



Research on the Cultivation of University Students' Media Literacy Based on DM of Student Behavior

Yushuai Qi^(✉)

Suzhou Vocational University, Suzhou 215104, Jiangsu, China
zxbtianyi@vip.com

Abstract. With the improvement of society nowadays, the amount of message is increasing day by day. How to find useful data in the huge data becomes something that must be considered. With the improvement of message technique, the application of message technique in teaching has become more and more extensive, and various schooling and teaching platforms have appeared. With the rapid improvement of online schooling platforms, a lot of data has been accumulated. The method used in this paper is DM, and schooling is one of the elements of a country's progress. Mining is called schoolingal message mining in schooling. The focus of schoolingal DM is to apply reasonable and fast methods to discover knowledge from schoolingal databanks. Through cluster analysis and other methods, the registration message of students when they log on the website and the results of each subject are analyzed, so as to classify students, discover students' interests, use the classification results to scientifically group various students, and form online learning exchange groups. Learn more in middle school. Through the research of this paper, we can find that the research on the learning experience and learning process of smarter classroom is increasing year by year.

Keywords: Student Behavior · DM · University Students' Media Literacy Training

1 Introduction

Network schooling is an advanced modern distance schooling method that arises on the basis of computer-aided schooling with the improvement of computer network technique. While the scale of schooling is expanding, a large amount of data is generated in the schooling process, making it more and more difficult for schoolingal decision-makers to grasp, and traditional data processing methods have been unable to cope with the accumulated data ocean [1]. Informatization schooling provides a strong guarantee for supporting the improvement strategy of schools, and has become an effective means to enhance the comprehensive strength and competitiveness of universitys and universities. Education informatization is a vital step to realize the modernization of schooling, which is conducive to improving the quality of all citizens and promoting the cultivation of innovative talents [2]. The schooling cloud platform is the product of the combination

of cloud computing and schooling. It is a high-speed, high-quality metropolitan area network that can reach every class in every school. The current manual data analysis has been unable to analyze large-scale data, so we need to find a new method to analyze the data. Data mining can help us gain a deeper understanding of data. How to more effectively discover and utilize the value of data, and let the data feed us useful message and laws, is a problem we have to face in this rapidly changing era, would be a huge waste [3]. Educational DM is the use of DM methods to extract valuable message from schooling data, and the message obtained through mining can help educators make better decisions and services. The platform is to virtualize the teaching resources and store them in the resource library to provide cloud services to teachers and students [4]. Online schooling overcomes the limitations of traditional schooling in terms of time, space and real-time interaction between the two sides of the teaching, so that learners are more flexible and free in learning content, learning form, learning time, learning place, etc., and fully mobilize learners. Positivity. These characteristics undoubtedly make online schooling the best way to achieve lifelong schooling [5].

2 Research on Theoretical Algorithms

2.1 Clustering Rules

Clustering is the grouping of a set of physical or abstract objects into categories by similarity, also known as “unguided classification”. Clustering refers to analyzing the relationship between data objects in a data set, and grouping data sets according to their degree of similarity, so that the instances in the group are similar, and the instances in various groups are various. The purpose is to make the distance between objects in the same category only as small as possible, while the distance between objects in various categories is as large as possible [6]. The greater the difference between groups, the higher the degree of similarity within the group, and the better the clustering effect. Compared with the classification problem, the instances of cluster analysis have no category labels, and it is impossible to develop models with known categories and apply them to uncategorized data objects. Therefore, clustering is also called unsupervised analysis. The “class” in cluster analysis is various from the “class” of classification, and a more accurate translation should be “cluster” [8]. The core problem of cluster analysis is the definition of similarity between data objects. The task of clustering is to assign all instances to several clusters, so that instances of the same cluster are clustered around the center of a cluster, and the distance between them is relatively close, while the distance between instances of various clusters is relatively far [9].

Different clustering problems have various measures of similarity, such as distance, density, connectivity, and so on. Therefore, a variety of clustering algorithms have been derived, such as partition-based clustering algorithms and density-based clustering algorithms. The uses of clustering are very broad. For example, in business, clustering can help market analysts to distinguish various consumer groups from consumer databanks, and summarize the consumption patterns or habits of each type of consumers. Clustering analysis has been achieved in the field of e-commerce. Widely used [8]. The division and clustering algorithm requires the researcher to give the number of clusters of the data set or specify the cluster center point in advance through the analysis of the problem

requirements, and to continuously optimize the clustering target through the algorithm iterative operation so that the objective function finally converges, that is, the clustering is obtained. Class results [7].

