Cluster analysis is a multivariate statistical analysis to classify samples or indicators [10]. According to the different objects can be divided into sample classification and variable clustering, that is, the classification of the samples and the classification of variables. Clustering is different from classification. Classification has prior knowledge that reference learning. For the classification model, the class label of the sample information is known.

The K-means algorithm is a stepwise clustering analysis of a large sample of data by a user-specified number of categories [11, 12]. The sample data is regarded as a point in the k-dimensional space, distance is used as an index to judge the distance relationship between data, and a higher execution rate is obtained at the expense of wasting multiple solutions. The basic idea of K-means algorithm is to aggregate each sample into its nearest average class, that is, to initially classify the data and then gradually adjust it to obtain the final classification result.

The K-means algorithm is implemented as follows. The dataset that needs to be clustered is

$$S = \{x_1, x_2, \dots, x_n\} \tag{1}$$

where x_i is the p dimensional data, p is the number of attributes. $\{c_1, c_2, \dots, c_k\}$ represents k cluster centers. The degree of similarity between any two data objects in the data set is represented by Euclidean distance as

$$d(x_i, y_j) = \sqrt{\sum_{m=1}^p |x_{im} - x_{jm}|^2} \tag{2}$$

The objective function used to determine whether the clustering result is good or bad is

$$J = \sum_{i=1}^k \sum_{j=1}^n d(x_j, c_i) \tag{3}$$

where the objective function J is a function about the clustering sample and the clustering center. The objective function is the sum of the squared errors of the cluster centers of all the data objects in the data set and the clusters of the respective clusters after the cluster iteration. The process of classical K-means clustering algorithm is shown in Fig. 1.

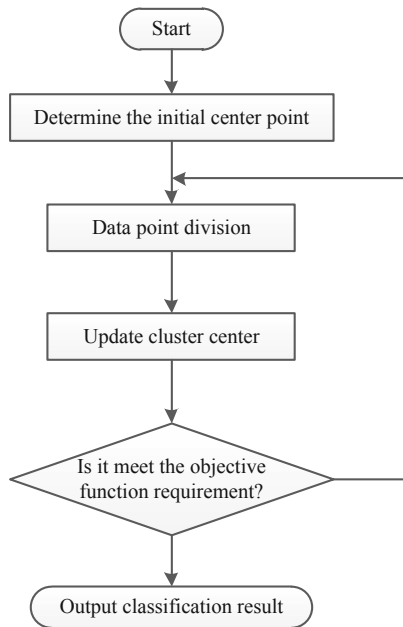


Fig. 1. K-means clustering algorithm

3 Cluster Evaluation Model of Ideological and Political Education

With the continuous development and popularization of higher education in China, more and more high school graduates have entered the university's school, which has brought a lot of influence on the school's teaching and management. In addition, because the study habits are not well developed, the creation of a learning atmosphere has brought great difficulties. The poor learning atmosphere has also exerted a great influence on students' interest in learning. These are the difficulties in the education and management of college counselors. Under normal circumstances, the relevant departments of the school have collected a large amount of data through the usual education and management work. However, at present, the processing of these data is only at a relatively low level of search and simple analysis, and no valuable and instructive information has been mined. How to tap into these "treasures" data and find valuable information among them is a key issue that both the school leaders and the majority of student workers are concerned about.

In order to more specifically understand the work of counselors, the school organizes students' comments on the counselor's work every semester, and fills in a "workbench quantification form" for counselors. How to make better use of it and find valuable information on the education, management, and school management of the counselors is a very meaningful thing for the optimization of ideological and political education in colleges and universities. School leaders and student workers must scientifically judge each stage of student management education, and dig out useful knowledge from it, so that they can achieve further guidance and management requirements.

3.1 Schematic Design

In the past, most of the information on the effectiveness of the counselor's work was used to find large amounts of information. This paper proposes a cluster analysis data mining method to deal with the data information of the effectiveness of the counselor. It can turn a large number of data into a clustering result, and then better use this type of data. The data mining process is shown in Fig. 2.

Step 1: Identify the data mining object and purpose. Clearly defining the problem and finding the goal of data mining is an important first step in the data mining process. Although data mining cannot predict the final result, it can foresee the problems studied. Therefore, it is necessary to avoid blind digging.

