Application Research of Big Data Mining in Personalized Teaching of Internet Education Platform

Chengcai Tan1*, Li Xue1
1Marxist College of Changchun Guanghua University, China, 130033

*Corresponding author e-mail: tanchengcai@ghu.edu.cn

Abstract. Big data is a valuable resource for the Internet education platform. Big data mining is an important technology for Internet education platforms to provide personalized services for learners [1-3]. Using big data related technology to explore the inherent rules among students, teachers, courses and grades can provide reference for decision-makers of education and teaching, and can also provide guidance for the school's teaching tasks and teaching plans [4, 5]. This paper briefly introduces the concept of data mining and personalized teaching, and studies the application of data mining technology in the personalized teaching of Internet education platform, in order to improve the school's teaching management level and students' academic performance.

Keywords: Big Data Mining, Internet Education, Personalized Teaching

1. Introduction
Comrade Xi Jinping has pointed out that education is an important way to improve the comprehensive quality of the people and promote their all-round development. Internet education is an effective means to promote starting point equity. It allows children in remote areas to "enter" famous schools, greatly broadening their horizons. Internet education has not only attracted the attention of experts and scholars in the field of education, but also attracted widespread interest from Internet companies. In recent years, personalized teaching services for users have become one of the important themes of Internet education. To provide personalized teaching services for learners requires effective integration of massive data of different types of Internet education platforms, and data mining and real-time analysis are carried out on this basis to discover the personalized characteristics, interests and needs of learners, so as to provide targeted teaching services.

2. The concept of data mining and personalized teaching
2.1. The meaning of data mining
Data mining refers to extracting information or knowledge that is useful and potentially valuable to a specific group of people from a large number of random and incomplete data. Figure 1 is a specific process of data mining. Data mining technology is divided into classification and prediction, cluster analysis and association analysis according to functions. Classification and prediction are two basic
forms of data analysis, mainly used for the extraction description of important data sets and the prediction of future data development trend models. Clustering is the classification of data based on its properties. Then, neat data after clustering can be obtained, some of which can directly reflect the intrinsic relationship of the object, and other results need to be further processed by other tools. The association rule is to use the certain rules contained in the data set to mine the data. It can be said that it is one of the most important models widely used in the field of data mining.

![Figure 1. General flow chart for data mining.](image)

2.2. The meaning of personalized teaching
Personalized teaching is a teaching model that expresses individuality and is student-centered. In the process of teaching, we should develop a personalized teaching strategy to improve learners' initiative in learning. In general, personalized teaching should meet the requirements of student-centered and learner personality. The essential characteristics of personalized teaching: First, personalized teaching emphasizes the diversity of students. Second, personalized teaching emphasizes the subjectivity of learners. The new student-centered teaching model should be based on the individual situation of each student and fully explore their potential. Diversified content, diverse forms and appropriate teaching methods enable each learner's potential to be fully explored and developed. When implementing personalized teaching, each student should be regarded as a subject with independent personality, and should respect the learners' hobbies, personal habits and learning goals. Only by formulating a special learning plan for the students according to the basic situation of the learners can the learners master the learning skills.

3. Personalized teaching service based on big data mining

3.1. Cluster analysis and its application
Cluster analysis is to cluster similar objects into clusters according to their attributes and characteristics, so as to grasp the internal laws of things. According to the background, identity, age and other personal information, browse, click, view, download the teaching resources, and cluster the learners through the learning behavior data. Analyze the characteristics and hobbies of different learning groups, and adopt targeted classification teaching to meet the individual needs of learners. For example, learners are clustered according to the scores of several courses, which are divided into four levels: excellent, good, medium, and unqualified. According to the different levels of learners, make
follow-up learning plans of different difficulty for them. Clustering teaching resources according to attributes such as type, language, title, keyword, and click rate can recommend relevant learning resources for learners. For example, "Shi Ji", "Han Shu", "Later Han Shu" and "Records of Three Kingdoms" enjoy a high reputation in the works of Chinese history, and they are called "Early Four Historiographies" and belong to the same category. For the "Shi Ji" learners to recommend the "Han Shu", "Later Han Shu" and "Records of Three Kingdoms" learning resources, can not only satisfy the interests of learners, but also help learners in the history of knowledge to a higher level. Figure 2 is a clustering analysis of the data set ex.Food_cal using the clustering method. The CLUSTER process implements hierarchical classification. The code is as follows:

```
PROC CLUSTER DATA=EX.NUTRITION OUTTREE=TREE METHOD=AVE CCC PSEUDO;
   VAR MAGNESIUM_MG PERCENT_WATER PROTEIN_G SATURATE_FAT_G;
   ID FOOD;
RUN;
```