As a module in DM, clustering can be used as a separate tool to discover some deep message distributed in the databank, and summarize the characteristics of each class, or focus on a specific class for Further analysis and cluster analysis can also be used as a preprocessing step for other analysis algorithms in DM algorithms [10]. Partition clustering is generally complete clustering, that is, each data point exists in a specific cluster, and is sensitive to the existence of outliers. This clustering method can be scaled according to the size of the data set, but the processing of large data will consume a high time cost.

2.2 Neural Network

The existence of some regularity between two or more variables in a dataset is called an association. Association analysis is the search for hidden networks of associations between variables. The problem described by the association rules is: in a given transaction set (existing in the form of a transaction databank), each transaction is an item set, and the close relationship between the data item sets in the transaction set can be found using mining algorithms. Represented by a series of association rules. The purpose of association rule mining is to discover meaningful connections between various items hidden in a large number of transaction data sets, help users find certain behavioral characteristics, and provide guidance for decision-making. According to conventional thinking, some variables in the data are sometimes difficult to be linked [11]. Such as: diapers and beer, without the analysis of large amounts of transaction data, this valuable message is difficult to find. A typical application of association rule mining is shopping basket analysis. This application is based on a business transaction databank, which records every business transaction at a certain commodity point of sale, such as a certain supermarket.

Support is an vital measure. A rule with low support may only appear by chance and is mostly meaningless. Confidence is used to determine the reliability of reasoning with the rule. Credibility describes the strength of the implications, while support describes the frequency of patterns in the rule. The rules with high credibility and strong support are called “strong rules”, and the essence of the association rule discovery task is to discover these strong association rules in the databank. In the databank, the rules that satisfy the minimum support and minimum confidence are called strong association rules.

In addition to shopping basket analysis, association rule mining is widely used in business decisions such as market planning, stock market transaction analysis, classification design, and many other practical applications such as service data analysis and network failure analysis. In association analysis, there are the following problems: when the support is set low, even if it becomes a frequent itemset for the occasional transaction, it will not have the desired effect in the subsequent recommendation. In the online schooling platform, association rules can be used for student source analysis, learning resource recommendation, learning material organization optimization, learner learning evaluation, adaptive adjustment of courses based on learner behavior, and evaluation of

schooling websites. The minimum support in the transaction databank will be used as the threshold for determining frequent itemsets, and the minimum confidence defines the minimum probability that the rule is established, and rules with a low probability are excluded. The rules obtained without considering the hierarchical relationship between data in reality are called single-layer association rules, while multi-layer association rules take into account the multi-level nature of data.

3 Cultivation of University Students' Media Literacy Based on Student Behavior DM

3.1 Data Mining

A DM technique produced in the late 1980s, known as knowledge discovery in databanks, is the analysis of data and data processing algorithms combined with traditional modern new techniques. A databank, data warehouse, World Wide Web or other message repository is a databank, data warehouse, spreadsheet or other type of message repository or a group of databanks. Data mining is the process of mining tacit knowledge from disordered and disturbing data using computer methods. That is, by analyzing the data, mining the previously unknown relationship, and explaining it to obtain valuable message. Education-based DM is an emerging discipline that focuses on developing and exploring unique data from schooling settings and using these methods to better understand students, and settings that enable them to learn. The essential difference between DM and traditional data analysis such as query, report, and online application analysis is that DM is to mine message and discover knowledge without explicit assumptions. The message obtained by DM should have three characteristics: unknown, valid and usable. Previously unknown message means that the message was not expected in advance, that is, DM is to discover message or knowledge that cannot be found by intuition, or even counter-intuitive message or knowledge. May be more valuable.

We can utilize DM techniques based on schooling systems: predicting students who drop out, relationships between students, university entrance exam results and their achievements, predicting students' academic performance, discovering relevant subjects of interest to students in the undergraduate syllabus, Possibly knowledge discovery in academic achievement, classification of students' performance according to their various learning styles in computer programming design, prediction of similarities and differences between schools. Data cleaning and data integration techniques can be performed on this data. Based on the user's DM request, the databank or data warehouse server is responsible for obtaining the relevant data. The domain knowledge of the knowledge base that is used to guide searches or evaluate the interestingness of result patterns. Such knowledge may include conceptual layers and user beliefs. In a short period of time, DM technique has achieved a qualitative leap from its emergence to its rapid improvement. Through systematic research on DM algorithms and application theory, it has developed rapidly from an isolated system problem to a comprehensive discipline. In schooling, the research that has been carried out so far includes: through cluster analysis and other methods, the registration message of students when they log on the website, and the results of each subject are analyzed, so as to classify students, discover students' interests, and

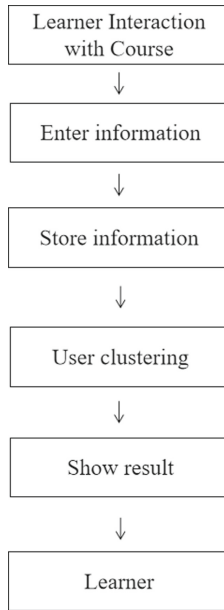


Fig. 1. Recommender system flow chart.

use the classification results to classify various students scientifically. Group, form an online learning exchange group, and learn more knowledge in exchange learning.