Step 2: Data collection. This process has heavy work tasks and it takes more time. In the usual education management, it is necessary to carefully collect data and information. Some of the data can be obtained directly, and some of the data must be obtained through research.

Step 3: Data preprocessing. The process is to convert the collected data set into an analyzable data model. The data model is prepared according to the algorithm. Different algorithms have different requirements for the data model.

Step 4: Data Clustering Mining. Clustering data mining is to divide the data model into multiple groups of similar objects. The process is mainly the input process of the data model and the selection and implementation process of the clustering algorithm.

Step 5: Analysis of clustering results. This process is to analyze and study the multiple group attributes composed of classes after clustering data mining.

Step 6: Application of knowledge. This process integrates the useful information obtained from the research into the counselor's management education. The counselor can use the conclusions to promote teaching management and form a good management policy.

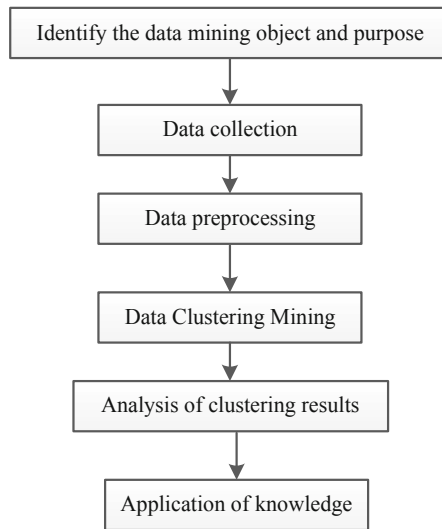


Fig. 2. Data mining process

3.2 Algorithm Implementation

We collected and compiled the “Quantification Form for the Assessment of Counselors’ Work” of the Inner Mongolia University for the Nationalities in 2016, and sorted out 1200 assessment quantification tables concerning the education and management of counselors. Use this data to evaluate the counselor’s ideological and political education management. After data mining of these data, some important conclusions can bring unprecedented guiding value for management and teaching.

According to the job performance of the counselor, it is divided into five categories of “excellent, good, qualified, poor, and poor”. Further measures were taken for the missing records, and finally a quantitative assessment of the work evaluation with complete evaluation records was obtained.

According to the four attributes of “management attitude”, “management ability”, “management method” and “management effect”, the data in the quantification table of the merger work assessment is reorganized. Through the above processing, we have

unified the quantitative assessment of the job evaluation to the four attributes of “management attitude”, “management ability”, “management method” and “management effect”. In the following, we use the clustering mining analysis method to analyze the four attributes of the 1170 data sample information.

By taking the process of preprocessing data sample information, the data samples used for the analysis process are listed in Table 1.

Table 1. Clustered data sample.

Management attitude	Management ability	Management method	Management effect
0.66	0.65	0.56	0.59
0.66	0.64	0.61	0.60
0.34	0.39	0.31	0.33
0.8	0.78	0.83	0.80
...

4 Experimental Results and Analysis

The k-means algorithm was used to analyze 1200 direct sample data and 1170 sample data obtained by data conversion. All sample data contains four types of attributes: management attitude, management ability, management method, and management effectiveness. Data mining clustering was performed on these four attributes respectively, and the initial K value was set to 3. The final mining results are shown in Table 2.

Table 2. Clustering results.

Cluster category	Management attitude	Management ability	Management method	Management effect	Sample number
Better	0.76	0.75	0.74	0.79	36
Medium	0.63	0.58	0.54	0.56	74
Poor	0.34	0.30	0.27	0.30	10

According to the results of the cluster analysis, the final scale distribution of the data included in each cluster is as follows. There are a total of 36 data samples in the better category, and 30% after deleting a standard sample. There are 74 samples in the medium category and 62% in the case of deleting a standard sample. There are a total of 10 data samples for the poor category and 8% for the deletion of a standard sample.

In order to further verify the final results of data mining, we once again obtained 248 people from the Department of Student Affairs to obtain the relevant comprehensive quantitative scores in participating in 10 activities. The total score is 100 points. These data samples are categorized on a scale of 0 to 100 points. A total of three

categories are included: more than 80 points (Better), 80 to 60 points (Medium) and less than 60 points (Poor). The final statistical results are shown in Table 3.

Table 3. Comprehensive quantitative score analysis results.

