Intelligent Information Teaching Management System Based on Data Mining Technology

Lan Wu and Ji Cheng*

{*Corresponding author: 15377500195@163.com}
{wulanforever2023@163.com}

Wuhan Business and Trade Vocational College, Wuhan, Hubei, 430205, China

Abstract: With the continuous development of internet technology in today's era, the internet continues to empower various fields,where more and more forms of online education have emerged. Among them, online learning systems are widely used in online teaching in many universities both domestically and internationally, allowing students to learn anytime and anywhere in the system. However, in experimental teaching in universities, students often encounter more nuanced problems, and they are unable to quickly and accurately locate the content they urgently need from a large amount of resources. This article designs and develops a management system for experimental teaching in vocational colleges by mining educational data information from universities. In order to verify the correctness of the system, functional, performance, and data synchronization tests were completed. After testing, the system can meet the teaching requirements of universities very well.

Keyword: Internet Technology, Data Mining, Intelligent Informatization, Teaching Systems

1. Introduction

With the continuous development of Internet information technology, innovative activities with the theme of "Internet plus+education" have emerged in endlessly. Experimental teaching itself is an important part of the teaching process, which greatly affects students' mastery of classroom theoretical knowledge and also affects their hands-on practical abilities. Vocational colleges mainly cultivate technical and skilled talents, with a greater emphasis on cultivating students' practical skills and practical work abilities. In recent years, under the background of vigorously developing vocational education, the state and vocational colleges have paid more attention to laboratory construction, invested a lot of funds in laboratory construction, and the number of Laboratory equipment has increased in a "blowout" manner, with an increase in the proportion of practical courses. Exacerbating the contradiction between the rapid growth of experimental equipment and the relative lag in laboratory informatization construction in vocational colleges.

ICSETPSD 2023, November 17-18, Coimbatore, India

DOI 10.4108/eai.17-11-2023.2342772

The online experimental teaching system is highly favored by users due to its convenience in time, location, and shared educational resources, which reduce students' learning barriers and costs. Traditional online experimental teaching systems provide users with massive resources, but there are also some problems and pain points [1]. Firstly, during the experimental teaching process, students may encounter many detailed problems, and students may not be able to accurately locate the answers they want in a short period of time, often over an hour of course. If long videos are divided into short videos, the number of resources will sharply increase, and relying on traditional database search response time will be very slow. Intelligent search technology can accurately find high matching resources in massive resources with extremely short response time [3]. Introducing intelligent search technology can address this issue in a targeted manner, allowing users to easily find resources of interest without much effort. This not only saves users time and costs, but also greatly improves the user experience. Thus, it further stimulates users' learning enthusiasm and enhances their subjective initiative in learning. Secondly, experimental teaching often receives a large number of visits at the same time. With the continuous innovation and development of smart experimental teaching and the iterative upgrading of the system, the number of users will rapidly increase, and the system will have the need to add functional modules or refactor. Therefore, ensuring the robustness, high availability, low coupling, and scalability of the system has become an urgent problem. At present, educational institutions in various countries have also invested in large-scale online open courses, and many foreign universities have also joined some large-scale MOOC open course projects, such as Coursera, edX, Udacity, etc. [4]. In recent years, the experimental teaching community has been exploring new models and methods of experimental teaching. Taking the specialized experimental center of electronic information and Telecommunications engineering as an example, the experimental center has invested construction of the new model of experimental teaching, and has continuously made innovation and development in hardware facilities, but there is still a large enabling space in software infrastructure [5]. The traditional experimental teaching mode in universities is limited by educational resources such as the number of teachers, software and hardware resources, and experimental teaching venues. The MOOC ization of experimental course content is a new way to open and share experimental teaching resources.

This paper designs and implements an intelligent information-based teaching management system based on data mining technology. It provides a detailed analysis and design implementation of key technical points, analyzes the functional and performance requirements of the system, designs various functional modules and a system architecture based on data mining, and conducts functional and performance tests on the system, achieving the expected requirements.