3.2 System Construction

With the rapid improvement of computer technique and Internet technique, the phenomenon of message “overload” in the network is becoming more and more serious. In the face of the “ocean” of message, the message used by users is only a drop in the ocean. After clarifying the purpose of DM, according to the problem to be solved, select variables that are closely related to it. The flow chart of the recommendation system is shown in Fig. 1.

When the user sends an HTTP request, the server will call these data sources according to the various requirements of various data sources. When the administrator is processing the statistical analysis of these data, the server will process the data in advance and delete unnecessary items in the project. Then, the processed data will be returned to the client, and then the components in the graphic will be used for the next graphic display on the client’s browser. In the model evaluation parameters, the accuracy rate, mean absolute error and root mean square error are extracted, and the comparison is made for the three classifiers of the experiment, as shown in Table 1.

As can be seen from the table, the prediction accuracy and other parameters of the SMO model are better than other models. Taking the SMO model as the object, 10 times ten-fold cross-validation was performed on the model, and the results are shown in Table 2.

Table 1. Performance comparison of classifier models.

Classifier	Classification accuracy	Mean absolute error	Root mean square error
J48	80.51%	0.165	0.42
Logistic	81.66%	0.20	0.37
SMO	83.75%	0.15	0.36

Table 2. Results data of ten cross-validation.

Accuracy%	Mean absolute difference	Standard deviation
84.16	0.1619	0.3946
83.97	0.1560	0.4089
83.89	0.1504	0.4019
84.82	0.1606	0.4081
84.80	0.1509	0.4001
83.06	0.1589	0.4046
83.79	0.1690	0.3945
83.15	0.1579	0.3951
83.19	0.1590	0.3958
00084.39	0.1549	0.4050

Through the outer cross-validation, the results obtained are more accurate than a single cross-validation.

The function of the schooling DM analysis system is to complete the data analysis related work in the schooling field. When administrators conduct in-depth DM, they will first process the data, so that the data used can be adapted to the DM algorithm used, and the results obtained after mining are finally converted into the corresponding formal format. Users need to spend a lot of time and energy in order to find the message that suits them. This problem also exists in online schooling. The learner does not know where the learning resources are placed. The interface of all learners is the same. In the process of finding the resources they need, the learners are easily “lost” and confused. Overwhelmed feeling. As shown in Figs. 2 and 3, we can find that the research on the learning experience and learning process of smarter classrooms is increasing year by year.

The statistical data set includes the estimated difficulty level of each question completed by the students during the learning process of this chapter, the knowledge points examined, the time spent answering each question (the system records the unit in seconds), and the students' answers. Right and wrong situation. The most common research proposition in the field of schooling DM is to extract key message from complex data to build a student model and predict the future professional performance of students. From

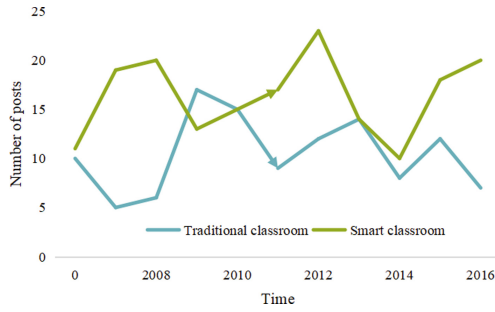


Fig. 2. Number of relevant documents in core journals.

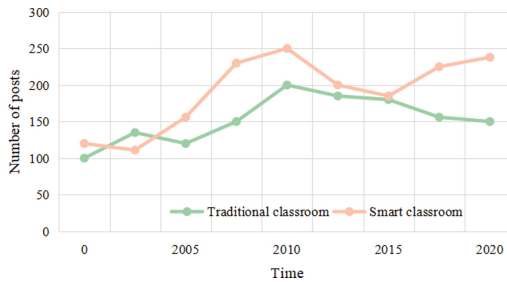


Fig. 3. Trend chart of domestic “learning experience” research.

the data point of view, the performance of students is directly related to their assessment scores, and in the ocean of data stored in the schooling message system, the scores of students in each course are the largest component.