Cluster category	Sample number	Proportion
Better	512	21%
Medium	1686	68%
Poor	282	11%

The statistical results show that the proportions of more than 80 points, 80 to 60 points and less than 60 points are 21%, 68% and 11% respectively. Comparing the analysis results in Tables 2 and 3, we can see that the clustered evaluation results are basically consistent with the integrated quantitative results, as it shown in Fig. 3. The experimental results confirm that such a data mining model based on the quantification table of the counselor’s job evaluation is a very successful model, which brings certain reference significance and certain guiding significance to the management and education work of college counselors.

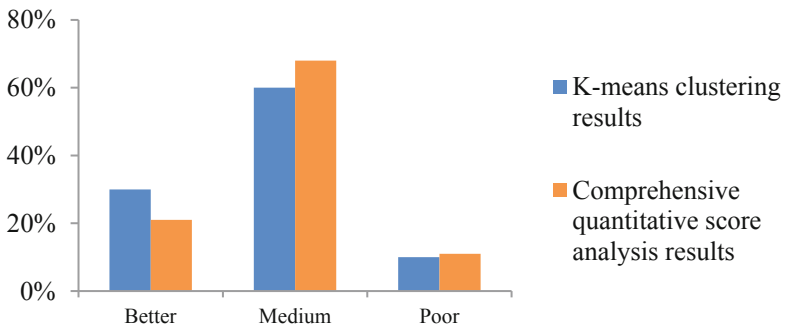


Fig. 3. Comparison of evaluation results

5 Conclusion

The application of clustering methods in data mining to the evaluation of ideological and political education in colleges and universities has been studied in this paper. The data analysis method of “work evaluation scale” of counselors based on K-means clustering was designed. The experimental results show that this method can effectively evaluate the ideological and political education of college counselors, and the evaluation results are more efficient and more objective. This study provides a new idea for the evaluation of ideological and political education in colleges and universities in the context of big data.

Acknowledgements. Inner Mongolia University for Nationalities Ideological and Political Theory Teaching and Research Project (SZ2014009). Inner Mongolia Autonomous Region Science and Technology Innovation Guide Project in 2018: KCBJ2018028.

References

1. Lei, Z.M.: Research on the Innovation of the Ideological and Political Education of College Students in the Background of the Omnimedia. *Theory & Practice of Education* (2016)
2. Wang, J.F., Yao, J., Guo, M.L.: Research on Cooperative Training Mechanism of College Students' Innovation and Entrepreneurship Education. *Value Engineering* (2018)
3. Zhao, J.L., Fan, S., Hu, D.: Business challenges and research directions of management analytics in the big data era. *J. Manag. Anal.* **1**(3), 169–174 (2014)
4. Samier, E.A.: Education in a troubled era of disenchantment: the emergence of a new Zeitgeist. *J. Educ. Adm. Hist.* **50**, 1–10 (2018)
5. Sakaue, K.: Informal fee charge and school choice under a free primary education policy: panel data evidence from rural Uganda. *Int. J. Educ. Dev.* **62**, 112–127 (2018)
6. Liang, J.F., Yuan, Z.H.: Adaptation and innovation: ideological and political education in colleges and universities in the era of big data. *Stud. Educ. Res.* **6**, 63–67 (2013)
7. Zhang, R., Dong, Z., Xia, X.: Innovation and exploration of ideological and political education in universities in the era of big data. *Sch. Party Build. Ideol. Educ.* **10**, 67–69 (2014)
8. Zhou, S.M.: Talking about the application of big data in ideological and political education in colleges and universities. *Stud. Educ. Res.* **5**, 124–126 (2014)
9. Gu, Q.L., Xu, J.Z.: Study on ideological and political education of college students based on big data thinking. *Theory Res.* **14**, 248–249 (2015)
10. Carvalho, A., Dall'Osso, M., Dorigo, T., et al.: Higgs pair production: choosing benchmarks with cluster analysis. *J. High Energy Phys.* (4), 126 (2016)
11. Hasib, A.A., Cebrian, J.M., Natvig, L.: A vectorized k-means, algorithm for compressed datasets: design and experimental analysis. *J. Supercomput.* **3**, 1–24 (2018)
12. Kwon, K., Shin, J.W., Kim, N.S.: Incremental basis estimation adopting global k-means algorithm for NMF-based noise reduction. *Appl. Acoust.* **129**, 277–283 (2018)