2. Data Mining and Intelligent Informationization Teaching

2.1 Decision Tree Technology

Decision tree technology is a classification algorithm in machine learning. It takes information entropy as the main reference, has the characteristics of high readability and rapid analysis, and is a kind of Supervised learning. The main process of building a decision tree is as follows [6]:

Step 1: Treat all data as a node and proceed to the next step;

Step 2: Select a new data feature from the existing data features and partition the nodes to proceed to the next step;

Step 3: After generating some child nodes, make a decision on each child node. If the requirement to stop splitting is met, proceed to the next step; Otherwise, you will proceed to step 2;

Step 4: Type with the highest proportion in the total number of nodes.

There are two types of data segmentation objects: discrete and continuous. When dividing discrete data, it is divided according to feature values, with each feature value corresponding to a node; When dividing continuous attributes, the first step is to determine the numerical value of the continuous attribute

Sort and then split into corresponding intervals, such as [0, 10], [10, 20], [20, 30]. Each attribute corresponds to an interval, and each interval corresponds to a node.

When using algorithms to establish a decision tree, attributes are divided. Assuming that the dataset is represented by S and A represents a specific attribute [7].

Step 1: Calculate category information entropy

The time uncertainties of different types of occurrences in each study sample.

$$Ent(S,A) = -\sum_{i=1}^{c} \frac{S_i}{i} \log_2 \frac{S_i}{i} \qquad (1)$$

Among them, S_i to S_c are the set of c sample examples generated by splitting the attribute A of c values into S.

Step 2: Analyze the dataset S through statistical analysis of various attributes to obtain information gain.

$$Gain(S,A) = Ent(s) - \sum_{v=1}^{v} \frac{Sv}{v} Ent(S_V) \qquad (2)$$

Among them, Ent (S) represents the calculation based on the entire attribute as a set, Ent (SV) represents the calculation based on the attribute value of attribute A as a set, and v represents the number of attribute values of attribute A.

Step 3: Calculate the information gain rate

$$GainRatio(S,A) = \frac{Gain(S,A)}{Ent(A)} \quad (3)$$

The higher the information gain rate, the higher the purity of the indicator. The attribute with the highest information gain rate is used as the root node, and steps 1, 2, and 3 are repeated to continue building the decision tree [8].

2.2 Indicator Weighting Method

The indicator used to solve the relationship between the various indicators multiple indicators. The weight of indicators reflects the importance of each indicator in the evaluation system and is the main factor affecting the accuracy of the evaluation system[9].

The subjective weighting method is currently widely used as a weighting method, which mostly assigns weights to each indicator based on the work experience of experts or researchers in a certain field, after subjectively evaluating the importance of each indicator [10]. Usually, multiple experts or scholars are organized to assign their own indicators and finally integrate them. The advantage of subjective empowerment method is that it is easy to operate and can greatly mobilize the personal experience and knowledge of managers. However, it also faces some problems, as relying on the personal experience of managers can produce corresponding subjective results, which are easily influenced by the personal psychological factors of managers, thereby affecting the correctness of the results [11]. At present, the common subjective weighting methods are: Delphi method method, analytic hierarchy process, binomial coefficient method and minimum scoring method [12].

Contrary to the subjective weighting method, the objective weighting method determines the weight based on the degree of variation between the original data. It analyzes and processes the evaluation data of each indicator to calculate the weight, which avoids subjectivity to a certain extent. However, the objective weighting method also has certain limitations. The method is too dependent on data to think about problems from the perspective of decision makers, which may lead to results that are quite different from the ideas of decision makers, And the calculation process is quite complex [13]. At present, common objective weighting methods include principal component analysis, entropy method, dispersion method, and CRITIC method [14].

Both subjective and objective weighting have own advantages and disadvantages. In order to balance decision-makers' preferences for attributes while also striving to reduce the subjective arbitrariness of weighting, and achieve the unity of subjective and objective weighting for attributes, thus making weighting more scientific and reliable, scholars have proposed a combined weighting method that combines subjective and objective weighting methods. The combination methods include multiplication synthesis method, entropy weight method, coefficient of variation method, correlation coefficient method, etc. [15].

3. Platform Architecture Design

3.1 Overall System Architecture

This system adopts a B/S structure, which is a working mode of browser request and server response. It is divided into three levels: client browser, web server, and database server. The advantage of hierarchical design is that each level is independent of each other and does not interfere with each other. Regardless of which level of optimization, maintenance, or upgrade operations are more convenient; At the same time, the client is isolated from the database, and users cannot directly access and operate the database. Through the effectiveness verification and user permission control of the middle layer web server, the security of the system is greatly improved.