There are three common distance similarity:

Manhattan:

$$D_{ij} = \sum_{k=1}^p |X_{ik} - X_{jk}| \tag{1}$$

Euclidean:

$$D_{ij} = \left(\sum_{k=1}^p (X_{ik} - X_{jk})^2 \right)^{1/2} \tag{2}$$

Minkowsky:

$$D_i(q) = \left(\sum_{k=1}^p |X_{ik} - X_{jk}|^q \right)^{1/q} \quad q > 0 \tag{3}$$

The smaller the distance between the two objects, the more similar the two objects are; the larger the distance between the two objects, the less similar the two objects are.

When designing this module, first construct an initial Job result entity, initialize the control message in the entity, and try to execute the Job content, including Job

initialization, data extraction process, data transfer process, data loading process, and Job destruction process, and Encapsulate the execution log of each step into the Job result entity. In order to solve the problem of users looking for resources in e-commerce, personalized recommendation services are widely used. Similarly, in online schooling, it is also necessary to use personalized recommendation technique to better promote learners' autonomous learning.

4 Conclusions

Educational DM has played a very vital role in the process of school schooling, and the huge and complex schoolingal data has brought difficulties to our work. How to make better use of schoolingal data for DM to make better decisions for universitys and universities? Good support analysis is another big challenge we face. With the continuous application of cloud computing technique in the field of schooling, it will bring great changes to schooling. As the main component of the platform, resources are obtained by analyzing a large amount of data in the platform to obtain helpful resources for various students' learning, and constantly improve the functions of the platform. With the help of widely used DM technique, it can provide effective assistance to teaching and learning by analyzing the situation of students completing homework online, thereby making up for the deficiencies in traditional schooling. To further establish a standardized databank, you can contact relevant leaders to integrate all aspects of the school's schoolingal data, which is convenient for future improvement. In the future schoolingal DM process, a unified data warehouse is imperative.

References

1. Abhishek K, Kumar V R. Parametric appraisal and optimization in machining of CFRP composites by using TLBO [J]. *Journal of Intelligent Manufacturing*, vol. 28, no. 8, 2017.
2. Boaroli L. Data Monitoring and Hardware Control for App Android by Bluetooth Communication for Laboratory Teaching in Electrical Engineering Courses[J]. *IEEE Latin America Transactions*, vol. 15, no. 1, pp. 31-39, 2017.
3. Coope C M. Does teaching by cases mislead us about morality?[J]. *Journal of Medical Ethics*, vol. 22, no. 1, pp. 46-52, 2019.
4. Fragneto R Y, Gaiser R. Labor analgesia: We need to better understand and educate our obstetric patients[J]. *Journal of Clinical Anesthesia*, p. 41:42, 2017.
5. Joachim, Walther, Shari, et al. A Model of Empathy in Engineering as a Core Skill, Practice Orientation, and Professional Way of Being[J]. *Journal of Engineering Education*, vol. 106, no. 1, pp. 123-148, 2017.
6. Lpc M, Cic D, Cic K, et al. "You Use It, You Clean It": A Multifaceted Approach to Educate and Engage Pediatric Healthcare Workers in Non-Critical Reprocessing[J]. *American Journal of Infection Control*, vol. 47(6):S22-S23, 2019.
7. Maha, Kamal, Eltayeb, et al. Distribution of white spot lesions among orthodontic patients attending teaching institutes in Khartoum.[J]. *BMC oral health*, vol. 17, no. 1, p. 88, 2017.
8. Malliarakis C, Satratzemi M, Xinogalos S. CMX: The Effects of an Educational MMORPG on Learning and Teaching Computer Programming[J]. *IEEE Transactions on Learning Technologies*, vol. 10, no. 2, pp. 1-1, 2017.

9. Mokgatle M, Madiba S. High Acceptability of HIV Self-Testing among Technical Vocational Education and Training College Students in Gauteng and North West Province: What Are the Implications for the Scale Up in South Africa?[J]. *Plos One*, vol. 12, no. 1, pp. e0169765, 2017.
10. Selleck C, Jablonski R, Miltner R S, et al. Partnering to Educate Nurses in Long-Term Care[J]. *The Journal of Continuing Education in Nursing*, vol. 51, no. 2, pp. 75-81, 2020.
11. Shiha G, R Soliman, Mikhail N, et al. An educate, test and treat model towards elimination of hepatitis C infection in Egypt: Feasibility and effectiveness in 73 villages - *ScienceDirect*[J]. *Journal of Hepatology*, vol. 72, no. 4, pp. 658–669, 2020.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

