Research on the application of university teaching management evaluation system based on Apriori algorithm

Yang Zhao1

1Business management and human resources, Wuhan Business University, Wuhan, Hubei, 430000, China

author's e-mail: 20150282@wbu.edu.cn

Abstract. University teaching quality evaluation is supported by big data, which can collect rich and comprehensive data materials as the basis of evaluation. However, it is difficult to find accurate and effective information from the huge data materials. This paper gradually establish "the whole process of implementation, multi stakeholder participation in the teaching quality evaluation system". This paper take "the collection of data to achieve standardization, the application of data to achieve diversification and the management of data to achieve institutionalization" management measures. This paper make full use of the advantages of big data to improve the teaching quality of undergraduate teaching in local colleges and universities. This paper mainly expounds the Apriori algorithm in association rules of data mining, and uses it to further analyze the teaching evaluation results, trying to dig out the potential valuable information. These information can be used as a reference for further decision-making of teaching management department, so as to effectively improve the quality of education and teaching and accelerate the realization of education modernization.

1. Introduction

In order to promote the implementation of the overall deployment of education informatization in the outline of the national medium and long term education reform and development plan (2010-2020), the Ministry of education has organized and compiled the ten year development plan of education informatization (2011-2020). The plan puts forward a clear goal for the future development of China's Education Informatization: the time node for the full completion of the education informatization system is 2020. At that time, it is required that the education informatization system must adapt to the national education modernization development goal; to achieve an environment where everyone can enjoy high-quality information education resources, We should create a service system that can support a learning society by realizing informatization; broadband network should be basically spread all over the education environment of any level and category to truly achieve full coverage; the level of informatization education management has been significantly improved, the development level of the integration of education and technology has been significantly improved, and the level of education informatization has reached a leading position in the world, Therefore, it is an inevitable trend for local colleges and universities to give full play to the advantages of big data to build teaching quality evaluation system.

At present, the traditional teaching evaluation system only makes simple statistics on the collected data to form the teaching evaluation results, and does not make full use of the more valuable information implied in the data to promote the teaching improvement. It is difficult to explain which factors are related to the teaching level, let alone effectively improve the teaching quality. Data mining
technology can make a comprehensive and in-depth analysis of the data. By looking for the hidden relationship behind the data, it can dig out the potential useful information, help the teaching management department to make correct decisions and judgments, so as to effectively improve the quality of education and teaching, and further accelerate the realization of the modernization of higher education.

At present, data mining has become a very popular research field, its technology has been widely used in the telecommunications industry, financial industry, health care, natural science and other aspects. In short, data mining is to mine or obtain potential information from massive, incomplete and random data, which is a meaningful part that people have not found in advance. Data mining is a process of discovering useful patterns in data. The purpose is to use the patterns found to help explain current behavior or predict future results.

2. Establish the evaluation concept of student-centered and data reference

In the information age of the 21st century, information technology presents digitization, networking and intellectualization, forming the basic characteristics of openness, sharing, communication and cooperation. Big data as a new technology architecture obtain massive data through high-speed capture and discovery analysis. Its connotation can be summarized as large amount of data, complex data categories and fast data update, With the help of big data, local colleges and universities have many choices in teaching quality evaluation. However, higher education administrators should not forget the original intention of evaluation and lose their way in massive data, We need to think about making good use of big data as the basis of teaching quality evaluation system to fully serve the growth and success of students. To pay attention to teaching quality is to pay attention to the most direct beneficiaries of Education -- students, that is to pay attention to their learning and development needs, and finally let higher education return to its ultimate value, that is, quality personnel training. Teaching evaluation can only be teacher centered, teaching mode design can only highlight the role of teachers, and students' role as learners is often ignored. Information technology enables teachers to get rid of repetitive and monotonous teaching activities, spend more time and energy to study efficient classroom, spend more time to study teaching methods, integrate various teaching resources and use a variety of teaching means, It is the development of information technology that has brought about fundamental changes in education itself. We should get rid of one-way teacher speaking and students listening, break the teaching mode that the classroom is the only classroom, jump out of the original rigid and fixed space-time education paradigm, and extend the teaching place from the classroom to a wider space under the classroom, even outside the school, Teachers have become developers and organizers of teaching activities from knowledge transmitters. Teachers have changed from mechanical use of educational resources to active integration, development and organization of various teaching resources, thus changing the role and status of students in teaching and driving the change of students' learning habits, learning content and learning motivation in an all-round way.

