Intelligent Computing Technology and Automation
Z. Hou (Ed.)

doi:10.3233/ATDE231298

Teaching Information Management System Design Based on Decision Tree

Shuaiting YUE[a,1]
[a] *Shaanxi Institute of International Trade&Commerce, Xi'an, 13123, China*

Abstract. In order to improve the informatization level of teaching management in colleges, a system solution based on decision tree and WebService is proposed. Based on the demand analysis and credibility analysis of the teaching management system, this paper puts forward the logical architecture of system and establishes a UML model for the key teaching management process. Then, the improvement of the decision tree algorithm is applied to the core teaching evaluation problem through the bias problem of simplifying the ID3 calculation formula and feature values. The detailed implementation process is provided for the key modules of the system. In the system test and analysis based on J2EE and MSSQL2016, the results show that the scheme can realize the sharing and mutual access of educational resources. Its rapid calculation of information gain and the selection of eigenvalues also meet our expectations and it has the advantages of usability, reliability and efficiency.

Keywords. teaching information management; WebService; decision tree; ID3; J2EE; MVC

1. Introduction

In recent years, with the rapid development of network technology and multimedia technology, the situation of educational informatization is becoming more and more intense, and people's research on network education is also deepening. A large number of educational resources such as courseware, video, case question bank, literature and website courses exist in the network. The traditional teaching management method is difficult to quickly convey and process all kinds of information, which requires the emergence of new teaching management methods among campuses On the premise of ensuring data and information security and unchanged original infrastructure, how to achieve data sharing among campus departments and seamless communication connection between platforms [1]. Regardless of definition or implementation, during the web service process, the server provides a machine readable description (usually based on WSDLs) to identify the web services provided by the server. Web services technology is a new distributed computing model based on open Internet, which represents the inevitable trend of next-generation network computing and enterprise application. It is accessed through general web protocols and data formats, such as HTTP, XML and soap. Once deployed to the Internet, users can call web services on any platform and in any programming language. Web services is regarded as the best means to solve the loose coupling integration of e-commerce, government affairs and other applications [2]. How to effectively utilize these data, analyze the factors that affect student cultivation, and improve student cultivation methods is the key to the

[1] Corresponding Author: Shuaiting Yue; Shaanxi Institute of Intrenational Trade&Commerce, Xi'an, 13123, China; yst2011@1163.com

development of universities. Therefore, it is necessary to combine the teaching data of universities and use data mining technology to convert current teaching management data into usable information. Through data analysis, the internal factors that affect student performance are excavated, providing theoretical guidance for the improvement of student cultivation methods in universities, Provide decision-making basis for teaching reform [3].

This paper discusses the feasibility of developing teaching management system by using the combination of WebService and J2EE, and makes preliminary preparations in combination with the needs of teaching, management and resource storage. We develop it through the standard framework of SSH + MVC. The core of the system design is establishing a management model for teaching analysis and evaluation based on decision tree algorithm, and extract corresponding rules. Finally, Apache Axis2 is used to implement web services to realize the basic function of publishing educational resources on the Internet. On this basis, the system is tested and the results show that it can better meet the current requirements of colleges and universities for teaching information management, and it has simple structure, low investment cost and good expansibility.

2. Feasibility Analysis of Teaching Information Management System

2.1 Demand Analysis

For the teaching management system of colleges and universities, research is an important link in the stage of demand analysis and feasibility study Before investigating the system, it is necessary to clearly and comprehensively understand the potential software crisis and various factors affecting the system, and then use scientific, rational and mature research methods to collect and analyze information. The main requirements of this teaching management system include the follows:

(1) Generally, only system administrators and A/C set supervisors are allowed to enter the system management interface; All data backup work shall be recorded in detail and archived by the system administrator; The administrator is responsible for the authority management of new and old users and the daily maintenance of system development and operation;

(2) Any student can register as a fixed user of the system through his real name; basic information, detailed information and contact information shall be filled in during registration;

(3) Registered users automatically access the corresponding management interface according to their own roles; Students, teachers and administrators search for course information and management information according to their identity

(4) The client runs under the whole mobile platform; The server side runs on Linux platform.