3.2 System Module Design

The experimental teaching management system in vocational colleges mainly consists of three types of users: teachers, students, and system administrators, with a large amount of information sharing needs and different permissions. So a more reasonable solution is to divide the class, students, teachers, courses, experiments, experimental reports, laboratory announcements, system administrators, and other information within the system into independent modules for management. Each module has different permissions open for use by system administrators, teachers, and students. This design approach not only improves system security but also avoids the possibility of unauthorized operations.

(1) Class Information Module

Module description: The system administrator adds a class during system initialization, enters the class number, class name, and class information in sequence, and can query based on these three keywords.

Module permissions: System administrators can add, delete, modify, and check; Teachers can only view and query; Students do not have permission to operate. Module constraint: Before deleting class information, all student information under that class must be empty.

(2) Student Information Module

Module description: The system administrator adds students when the system initializes, enters student number, name, class, login password, and student information in sequence, and can query keywords according to student number, name, class, and student information.

Module permissions: System administrators can add, delete, modify, and check; Teachers can view and query; Not available for students. Module constraint: Each student can only belong to a unique class, and the class number information in the student ID must be consistent with the class number information of the class they belong to.

(3) Teacher Information Module

Module Description: The system administrator adds a teacher during system initialization, sequentially entering the teacher ID, name, login password, whether the teacher information is enabled, and can query keywords.

Module permissions: System administrators can add, delete, modify, and check; Teachers can view the list of in-service (enabled) teachers; Students can view a list of in-service (enabled) teachers.

Module constraint: None.

(4) Course Information Module

Module Description: The system administrator adds courses during system initialization, sequentially entering the course number, course name, and grade, and can query based on the grade and course keywords.

Module permissions: System administrators can add, delete, modify, and check; Teachers can only view and query; Students can only view and query.

Module constraint: Before deleting a course, all experiments under the course must be deleted first.

(5) Experimental module

Module Description: In the Add Experiment interface, the teacher can enter the experiment number, experiment name, teacher, release time, experiment status, course, experiment cover image, experiment content, and search for the experiment name. The experiment list is arranged in the form of a first level classification for the course, and a second level classification for the experiments under the course.

Module permissions: System administrators can add, delete, modify, and check; Teachers can add, delete, modify, and check the experimental projects they are responsible for; Students can view it.

Module constraint: Before deleting an experiment, the experimental report belonging to that experiment must be deleted first.

(6) Experimental Report Module

Module description: Students participate in an experiment on a class basis. In the filling interface, the system automatically generates an experiment report number, experiment name, experiment status, course, student ID, and name. Students only need to fill in the experiment content results through a text editor, select whether the report is completed, and complete the submission of the experiment report; Teachers review reports on a class by class basis.

Module permissions: System administrators can view and delete, but cannot add or modify; Teachers can query and view all experimental reports and review them; Students can only view all experimental reports and grades they have participated in.

Module constraint: Only when the experiment is in an open state can experimental reports be submitted;

(7) Laboratory Announcement Module

Module description: The system administrator can add laboratory announcements as needed, enter the announcement title, publisher, publication time, viewing level, and announcement content in sequence, and can query keywords based on the announcement title.

Module permissions: System administrators can add, delete, modify, and check; Teachers can view; Students can view it. Module constraint: laboratory announcements are visible to system administrators according to the viewing Elo rating system; Visible to system administrators and teachers; All system administrators, teachers, and students are visible.

(8) Administrator module

Module Description: The system administrator can add other administrators as needed, enter the administrator account, Chinese name, password, and confirm password in sequence.

Module permissions: System administrators can add, delete, modify, and check; Teachers are not available; Not available for students. Module constraint: A unique system administrator account cannot be deleted.

4. Platform Testing Analysis

4.1. System Performance Testing

The system response time mainly tests the web page opening speed and login time when the system uses different browsers. This system is tested on the client using three commonly used browsers, IE, Chrome, and Firefox. The corresponding time for logging in or opening web pages in different browsers is not significantly different, all of which are around 500ms.