Under the background of big data of education informatization, teachers can make exploratory analysis on students' course learning, homework completion and course testing, collect feedback information, and timely put forward targeted improvement suggestions on students' existing problems; carry out effective intervention guidance, improve teaching and meet students' diversified learning needs, Education relies on the support of information technology to strengthen the leading role and leading function of teachers, so as to change the status of students in learning, from passive receivers of knowledge learning to active explorers of knowledge, Using advanced information technology to participate in educational activities, enrich learning content, sublimate teaching value and improve teaching effect.

Data mining is originated from knowledge discovery in database 8 (KDD). It is a process of extracting latent and useful information and knowledge from a large number of, incomplete, noisy, fuzzy and random practical data. Through data mining, some valuable knowledge or high-level information in teaching management can be extracted from teaching management database to provide decision support for teaching managers and improve teaching quality.
The data mining process consists of data preparation stage and data mining stage. The main work of the data preparation stage includes the collection, extraction and transformation of the original data. In this paper, the original data mainly refers to a large number of students' teaching evaluation information and students' scores and so on. Data preprocessing is an indispensable link in the process of knowledge discovery. It can improve the quality of data, and thus help to improve the accuracy and performance of data mining. In the process of data mining, it is very important to select the appropriate mining algorithm for the formation of mining results and knowledge.

Teaching evaluation is an important part of teaching process. It is a process of studying the value of teachers' teaching and students' learning. Therefore, a complete teaching evaluation should include the evaluation of teachers, students, teaching content, teaching methods, teaching environment, teaching management and many other factors in the teaching process, especially the evaluation of students' learning effect and the evaluation of teachers' teaching process. The evaluation of students' learning effect can often be made through examination or test. The teaching evaluation of teachers should include teachers' peer evaluation and students' evaluation of teachers. However, the current teaching evaluation mostly adopts the method of students' evaluation. Students' evaluation is affected by subjective factors and can not fully reflect the actual teaching effect. In this way, only one-sided evaluation results can be obtained and the guiding role of teaching evaluation can not be fully played. Therefore, it is necessary for us to apply data mining technology to teaching evaluation in Colleges and universities, to help college teachers and administrators analyze the factors affecting teaching quality, to find out the problems existing in teaching management, to provide decision support for managers, and to provide important basis for improving teaching methods.

Data mining has high requirements for data, which requires data to be clean, accurate and complete. However, there are many problems in the original teaching data. There are many teaching departments in Colleges and universities, often different teaching departments use different databases and file systems, the data is messy, lack of unified standards and definitions; some of the same transactions may have multiple physical descriptions in the database, there is redundancy; in addition, there are data loss, uncertainty and incompleteness due to design defects or human reasons. Data cleaning is mainly used to remove the noise data in the data source and the data that has nothing to do with the mining topic.

3. Construction of teaching management execution evaluation model

Data integration is to merge multiple files or heterogeneous data in database, such as teacher information table, student information table, student score table, student evaluation table, etc. In addition, data integration helps to solve the problem of semantic ambiguity, eliminate ambiguity and unify data types. For example, "multimedia technology foundation" is called "multimedia" in some tables, but it actually refers to the same course. At this stage, it is required to unify it.

This step will be responsible for narrowing the scope of data and improving the quality of data mining. With the continuous increase of data in data warehouse, in the process of data mining, the time of data mining is longer and longer, and there are more and more rules. End users will face a mountain of rules, and even some rules can not be mined because of the "dilution" of the overall data. Therefore, we need to mine different parts according to the direction of interest, so as to improve the speed of mining. Finally, the data is transformed into a form suitable for data mining, which mainly involves smoothing, clustering, generalization, normalization and attribute construction of data transformation. Data mining is a basic step of knowledge mining, which needs to choose different mining methods according to different tasks. Clustering is based on similarity, according to the principle of "maximizing class similarity, minimizing class similarity", using mathematical methods according to some similarity or difference index, quantitatively determine the relationship between samples, and cluster samples according to the degree of this relationship. Therefore, after clustering, the objects in the same class have high similarity, but are different from the objects in other classes. For example, through the clustering of teachers, we can find out what kind of teachers are most popular with students, and what are excellent teachers, good teachers and qualified teachers.
There are many comprehensive evaluation methods, such as rough set theory identification method, analytic hierarchy process, AHP fuzzy evaluation method, expert scoring method and so on. AHP fuzzy evaluation method is widely used because it is easy to operate and calculate. In this paper, AHP fuzzy evaluation method is used to build the evaluation model of teaching management execution.