(5) The system establishes the log query function to record and monitor the core data. At the same time, we should pay close attention to whether the operation behavior of key personnel contacting this kind of data meets the system specifications.

2.2 Feasibility Analysis

Starting from the plan of system development, this paper discusses the feasibility of system development force, and demonstrates whether various technical means adopted in the system scheme can be fully realized, as well as the impact of technical development on system construction [4-6]. In constructing components, it should be allowed to effectively develop, test and introduce new system functions without hindering services. J2EE includes all aspects of Java Web development, such as JSP, servlet, JDBC, JNDI, JavaBean, EJB, etc. Database development tools include MSSQL, Oracle, mysql, DB2, etc. These can be used in the development of management information system and database.

2.3 Decision Tree and Its Implementation in Teaching Evaluation

Decision tree algorithm is a machine learning algorithm based on tree structure, used to construct decision tree models from data. The decision tree model classifies or predicts data through a series of judgment conditions. In the decision tree, each internal node represents a feature or attribute, each branch represents a value of the feature, and each leaf node represents a category or prediction result. Decision trees construct a mutually exclusive and complete decision tree by learning sample data and using a certain generation algorithm, and then classify the processed samples. By constructing a decision tree model, the teaching quality of teachers can be evaluated based on characteristics such as teacher's teaching evaluation, student feedback, and teaching resource usage, such as high, medium, or low teaching level. At the same time, the decision tree algorithm has the advantages of strong interpretability, easy understanding and implementation, making it widely used in teaching evaluation in universities.

ID3, as a classic decision tree algorithm, is based on information entropy to select the best test attribute. It selects the attribute with the maximum information gain value in the current sample set as the test attribute [7]. The division of the sample set is based on the values of the test attributes, and the number of sub sample sets can be divided according to the number of values of the test attributes; At the same time, new leaf nodes grow on the nodes corresponding to the sample set on the decision tree. According to information theory, the ID3 algorithm uses the uncertainty of the partitioned sample set as a measure of the quality of the partitioned sample subset, and uses the "information gain value" to measure uncertainty - the larger the information gain value, the smaller the uncertainty, which prompts us to find a good non leaf node for partitioning.

We assume a data sample set S, which contains s data samples. Assuming that the category attribute has m different values (judgment indicator): $\mathrm{C_i}\ (\mathrm{i}=1,2,3,...,\ \mathrm{m})$. $\mathrm{S_i}$ is the number of samples in $\mathrm{C_i}$, and the total information entropy for a sample set is:

$$\mathrm{I}(\mathrm{s}_1,\mathrm{s}_2,...,\mathrm{s}_m)=-\sum_{i=1}^{m} P_i \log_2 P_i \tag{1}$$

Among them, P_i represents the probability that any sample belongs to $\mathrm{C_i}$, which can also be estimated using $\mathrm{s_i}/s$. We assume that an attribute A has k different values $\{\mathrm{a}_1,\ \mathrm{a}_2,...,\ \mathrm{a}_\mathrm{k}\}$, and use attribute A to divide the data sample S into k subsets $\{\mathrm{S}_1,\ \mathrm{S}_2,...,\ \mathrm{S}_\mathrm{k}\}$, where $\mathrm{s_j}$ contains samples from the set S where

attribute A takes the value of a_j. If attribute A is selected as the test attribute, then these subsets are new leaf nodes grown from the nodes of set S.