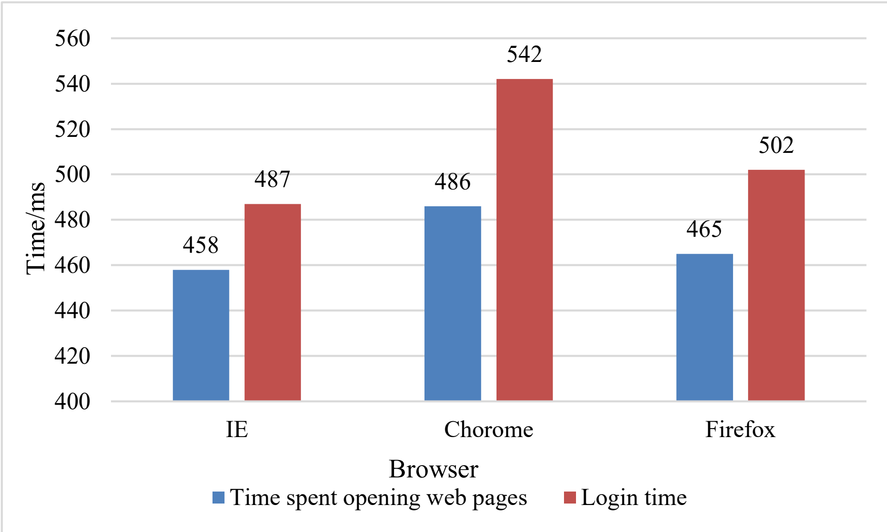

Fig.1 System Performance Testing

4.2. Stress Testing

Set the number of threads to 200, 250, 300, 350, 400, and 450, all accessing the interface within 1 second. The test result data is shown in Table 1.

The performance requirement of the experimental center system is 400QPS, and the average response time of the interface should be below 500ms. The request error rate during low peak periods is required to be around 1%, and the request error rate during peak periods is required to be below 10%. Satisfy system performance requirements with QPS not exceeding 400.

Table 1. Stress Test Results

Configuration		Test result	
Number of concurrent threads	Pressor phase	Response time	Error rate
200	1	23	0%
250	1	29	0%
300	1	134	0%
350	1	361	0%

400	1	345	0%
450	1	795	13%

4.3. Synchronous Performance Testing

Select files with sizes of 100M, 500M, 1G, and 10G for three transmission tests. The test results are shown in Figure 2, with an average time of 0.92, 4.35, 8.89, and 87.49 seconds, respectively.

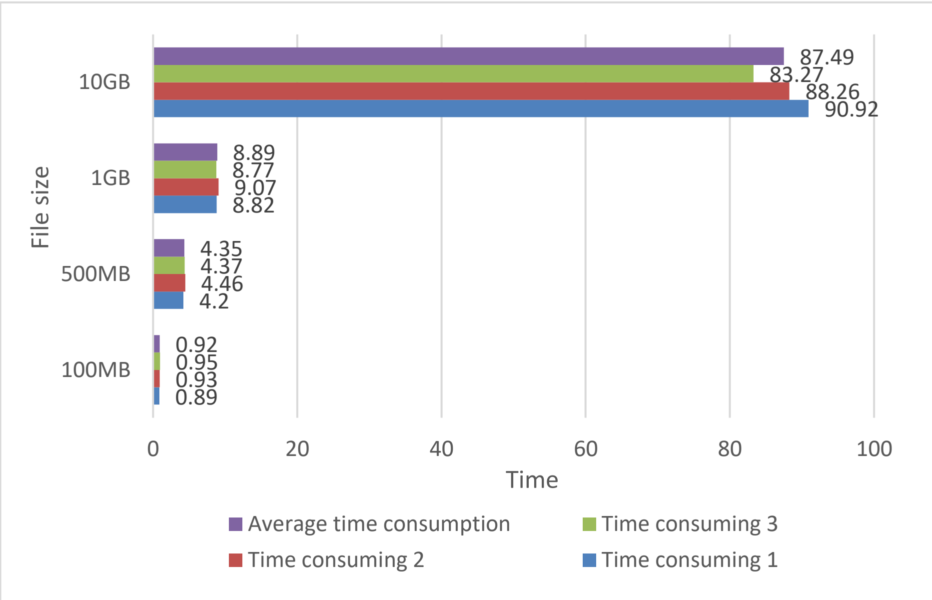

Fig.2 Synchronization Performance Test Results

5. Conclusions

With the development of "Internet plus education", the informatization construction of higher vocational colleges is more comprehensive. Experimental teaching is an important part of the teaching process, and should take a place in the informatization construction. This article designs an information management system based on data mining. This system can not only reduce the workload of teachers, but also stimulate students' learning enthusiasm, laying the foundation for improving teaching quality. At the same time, after testing the functionality and performance of the management system, as well as data synchronization, the test results meet the design expectations.