3.1 weight calculation method

In order to facilitate the evaluation, it is necessary to transform the evaluation results into measurable indicators and use quantitative analysis method for comprehensive evaluation. The key of comprehensive evaluation is to determine the weight of each evaluation index. In the hierarchical structure model, the scale method is used to compare the evaluation indexes to determine the weight, and the judgment matrix $B'$ is obtained.

$$B' = \begin{bmatrix} b_{11} & ... & b_{1n} \\ ... & ... & ... \\ b_{m1} & ... & b_{mn} \end{bmatrix}$$  \hspace{1cm} (1)

The square root method is used to calculate the weight value, and the following formula is used for calculation:

$$B'W = \tau_{\text{max}}W$$  \hspace{1cm} (2)

Where, $\tau_{\text{max}}W$ is the maximum eigenvalue, $B'$ is the judgment matrix, $W$ is the weight of single factor sorting, $W$ is the weight of single factor sorting.

3.2 membership calculation and comprehensive evaluation

Construct factor set $u = \{U_1, U_2, U_3\}$, comment set $V = \{V_1, V_2, V_3, V_4, V_5\} = \{90, 80, 70, 60, 0\}$. Experts are invited to score each evaluation index of different evaluation sets, and the proportion of different factor sets is counted, that is, the membership degree $r$ of each evaluation set of factor sets.

$$R = \begin{bmatrix} r_{11} & ... & r_{1n} \\ ... & ... & ... \\ r_{m1} & ... & r_{mn} \end{bmatrix}$$  \hspace{1cm} (3)

According to the membership matrix and weight matrix, and using the formula for comprehensive evaluation.

$$B' = W \cdot R$$  \hspace{1cm} (4)

4. Data mining process of teaching evaluation system

Classification algorithm is one of the most important data mining algorithms. The main classification algorithms are decision tree algorithm and association rule algorithm.

Then classification algorithm, k-nearest neighbor classification algorithm, Bayesian classification algorithm and classification algorithm based on fuzzy logic, genetic algorithm, rough set and neural network. Data mining is a process of continuous feedback. Generally, it can be divided into three stages: data preparation, data mining, pattern evaluation and representation. (1) Data preparation and data preprocessing. The precondition of data mining is the preparation of data. The evaluation data of teachers come from the system database, including peer evaluation, student evaluation, teacher self-evaluation and expert evaluation. This part of data can be obtained by reading the database table. On the other hand, it is the data obtained by questionnaire. The preprocessing of data will transform the sample set data into a format suitable for data mining. (2) Data mining. First of all, according to the nature of the data, choose the appropriate technology in the neural network, induction technology, clustering analysis technology, association analysis technology and many other algorithms. Secondly, select the appropriate algorithm, such as: ID3 algorithm, BP algorithm, Apriori algorithm. Next, use the selected technology and algorithm to mine the preprocessed data. (3) Pattern evaluation and representation. Through data mining, the original data can be transformed into a form that is easier to understand and clear relationship. Statistical methods are used to evaluate the analysis results, so as to get the most suitable model, and predict the possible future situation for decision-makers to make decisions.
Teaching evaluation is an important part of the teaching process, is a scientific and reasonable analysis method, mainly studies the value process of teachers' teaching and students' learning. Therefore, the data mining technology is applied to the teaching evaluation of colleges and universities. From a large number of data analysis, we find the key factors that affect the teaching evaluation results of students, experts and peers, and analyze the relationship between teaching behavior and teaching effect. To help college teachers and administrators analyze the factors affecting the quality of teaching, find out the problems in teaching management, provide decision support for teaching managers, and provide basis for improving teaching methods. (1) The application of data mining in teaching evaluation is conducive to the establishment of a scientific and feasible teaching quality evaluation index system. Through the special department evaluation, we can have an accurate value judgment of teaching quality. At present, there is no reasonable way to evaluate the teaching level in Colleges and universities, and the evaluation results are not reasonable and can not achieve the expected effect. Therefore, through data mining technology, to judge the teaching level of teachers, promote teachers to improve teaching technology, improve the quality of teaching. (2) Introducing data mining technology into the process of teaching evaluation can guide teaching management and improve teaching quality. Using scientific means to analyze and deal with the information obtained, find out the centralized characteristics of teachers with good teaching quality. The home manager can correctly and timely take active and powerful measures for guidance, so as to promote the improvement of the quality of teaching work. (3) Introducing data mining technology into the process of teaching evaluation can promote teaching reform. Through data mining, teachers can judge whether the teaching methods and means are reasonable and whether the key points and difficulties of teaching are clear according to the data, so as to adjust the teaching strategies and improve the teaching means. Teachers should change their teaching ideas and optimize the teaching process as soon as possible.

