Analysis and Research on Ideological and Political Learning Outcomes of College Students Based on Education Data Mining

Jiang Xiaoguo

Chengdu Agricultural College, Chengdu 611130, China

Keywords: Data Mining, Clustering Analysis, Partition Method, K-Means Algorithm

Abstract: It is Still in the Exploration Stage for the Application of the Existing Education Management System in the Ideological and Political Education Management of Colleges and Universities, Especially for the Research on the Ideological and Political Education Management of Colleges and Universities in the Management of Ideological and Political Education. This Paper Aims At the Topic of Education Management of Ideological and Political Thoughts in Colleges and Universities in Education Management of Ideological and Political Thoughts. Combining with the Clustering Analysis in Data Mining Technology, This Paper Carries out Research of the Ideological and Political Learning Results of College Students, Which is the Most Important Link in This Topic. the Traditional Analysis Method is Based on the Evaluation of Calculating the Absolute Score, Which Has Some Defects in the Objectivity and Accuracy of the Evaluation Results. Therefore, This Paper Adopts the Idea of Clustering Analysis in Data Mining and K-Means Algorithm to Analyze the Ideological and Political Learning Achievement Data of College Students, Which Can Effectively Overcome the Defects and Deficiencies of Traditional Analysis Methods.

1. Introduction

With the Continuous Growth of the Information Industry and the Progress of Industrial Technology, the Application Scope of Database Management System is Becoming Larger and Larger, and the Scale of Database to Be Processed by Database Management System is Also Increasing, and People Accumulate More and More Data. in Recent Years, Due to the Gradual Increase in the Number of Students on Campus and the Increasing Number of Teachers, Education Management Database Has Accumulated a Lot of Data. Many Schools Have Researched and Implemented Education Management System Based on Their Actual Situation, and Education Management System of Each University is Different. the System is Generally Composed of Edit Menu Bar, System Backup Maintenance Part, and System Function Interpretation Part. Education Management System Can Realize the Functions of Coordinating Data Information, Researching Histogram, Producing Statements, Printing Conclusions, Etc. [1]. It Has the Characteristics of Simple Operation Procedure, Convenient Use, High Execution Efficiency and Good Analysis Effect, Which is in Line with the Research and Analysis of Teaching Information in the University, and At the Same Time Fills the Loopholes of Education Management System.

However, the Application of the Existing Education Management System in the Ideological and Political Education Management of Colleges and Universities is Still in the Exploratory Stage, Especially the Research on the Ideological and Political Education Management of Colleges and Universities in the Management of Ideological and Political Education. the Traditional Analysis Method is Based on the Manual Calculation of Absolute Score for Comparison Analysis Method, This Method in the Evaluation of the Objectivity, Accuracy and Other Aspects of the Existence of Some Defects and Deficiencies. It Would Be Unfair to Evaluate the Ideological and Political Learning Results of College Students According to the Traditional Analysis Method, and It Would Not Be Effective and Appropriate to Evaluate the Effect of Students' Learning Management.

As an Important Part of Education Management Database, Education Management of Ideological and Political Institutions in Colleges and Universities is an Important Standard to Evaluate the Quality of Education Management of Ideological and Political Institutions, as Well as an Important Sign to Evaluate Whether the Education Management of Ideological and Political

Copyright © (2019) Francis Academic Press, UK 32

Institutions is Effective in Conducting Education Management of Students [2]. Therefore, How to Evaluate the Quality of Education Managers Scientifically, Accurately, Fairly and Fairly Has Been a Research Topic for Education Workers for Many Years. with the Deepening of Education Management System Reform, Especially after the Student Ideological and Political Instructor Management System Becomes the Mainstream, the Evaluation Method of the Management Quality of Ideological and Political Education Managers Has Developed from a Single Meeting Discussion and Manual Evaluation in the Past to the Form of Quantitative Work Assessment Table Widely Used Today.

The Evaluation Method of Education Manager Work Assessment Quantitative Table is Suitable for Vertically Measuring the Degree of the Education Manager's Mastery of the Students under Management. However, When It is Necessary to Compare the Level Differences between Education Learners of Ideology and Politics Horizontally and Measure the Deviation of a Certain Ideological and Political Education Learner from the Overall Management Level, the Information Contained in the Quantitative Table of Work Assessment is Obviously Insufficient [3]. in This Case, We Can Use Systematic Statistical Means to Process the Quantitative Assessment Table of Education Managers of Ideology and Politics, and Obtain a Set of Relatively Abstract Assessment Data Aiming At Specific Assessment Criteria.

Aiming at the above problems, this paper introduces clustering analysis into the application of education management of ideology and politics in colleges and universities. It applies data mining technology to conduct automatic data analysis and mining of a large number of evaluation data to find out useful information, which can effectively overcome the defects and deficiencies of traditional analysis methods.