Assuming S_{ij} is the number of samples with category C_i in subset S_j, the information entropy value of the samples divided based on attribute A is:

$$E(A)=\sum_{j=1}^{k}[\frac{s_{1j},s_{2j},...s_{mj}}{s}\times I(s_{1j},s_{2j},...s_{mj})] \quad (2)$$

The information entropy gain obtained by dividing the sample set S using attribute A is:

$$Gain(A)=I(S_1,\ S_2,...,\ S_m)-E(A) \quad (3)$$

3. Teaching Information Management System Design

3.1 The Overall Structure

The system as a whole adopts JavaWeb, front-end framework bootstrap, based on SSM framework, MVC three-tier design, based on B/S mode, Eclipse or idea as development tools. The specific design is shown in figure 1. To keep different course data resources from interference with each other and ensure that there will be no impact on each other when publishing a large number of courses [8]. The server uses Ubuntu server LST 16.04, and the local client system is Windows 10. The course video files store big data through HDFS, and the large files are divided into blocks and stored in a distributed way, which solves the problem that a single machine cannot store large files. The main structure and process of the software are described with the framework of main program, subroutine and sub process, and the input and output link relationship between each framework is defined and debugged to ensure the high coupling between each development module in the development process.

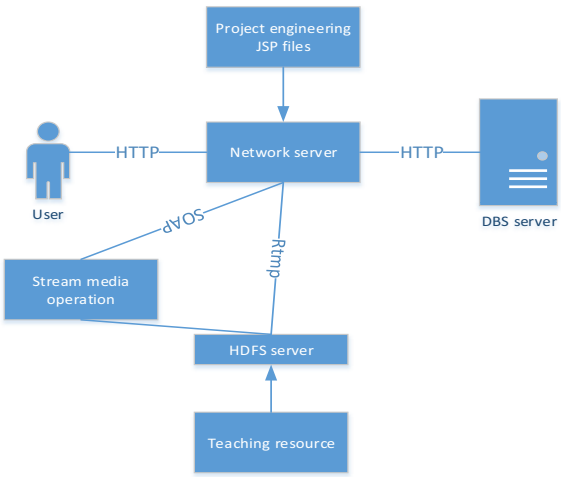

Figure 1. Technical framework of teaching management system.

3.2 Key Module Design

(1) Improvement of ID3 Algorithm

There are multiple ways to improve the ID3 algorithm, and we use mathematical formula derivation to simplify the calculation formula of information gain. Let a feature attribute have X forward and Y reverse directions, and its information entropy formula is:

$$H(D,A)=\sum_{i=1}^{n}\frac{x_i+y_i}{x+y}\left(-\frac{y_i}{x_i+y_i}\ln\frac{y_i}{x_i+y_i}-\frac{x_i}{x_i+y_i}\ln\frac{x_i}{x_i+y_i}\right) \tag{4}$$

Simplify the external constants by using the formula

$$\ln(1+x)\approx x \tag{5}$$

$$\ln\frac{y_i}{x_i+y_i}=\ln\left(1-\frac{x_i}{x_i+y_i}\right)\approx\frac{x_i}{x_i+y_i} \tag{6}$$

$$H(D,S)=\sum_{i=1}^{n}\left(-y_i\ln\frac{y_i}{x_i+y_i}-x_i\ln\frac{x_i}{x_i+y_i}\right)\approx 2\sum_{i=1}^{n}\frac{x_iy_i}{x_i+y_i} \tag{7}$$

Finally, we can use the following equation to calculate the average entropy, selecting the root or inner node with the lowest value.

$$\sum_{i=1}^{n}\frac{x_iy_i}{x_i+y_i} \tag{8}$$

(2) Course management module

The course management module page consists of the main functions of displaying courses, listing, course name and status, querying courses, putting courses on and off shelves, etc. Display the corresponding chapter and class information of the course in the form of drop-down box of tree structure. At the same time, the module also provides teachers with more than ten kinds of curriculum activities based on the teaching design concept of social constructivism, including discussion area, learning log, chat room, learning resources, unit summary and online survey. Through the data service layer, the data is obtained from the database and handed over to the action layer for operation.

(3) Work module design

The UML diagram of functional design of this module is shown in figure 2.