![VIEWTABLE: Ch13.Nutrition](image)

**Figure 2.** Application of cluster analysis in food nutrition course.

### 3.2. Classification and prediction and their applications

Classification techniques can be used for learner performance predictions, learning needs predictions, and learning risk warnings. For example, the academic performance is divided into 4 grades of excellent, good, middle and poor. According to learners' background, past transcript scores, activity logs, etc., the academic performance and grade of new courses are predicted. The learning courses are classified by content, and the new learning needs are predicted according to learners' background, previous learning contents, activity logs and learning paths, so as to push the required learning courses. The learning risk is predicted based on the learner's academic performance, participation level, and classroom performance. If the learner falls into the high-risk category, the learner should be intervened in time to provide personalized guidance. In recent years, the scene of resource recommendation services has also attracted attention. Through scene modeling, scene classification and learner scene recognition, teaching resources suitable for learners can be provided. For example, if the device scene of the learner is found to be a mobile terminal, a micro-course video with a shorter duration is recommended to facilitate the fragmentation learning. If it is found that the learner's device scene is a desktop computer, a MOOC video with longer duration and greater difficulty, is recommended to them so that they can concentrate on learning for a long time.
3.3. Association analysis and its application

Association analysis is to find interesting association relationships among data items and express them as association rules. Through the correlation analysis of students' learning data, students' learning mode is understood, and then the best learning plan is selected. Relevance is one of the characteristics of digital education resource services in the context of big data. Teaching resources are interrelated in content and topic. Learners and teachers can set up real-time or non-real-time discussion groups according to their interests and learning needs to conduct question-and-answer communication and discussion, so as to generate rich correlation. Through the correlation analysis of teaching resources and learners, the correlation among teaching resources, learners and teaching resources and learners can be found, and then relevant learning topic setting and relevant learning resources pushing can be carried out for learners. Through the association rule mining technology, the student's grades can also be extracted and analyzed, and then the links between the courses can be found. This extracts valuable and relevant information to help students, teachers, and system decision makers.

4. Conclusion

In the era of "Internet +", the production and lifestyle have undergone tremendous changes, and education has also undergone changes. Big data, which is mainly composed of learners, teachers and teaching resources, is a valuable resource for the Internet education platform. Use big data mining technology to deeply analyze the big data gathered by the platform. Data mining, collation and analysis of massive data can provide effective decision-making reference. Internet education can provide individualized teaching services for learners. The Internet education platform makes full use of learning resources, and the role of teachers can be fully demonstrated.

Acknowledgement

Ideological and political special project of jilin education department in 2018: theoretical and practical research on the cultivation of education system in applied colleges and universities (Project No. JJKH20181382SZ).

References

[1] You H, Qian F, Huang X, et al. Research and Practice of Building a Personalized Recommendation System for Mobile Game Platform Based on Big Data Mining[J]. Telecommunications Science, 2014.

[2] Zeng Z, Rao H K, Liu A P . Research on personalized referral service and big data mining for e-commerce with machine learning[C]// 2018 4th International Conference on Computer and Technology Applications (ICCTA). IEEE, 2018.

[3] Wang G Y. The Application and Research of Personalized Network Teaching Resources Database Based on Data Mining[J]. Advanced Materials Research, 2013, 756-759:2475-2478.

[4] Lin C F, Yeh Y C, Yu H H, et al. Data mining for providing a personalized learning path in creativity: An application of decision trees[J]. Computers & Education, 2013, 68(4):199-210.

[5] Reddy P K, Reddy B B, Rao D R. A Model of Virtual Crop Labs as a Cloud Computing Application for Enhancing Practical Agricultural Education[C]// International Conference on Big Data Analytics. 2012.