2. Data Mining Technology and Research Status

Data mining is to extract the knowledge that people are interested in from a large amount of data, which is implicit, previously unknown, but potentially useful information. The extracted knowledge can be expressed as concepts, rules, rules, patterns, constraints, visualization and other forms. Data mining means the process of finding patterns in sets of facts or observations [4]. The objects of data mining can be structured, semi-structured or even heterogeneous data distributed on the network. Data mining is an interdisciplinary subject, which elevates people's application of data from simple query at a low level to knowledge mining from data and provides decision support. The task of data mining is to discover knowledge from a large amount of data. Knowledge, as the result or crystallization of human cognition, includes both empirical knowledge and theoretical knowledge. These knowledge can be directly provided to decision-makers to assist the decision-making process^[9]; Or provide it to the domain experts to revise the existing knowledge system of the experts; It can also be transferred to the knowledge storage mechanism of the application system as a new knowledge, such as expert system, rule base and so on.

Data mining is a process in which multiple steps connect and interact with each other repeatedly. The main steps of data mining are: problem understanding and proposal, data preparation, data collation, model building, evaluation and interpretation. The above steps will not be done in one go, and some or all of them will be repeated. Data mining makes knowledge-based decisions by predicting future trends and behaviors. The knowledge that is mined is also known as a pattern. There are many types of patterns, which can be divided into two categories by function: predictive patterns and descriptive patterns. Mining methods include association analysis, sequence pattern analysis, classification analysis, clustering analysis, etc. The research of data mining technology has a long history abroad. Data mining technology and related decision support system have developed rapidly, and have brought surprising profits to many industries such as retail industry, public service industry, banking and securities industry, and many schools and research of data mining technology [5]. American express detection using neural network database of hundreds of millions of records, how to identify individual consumers where card transaction, the obtained each card users' propensity to buy value, according to these values, the American express company will

individual cardholder purchases matching sellers of history and the relationship and the information attached at the back of the monthly report, it saves costs and provides more value to the cardholder's analysis.

3. Application of Data Mining Technology in Education Management of Ideological and Political Learning of College Students

3.1 Status and Nature of Education in Ideological and Political Learning of College Students

According to the opinions of the propaganda department of the CPC central committee and the ministry of education on further strengthening and improving ideological and political theory courses in colleges and universities as well as the implementation plan, ideological and political theory courses in colleges and universities undertake the task of systematically implementing Marxist theory education on college students and are the main channel of carrying out ideological and political and political education on college students.

The state has always attached great importance to education teaching of ideological and political theory courses. The ministry of education has continuously raised the requirements for the reform of ideological and political courses, so that ideological and political courses can truly become excellent courses that students really love, benefit from and remember all their lives. Through actively exploring the reform of teaching methods in colleges and universities and ideological and political course teachers, ideological and political course teaching has achieved obvious results and been constantly affirmed and recognized by students [6]. According to the survey results, 26. Seven percent of college students oppose politics

The teaching content of the course is very familiar, 63% of college students have a good understanding of the teaching content of ideological and political courses, 7.5% of college students do not fully understand the content of ideological and political education, only 2.8% of college students do not know the content of ideological and political education.

Ideological and political teaching explores the implementation of problem-oriented teaching method to solve students' ideological confusion. Collect and sort out students' ideological confusion problems through multiple channels, design teaching objectives around the ideological confusion of young students, and carry out thematic teaching design. In this way, the traditional teaching method of "reading according to the text" and "filling the classroom" has been changed, and the pertinence of teaching has been effectively improved. By encouraging students to move from small classes in school to large classes in society and finding and analyzing problems in social observation and investigation, students' understanding of the theory of ideological and political courses has been effectively improved. The survey results show that college students' attitudes towards ideological and political education are relatively recognized and affirmed.

3.2 Problems Existing in Ideological and Political Learning of College Students

In the learning process of ideological and political courses, quite a number of students have two main problems: first, students' learning objectives of ideological and political courses are deficient or vague. The education thinking pattern formed in the long-term study leads to the lack of vision expansion awareness and learning initiative in the study of ideological and political courses. Learning methods are simple, mostly for the purpose of rote memorization, to cope with the test, the lack of understanding and use. Secondly, students have obvious slackness in the study of ideological and political courses, lack of fighting spirit, poor mental state, and lack of independent thinking ability, so they have deep prejudice against ideological and political courses. The students are not enthusiastic about the study of ideological and political courses is passive, reluctant and helpless in learning attitude. The study of ideological and political courses is regarded as an external burden and a "credit" task that must be completed. There are serious problems of students think that they are not interested in ideological and political courses and they do not want to learn them, let

alone unwilling to learn them, which seriously affects the practical effect of ideological and political courses.