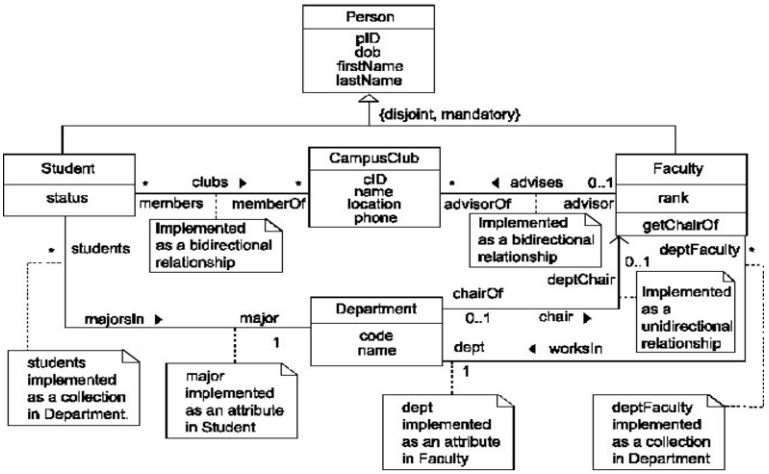

Figure 2. UML class diagram of job management.

After the teacher's account is registered and logged in, create a class, enter the course name, course nature and other information, assign homework or manage the homework uploaded by students. This process can instantiate an httpget object, pass the requested URL address to the httpget object through the construction method, and the request is captured by the spring front-end control servlet dispatcher servlet; Dispatcherservlet parses the request URL and feeds back the result to the user. *Loggingcache* stores the parity information that needs to be modified into the log file in the form of pipelining, so as to aggregate multiple writes of small blocks into one write of large blocks.

(4) Apache Axis2 extensible module

Apache Axis2 is a subsequent version of axis and a new generation of soap engine. Without any configuration, a simple POJO can be directly published as a WebService, in which all public methods in the POJO will be published as WebService methods and the client code needs Axis2 jar packages. If we are not sure which packages are needed, we should protect all Axis2 jar packages; Soap request messages flow through outflow We will create a new ordinary java project and two classes, one is module and the other is handler. The module class must implement org apache. axis2. modules. Module interface, we don't need to implement any interface method, unless the extension module has initialization and destruction operations. Due to the use of modules, modules with other functions can be added to the system without changing the configuration file. Part of the key codes of such process are described as follows:

```
<servlet>
<servlet-name>AxisServlet</servlet-name>
<servlet-class>org.apache.axis2.transport.http.AxisServlet</servlet-class>
<init-param>
<param-name>contextConfigLocation</param-name>
<param-value>classpath:conf/axis2.xml</param-value>
</init-param>
  <load-on-startup>2</load-on-startup>
</servlet>
  <servlet-mapping>
<servlet-name>AxisServlet</servlet-name>
<url-pattern>/services/*</url-pattern>
  …
```

3.3 DBS Design

After connecting the database management system through JDBC, a connection object is obtained. From this object, you can obtain various information about the database management system, including various tables in the database, columns in the table, data types, triggers, stored procedures and other information. This coefficient uses the database standard language server 2016 database. The data will not be stored elsewhere in the database. Taking the user registration data table as an example, its logical structure is shown in table 1.

Table 1. User registration data list

Fieldname	Description	Datatype	Length	Remarks
UserID	User code	Int	41	Key
UserPWD	User name	varchar	15	Not null
UserName	Password	varchar	20	Not null

Usertype	User type	varchar	10	Not null
UserLevel	Access privilege	varchar	10	Not null

4. System Implementation and Test analysis

After the data preparation stage, an algorithm is first used to establish a decision tree, and an important concept in information theory is introduced here. Information gain is used as a measure of attribute selection, and the attribute with the highest information gain is selected as the selection criterion for attribute selection. In this article, the effective sample size of the system is 300, and there are a total of 6 attributes, including the Advanced Mathematics course group(AM), English course group(En), Computer Network(CN) course group, Political course group(Pl), Physics course group, and Physical education course group(PE). Each attribute value is in {A, B, C}, Calculate the expected amount of information for the " "advanced mathematics" attribute in the sample according to the formula in the previous section. At the same time, by calculating the information gain of each attribute, identify the attribute with the highest information gain value as the test attribute.