Decision tree is a method to approximate the value of discrete function, and it is also one of the most widely used inductive reasoning algorithms. The decision tree uses top-down recursion to generate a tree structure similar to the flow chart. A tree graph used to represent a series of judgment processes that people make in order to make a decision. This method is used to show rules such as "under what conditions, what value will be obtained". The decision tree consists of decision nodes, branches and leaves. The root node of the tree represents the beginning of the classification, the leaf node represents the end of an instance, the middle node represents an attribute in the corresponding instance, and the edge represents the possible attribute value of an attribute. In the process of traversing the decision tree from top to bottom, there will be a test in each node. Different test results on each node lead to different branches, and finally reach a leaf node. In the teaching management of colleges and universities, if we want to analyze which factors have an impact on students' performance, we can use decision tree induction, decision tree algorithm and so on.

According to the general steps of data mining, we design the data mining flow chart of teaching evaluation system. The teaching evaluation database includes course information, students' evaluation of teaching, test paper analysis data, expert evaluation data, etc. It is used in the database after data preprocessing to mine data, output rules, analyze the output results, and explain the results by the
teaching management experts. Finally, it forms a report to feed back to the teachers and teaching management personnel. The data mining flow chart of teaching evaluation system is shown in Figure 1.

Apriori algorithm is a frequent itemset algorithm for mining association rules. Its core idea is to mine frequent itemsets through two stages: candidate set generation and downward closure detection. And the algorithm has been widely used in business, network security and other fields. In the apriori algorithm, the basic idea of finding the maximum item set is: the algorithm needs to process the data set in multiple steps. The first step is to count the frequency of all the item sets with one element, and find out those item sets that are not less than the minimum support, that is, the one-dimensional maximum item set. Start the loop from the second step until no more maximum project sets are generated. The circulation process is as follows: in step k, according to the (K ~ 1) dimension maximum item set generated in step k-1, the candidate item set is generated, and then the database is searched to get the item set support of the candidate item set, which is compared with the minimum support, so as to find the k-dimension maximum item set. We extract more relevant data from the database of teaching evaluation system. First, we preprocess the data, and then we mine the data by using the apriori algorithm. The pattern is found to be backward and the rules are output. The teaching management experts will further analyze and explain the mining results with reference to the students' scores, test paper analysis and expert evaluation data, and finally feed back the information. For teachers and administrators, teachers can refer to the evaluation report to improve their teaching work, and administrators can also use feedback as a reference for teaching reform and next decision-making.

5. Conclusion
As a new data analysis technology, data mining in the evaluation of teaching quality in Colleges and universities, analyzes the data related to teaching in Colleges and universities, so as to find out the problems existing in college teaching and find out the key factors affecting the teaching effect, which can not only improve efficiency, reduce costs, but also be more conducive to data sharing, ensure the safety of data, and better serve the students. The teaching and development of colleges and universities. At the same time, it is also conducive to further promote the development of university education information construction.

References
[1] Clark Glymour, David Madigan, Daryl Pregibon, Padhraic Smyth. Statistical Themes and Lessons for Data Mining. Data Mining & Knowledge Discovery, 1997, 1(1).
[2] Han Jiawei. Data mining techniques. ACM SIGMOD Record, 1996.
[3] Imielsniki T, Agrawal R. Database mining: a performance perspective. IEEE Transactions on Knowledge & Data Engineering, 1993, 5(6).
[4] Myra Spieioonlou. The Lobarious way from data mining to web mining. hat Journal of Comp Sys; scid&Eng, SpecM Issue on “Semantics of the web”, 1999.
[5] Jaideep Srivastava; Robert cooley, Mukund Deshande etal Web Usage Mining: Discovery and Applications of usage patterns from web Data. SIGKDD Explorations, 2000.