3.3 Analysis and Solutions of Ideological and Political Learning Outcomes of College Students

K-means algorithm is a typical distance-based clustering algorithm. Distance is used as an evaluation index of similarity. That is to say, the closer two objects are, the greater their similarity will be. This algorithm considers that clusters are composed of objects close to each other, so it takes compact and independent clusters as the final target [7]. The selection of k initial clustering center points has a great impact on the clustering results, because in the first step of this algorithm, arbitrary k objects are randomly selected as the initial clustering center to initially represent a cluster. In each iteration, the algorithm reasserts each object to the nearest cluster according to its distance from each cluster center for each remaining object in the dataset. When all data objects are examined, an iterative operation is completed and the new clustering center is calculated. If the value of J does not change before and after one iteration, the algorithm has converged.

In the past, most of the data information about the ideological and political learning outcomes of college students used to find a large number of materials. In this paper, the data mining method of clustering analysis is proposed to process the data information of ideological and political learning results of college students, which can transform a large number of data into clustering results and make better use of such data.

3.4 System Process Model

When selecting the software process model, the requirements of the system are relatively clear and not easy to change, so waterfall model was selected as the process model in the early stage of development [8]. Waterfall model is widely used in the development of software projects. It divides the software development process into distinct and independent stages such as description, design and development validation, and is applicable to software projects with clear requirements and not easy to change. However, in the process of development, due to my limited ability of demand analysis, I did not do enough demand analysis of the task book, and I often needed to discuss with my tutor about the requirements, which led to frequent changes in the requirements, and the demand analysis stage was excessively prolonged. Moreover, the waterfall model is clearly not appropriate when colleges and mentors need to check progress frequently, so incremental development models are used instead.

3.5 System Architecture

Container architectures are commonly used when a system needs to generate large amounts of data and persist it. The advantage of container architecture is that the functional components are independent of each other, they do not need to know about the existence of other components, and are not affected by other components at runtime. Changes to one component can be propagated to all other components, and all data can be managed consistently. In this system, the adoption of container architecture can make the mining results be kept for a long time, and the ideological and political learning information of college students can be managed conveniently. The interaction between the components is not strong, and the function of other components can be prevented from being affected by errors of one component.

The system includes five components: import data, data preprocessing, association rules, classification and clustering [9]. The student information Shared by all components and the mining results generated by each component are stored in the project container, which improves the efficiency of data transmission. There are several sub-functional modules under each component

First, randomly select K objects as the initial clustering center. Then calculate the distance between each object and each seed cluster center, and assign each object to the nearest cluster center. The cluster centers and the objects assigned to them represent a cluster. Once all objects are allocated, the cluster center of each cluster is recalculated based on the existing objects in the cluster. This process repeats until some termination condition is met. Termination conditions may be any of the following:

1) no (or minimum) objects are reassigned to different clusters.

2) there is no (or minimum) change in the clustering center.

3) the sum of squares of errors is locally minimum.

Divide the sample set into k classes, which minimizes the following formula

$$E = \sum_{j=1}^{k} \sum_{x_i \in \omega_j} \left\| x_i - m_j \right\|^2$$
(1)

 m_j is the center of mass of the JTH class. If you want to design an algorithm to obtain the global optimal solution, you must complete C(n,k) sub-clustering and find the clustering result that makes E the smallest. K-means clustering is an algorithm to find the local optimal solution.

Distance measurement is carried out by using Euclidean distance. The formula of central European distance in n-dimensional space is:

$$D = \left(\sum_{i=1}^{n} \left(x_{i} - m_{i}\right)^{2}\right)^{1/2}$$
(2)

This paper extracts the characteristics of ideological and political learning achievements based on the gray scale symbiosis matrix of learning achievements [10]. According to the symbiosis matrix, 16 kinds of characteristic statistics such as entropy, contrast, energy, correlation and variance can be calculated to extract the information in ideological and political learning results. The features of contrast, energy, correlation and homogeneity are selected in this paper. Feature extraction steps are as follows:

(1) Set the window matrix with radius d and size (2d+1) * (2d+1), calculate the gray scale symbiosis matrix in the window by traversing the score, and map it to the position represented by the center of the window.