References

[1] Ping T , Han W , Peng L ,et al.Teaching Management System with Applications of RFID and IoT Technology[J].Education Sciences, 2018, 8(1):26-.DOI:10.3390/educsci8010026.

[2] Wang F .Thoughts on building an intelligent integrated management system of agricultural environment[J].Journal of Physics Conference Series, 2021, 1852(3):032047.DOI:10.1088/1742-6596/1852/3/032047.

[3] Chen X , Metawa N .Enterprise financial management information system based on cloud computing in big data environment[J].Journal of Intelligent and Fuzzy Systems, 2020, 39(1):1-10.DOI:10.3233/JIFS-189007.

[4] Gong W , Tong L , Huang W ,et al.The Optimization of Intelligent Long-distance Multimedia Sports Teaching System for IOT[J].Cognitive Systems Research, 2018, 52(DEC.):678-684.DOI:10.1016/j.cogsys.2018.08.005.

[5] Xinhan N .Intelligent analysis of classroom student state based on neural network algorithm and emotional feature recognition[J].Journal of Intelligent and Fuzzy Systems, 2020, 40(1):1-12.DOI:10.3233/JIFS-189545.

[6] Somasundaram M , Junaid K A M , Mangadu S .Artificial Intelligence (AI) Enabled Intelligent Quality Management System (IQMS) For Personalized Learning Path[J].Procedia Computer Science, 2020, 172:438-442.DOI:10.1016/j.procs.2020.05.096.

[7] Anjali P , Harshita N , Akash M D N ,et al.VMATE – Intelligent E-Learning Management System for Educational Institutions[J].IOP Conference Series Materials Science and Engineering, 2021, 1042(1):012011.DOI:10.1088/1757-899X/1042/1/012011.

[8] Zhao L , Chen L , Liu Q ,et al.Artificial intelligence-based platform for online teaching management systems[J].Journal of intelligent & fuzzy systems: Applications in Engineering and Technology, 2019(1 Pt.1):37-37.

[9] Diao S .The Reform of Teaching Management Mode Based on Artificial Intelligence in the Era of Big Data[J].Journal of Physics: Conference Series, 2020, 1533(4):042050 (7pp).DOI:10.1088/1742-6596/1533/4/042050.

[10] Almufarreh A , Arshad M , Mohammed H .An Efficient Utilization of Blackboard Ally in Higher Education Institution[J].Intelligent Automation and Soft Computing, 2021, 29(1):73-87.DOI:10.32604/iasc.2021.017803.

[11] Hong H , Tsangaratos P , Ilia I ,et al.Application of fuzzy weight of evidence and data mining techniques in construction of flood susceptibility map of Poyang County, China[J].Science of The Total Environment, 2018, 625(JUN.1):575-588.DOI:10.1016/j.scitotenv.2017.12.256.

[12] Marozzo F, Talia D , Trunfio P .A Workflow Management System for Scalable Data Mining on Clouds[J].IEEE Transactions on Services Computing, 2018, 11(3):480-492.DOI:10.1109/TSC.2016.2589243.

[13] Marco,Diana.Studying Patterns of Use of Transport Modes through Data Mining: Application to U.S. National Household Travel Survey Data Set:[J].Transportation Research Record, 2018, 2308(1):1-9.DOI:10.3141/2308-01.

[14] Shi Y , Fu D , Zhou X ,et al.Data mining to online galvanic current of zinc/copper Internet atmospheric corrosion monitor[J].Corrosion science, 2018, 133(APR.):443-450.DOI:10.1016/j.corsci.2018.02.005.

[15] Sattarian M , Rezazadeh J , Farahbakhsh R ,et al.Indoor navigation systems based on data mining techniques in internet of things: a survey[J]. Wireless Networks, 2018(2):1-18.DOI:10.1007/s11276-018-1766-4.