Analysis of Students' Learning Behavior Based on Association Rule Mining Algorithm in Moodle Network Platform

Shaozhen Huang

Hainan Open University, Haikou, Hainan, 570100

Keywords: Moodle Internet Platform; learning behavior; Association Rule

Abstract: Based on the learning behaviors of learners in the curriculum management system, the assessment of learning performance is carried out to provide curriculum improvement and learning suggestions. The evaluation of learning materials and online courses provides feedback for teachers and students of E-learning courses. Quality is very important. However, because many curriculum management systems do not provide specific tools that allow teachers to track and evaluate all learners' behavioral activities throughout the entire process, it is very difficult to select valuable information when faced with large amounts of system data. Educational data mining is an effective way to solve this problem. Based on the introduction of E-learning data mining process, this paper focuses on the application of data mining technologies such as statistics, visualization, classification, clustering, and association rule mining in Moodle systems.

1. Introduction

Data mining is an interdisciplinary field that incorporates multiple computing paradigms such as decision tree construction, criteria induction, artificial neural networks, case learning, Bayesian learning, logical design, statistical algorithms, and more. Some very useful data mining techniques and methods in CMS are statistics, visualization, clustering, classification, and association rule mining. By using these methods, students can use the system's data to reveal new and interesting useful knowledge. This article will use the free course management system Moodle as an example to explore the application of data mining technology in distance learning.

2. Data Mining Process in E-learning

Traditional E-learning course development is a difficult task. Developers (usually course teachers or online teachers) must choose the content to be displayed, determine the structure of the content, and establish the most appropriate content for each class of potential users of the curriculum. Due to the complexity of these decisions, directional design is almost impossible. In many cases, it may be necessary to evaluate and revise course content, structure, and navigation based on student use information. What is more, it is necessary to adopt a continuous experience evaluation method. To facilitate this activity, you can use data analysis methods and tools to observe student behavior, help teachers identify possible errors and results, and take appropriate improvement measures. Traditional data analysis in Elearning is based on the assumption that the user starts with the problem and then seeks out the data to confirm his intuition. The premise of using this method is that the impact factors and data involved are moderate. But when users want to find more complex schemas from different relationships of data, it is very difficult to adopt this method. In contrast, using data mining techniques, you can automatically reveal the hidden information in the data, because its assumptions come from data rather than artificial settings. [1] Data mining constructs an analysis model that can reveal patterns and trends that students should pay attention to in their use of information.

3. Preprocessing of Moodle data

Moodle is an open source course management system that helps teachers create an effective

online learning community. Moodle is yet another alternative to patented commercial e-learning solutions that can be shared for free under open source certification. Moodle has been widely used in many universities and colleges around the world. Any organization has all rights to use the source code and can modify the source code if necessary. Moodle's modular design makes it easy to create new lessons and add content that learners need to learn. Its design is designed to support the social constructivist learning model. Social constructivism believes that when students interact with learning materials and create new materials for other learners, they learn best when they interact with other learners. Moodle doesn't need to use this learning mode in the course, but it provides good support for this mode. Moodle has a flexible array of teaching activity modules and resource modules that can create five types of static course materials: text, web pages, page links, course catalog browsing, and a table displaying any text or picture; six types of interactive course materials: extracurricular activities, Words, logs, homework, quizzes, surveys, reviews; interactions among five students: chat, forums, glossary, WIKI, and workshops. [3] Moodle details all student activity logs. It records every click the student makes for navigation and has an ordinary log browsing system embedded. Log files can be categorized by course, participant, date, and activity. Teachers can use these logs to determine who is actively involved in a course activity and when. For exam activities, teachers can not only obtain test scores and test time, but also analyze each student's answer in detail and analyze the question itself. Teachers can easily get a full report of learners' individual behavioral activities and all participants in a particular activity. Each student's learning activity report is available, and describes the details of the relevant module (last visit time, number of visits) and the details of each student's participation. The log can also show the behavior of classroom activities on different dates or times. This helps teachers check if each student has completed a specific task or maintains enough online time for a particular activity. Moodle does not store logs as text files. In fact, it stores logs in a relational database (MySQL and PostgreSQL are the best support, and Oracle, Access, Interbase, and other databases can also be used). Among all services provided by CMS, collecting detailed access information and high-end usage information, the database is more powerful and more flexible than traditional text log files. The Moodle database has about 145 associated tables, but not all information is required. Preprocessing Moodle data is the conversion of system data into a format suitable for use by a data mining algorithm. In particular, the task of data preprocessing is usually done by managers rather than teachers. Managers usually need to manually complete data preprocessing such as data cleaning and purification, user confirmation, time confirmation, path improvement, transaction confirmation, data conversion and concentration, data integration, and data simplification. Moodle, like many other course management systems, can use user authentication to authenticate identity of logged-in users, thus avoiding the identification of representative users and groups. Therefore, the task of purifying and preprocessing the data of the Moodle system is less than other network systems.