(2) Based on gray scale symbiosis matrix, the contrast, energy, correlation and homogeneity of four texture features are calculated by using formula (3) and formula (4), and the values are returned to the corresponding pixel center.

$$f = CON = \sum_{n=0}^{G-1} n^2 \left[\sum_{i=1}^{g} \sum_{j=1}^{G} p(i, j)\right] \qquad \left|i - j\right| = n \tag{3}$$

$$f = ENG = \sum_{i=1}^{G} \sum_{j=1}^{G} p^{2}(i, j)$$
(4)

$$f = COR = \frac{\sum_{i=1}^{G} \sum_{j=1}^{G} (i \cdot j) p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y}$$
(5)

$$f = HOM = \sum_{i=1}^{G} \sum_{j=1}^{G} \frac{p(i,j)}{1 + |i - j|}$$
(6)

Contrast can be understood as the contrast of achievement; Energy is a measure of the uniformity of gray distribution of grades. The image presents coarse texture, and the value of energy is relatively large. Correlation is a measure of linear relationship of gray level, which reflects the extension length of a certain gray level value along a certain direction. Homogeneity is a measure of local grayscale uniformity. If the local grayscale is uniform, the value of homogeneity is larger. In addition, in view of the window on the edge of the grades can't get the complete information and cannot get the texture eigenvalue problem, considering the contact involved in calculation of characteristic value, this paper adopted the regional average algorithm, to the radius of d: the window of the 3 matrix as an example, the edge pixel values approximate is expressed as the average value of texture characteristics associated with them. That is:

$$X = mean(\sum_{i=i-d}^{i+d} \sum_{j=j-d}^{j+d} p(i,j))$$
(7)

Where (i, j) is the spatial coordinate of x.

Randomly select K points as the initial clustering center, calculate the distance between each sample and the clustering center, classify the sample to the class where the nearest clustering center is located, and calculate the new clustering center for the adjusted new class. If there is no change in the clustering center of the two adjacent clusters, it means that the sample adjustment is finished and the clustering criterion function has converged.

In k-means clustering algorithm, an important step is to select the initial clustering center. Generally, K samples of the sample set to be clustered are randomly selected. The performance of clustering is related to the selection of the initial clustering center. Once these K samples are selected irrationally, the complexity of calculation will be increased and the clustering process will be misdirected to obtain unreasonable clustering results. The rough set theory provides the number and mean of initial classes required by k-means clustering, which improves the clustering efficiency and classification accuracy.

4. Conclusion

This paper aims at the research topic of education management of ideological and political thoughts in colleges and universities in education management of ideological and political thoughts. Combining with the clustering analysis in data mining technology, this paper carries out research and analyzes the ideological and political learning results of college students, which is the most important link in this topic. The traditional analysis method is based on the evaluation of calculating the absolute score. This method has some defects in the objectivity and accuracy of the evaluation results. Therefore, this paper adopts the idea of clustering analysis in data mining and k-means algorithm to analyze the ideological and political learning achievement data of college students, which can effectively overcome the defects and deficiencies of traditional analysis methods.

References

[1] Goetz, E. T, Cifuentes, et al. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis[J]. Computers & Education. no. 70(1). pp. 29-40.

[2] Steele D J, Medder J D, Turner P. A. (2010). Comparison of learning outcomes and attitudes in student- versus faculty-led problem-based learning: an experimental study[J]. Medical Education. no.34(1). pp23-29.

[3] Curter J E, Esche S K, Chassapis C, et al. (2011). Process and learning outcomes from remotely-operated, simulated, and hands-on student laboratories[J]. Computers & Education. no. 57(3). pp.0-2067.

[4] Gauthier A, Corrin M, Jenkinson J. (2015). Exploring the influence of game design on learning and voluntary use in an online vascular anatomy study aid[J]. Computers & Education, no. 87. pp24-34.

[5] GLEASON, Mary, MCCREARY, et al. (2010). Science-based and stakeholder-driven marine protected area network planning: A successful case study from north central California[J]. Ocean & Coastal Management, no.53(2). pp52-68.

[6] Duis J M, Schafer L, Nussbaum S, et al. (2013). A Process for Developing Introductory Science Laboratory Learning Goals to Enhance Student Learning and Instructional Alignment[J]. Journal of Chemical Education. no. 90(9). pp1144-1150.

[7] Getha-Eby T J, Beery T, O'Brien B, et al. (2015). Student learning outcomes in response to

concept-based teaching[J]. Journal of Nursing Education. no.54(4). pp193-200.

[8] Morin K H, Black J P. (2015). Student Learning Outcomes[J]. Journal of Nursing Education. no. 54. pp75-76,78.

[9] Rockinson Szapkiw A J, Courduff J, Carter K, et al. (2013). Electronic versus traditional print textbooks: A comparison study on the influence of university students' learning. [J]. Computers & Education. no. 63(2). pp.259-266.

[10] Miller C L, Louise L F. (2009). Student learning outcomes after integration of quality and safety education competencies into a senior-level critical care course[J]. Journal of Nursing Education. no. 48(12). pp678-685.