4. Application of Data Mining Technology in Moodle System

Data mining is an efficient process of excavating non-obvious value patterns from large data sets. [4] There are many common and special data mining tools. Some commercial mining tools such as DBMiner, SpssClementine, and DB2 Intelligent Miner. Some popular mining tools are Weka and Keel. There are also some specific educational data mining tools, such as Mining for joint and pattern mining; MultiStar for joints and classification; Tools based on student's academic performance for performing qualitative analysis; [5] EPRules for federation; KAON for clustering and text mining; Synergo for statistics and visualization; GISMO for visualization; Listen for visualization and browsing; TADAEd for observation and mining; O3R for sequential pattern mining MINEL for Learning Approach Mining; CIEoF for Joint Rule Mining; Simulog for Unexpected Behavior Patterns; and Sequential Mining for Pattern Mining. Since the Weka and Keel systems are free and can be executed in Java, using the characteristics of ARFF array external representation format files, this article will use them for data mining. Both of these softwares are easily accessible from the Internet, and you don't have to worry about data format problems when using them, and you can use the same programming language for correction if you need to. Weka is

an open source software that provides data mining algorithms such as machine learning and preprocessing, classification, regression, clustering, joint control, and visualization. Keel is an open source software tool for building and using different data mining models such as data preprocessing algorithms, decision trees, rule extraction, descriptive induction, statistical methods, neural networks, and multiple classifier evolution systems. Data mining has some representative forms, such as probability, rules, and trees, as well as tasks and methods derived from machine learning, statistics, visualization, and artificial intelligence. Below, we will introduce specific application examples of data mining technology in the Moodle system. Statistics on student use methods are usually the starting point for E-learning systematic reviews. Usage statistics are available using standard tools for analyzing web server logs, such as AccessWatch, Analog, Gwstat, and WebStat. However, there are some specific statistical tools for educational data, such as Synergo, ColAT. A simple example of usage statistics in the E-learning system is measurement, such as number of visits and visits per page. Some statistics show the distribution of learners over the past period of time and the courses most frequently visited; other statistics show the number of visits and durations per quarter, the most frequently searched and searched words; there are also reports showing support systems and weekly and monthly reports on user behavior trends. There are also specific statistical tables showing the average number of constraint violations and the average complexity of the problem. More advanced statistical methods, such as correlation analysis between variables, can be used to infer attitudes that affect student learning or to predict student achievement. Regression analysis can be used to predict knowledge and morals that cause unsatisfactory test scores. It can also be used to predict whether students can answer questions accurately and predict grades at the end of the school year.

Moodle only displays statistics such as ratings, quiz modules, and more. Teachers can use the scale to assess forums, assignments, quizzes, courses, journals, and seminars to evaluate students' academic performance. Moodle comes preloaded with two scales, one is a numeric scale (from 1 to 100), and the other is a scale used to indicate the relationships between certain items in the course. In order to take effective measures to observe students' progress, teachers can customize rating scales (classification levels, assignment scope levels, weight levels, hide or show grades for students). In addition, Moodle has statistical test reports based on project analysis. The report in some way shows process test data that is suitable for analyzing and judging the performance of each project issue itself. Statistical parameters are calculated based on the interpretation of classical test theory: Capability Index (F.I.), Standard Deviation (S.D.), Discrimination Index (D.I.), and Discrimination Factor (D.C.). Based on this, teachers can identify which problems are most difficult (F.I. value is small) and which are the simplest (F.I. value is large); which problem has the best discrimination (large values of D.I. and D.C.). This information can be downloaded in text format or exported in Excel format for drawing and analysis. Using this information, teachers can maintain the test continuously. For example, teachers can delete and modify issues that are too difficult or too simple. When some test questions are too difficult, or they have their own syntax and semantic errors, teachers need to correct them. Or when some test questions are too easy, almost all students respond perfectly, and test questions do not have a high ability to distinguish between good and bad students.

5. Conclusion

This article mainly discusses the application of several data mining techniques such as statistics, visualization, clustering, classification, and association rules mining in the Moodle curriculum management system. Research shows that applying data mining tools to the Moodle curriculum management system can provide distance education teachers with important information such as how to promote student learning, improve online courses, and conduct course management. Through this information, teachers can further improve teaching methods and improve teaching strategies. Improve the effectiveness and quality of distance learning. This study only describes the application of various data mining technologies in the Moodle curriculum management system. In fact, these technologies can also be combined to obtain important information that is worth paying

attention to in the CMS system.

Acknowledgements

References

[1] Chang, F. C. I., Hung, L. P., & Shih, T. K. A new Courseware for Quantitative Measurement of Distance Learning Courses [J]. Journal of Information Science and Engineering, 2003, (19): 989~1014.

[2] Merceron, A., & Yacef, K. TADA-Ed for Educational Data Mining [J]. Interactive Multimedia Electronic Journal of Computer Enhanced Learning, 2005, 7(1): 265~343.

[3] Feng, M. & Heffernan, N. Informing Teachers live About Student Learning: Reporting in the Assistment System [J]. Technology, Instruction, Cognition, and Learning Journal, 2006, (3): 1~8.

[4] Tang, T. & McCalla, G. Smart Recommendation for an Evolving E-learning System [J]. International Journal on E-Learning, 2005, 4(1): 105~129.

[5] Chen, G. Liu, C., Ou, K., & Liu, B. Discovering Decision Knowledge from web log Portfolio for Managing Classroom Processes by Applying Decision Tree and Data Cube Technology[J]. Journal of Educational Computing Research, 2000, 23(3): 305~332.