Select the attribute with the highest information gain as the node for this attribute selection. Calculate the attribute information gain of each branch attribute node using the same method as before. Repeat the above steps to complete the division of each branch, and finally establish a decision tree model, as shown in Figure 3.

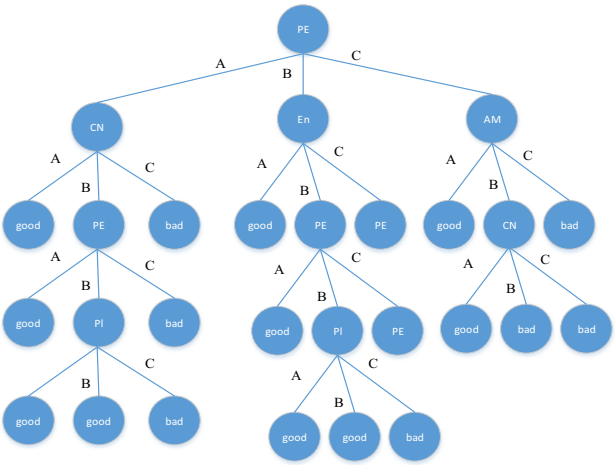

Figure 3. User login teaching management system main interface.

According to the previous module function analysis, the interface design of this system is mainly divided into two parts The front desk student user interface is mainly for users to browse and query course information online. After logging in, users can get different learning interfaces according to their permissions, as shown in figure 4. The main page and main menu describe the overview of the course and the operations that can be performed in the key areas. According to the actual needs, the system provides more simple and easy-to-use system operation and more comprehensive and perfect resource services. It can carry out many common functions, such as course video learning, homework submission, online interaction, forum posting, background

management and so on. To make administrators and users more intuitively understand the progress of teaching activities, the system realizes a visual data statistics interface. This module mainly uses the open-source front-end graphics rendering library Echarts to realize graphical data display. Figure 5 depicts the function of visual data statistics by taking the information of teachers and students as an example. It can be seen that the system realizes the automation of sorting and other calculation and statistics; This interface can query information by class and individual; Information can be released to the network to realize data sharing, so as to quickly and intuitively grasp the progress of current teaching activities.

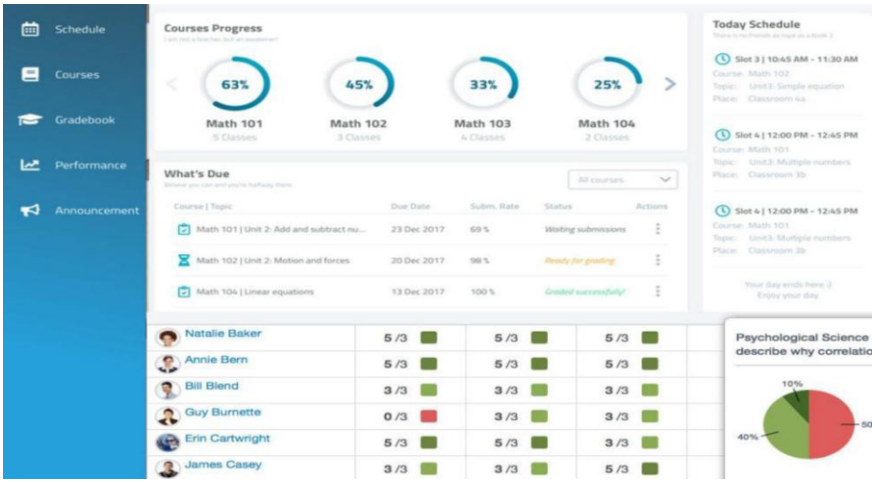

Figure 4. User login teaching management system main interface.

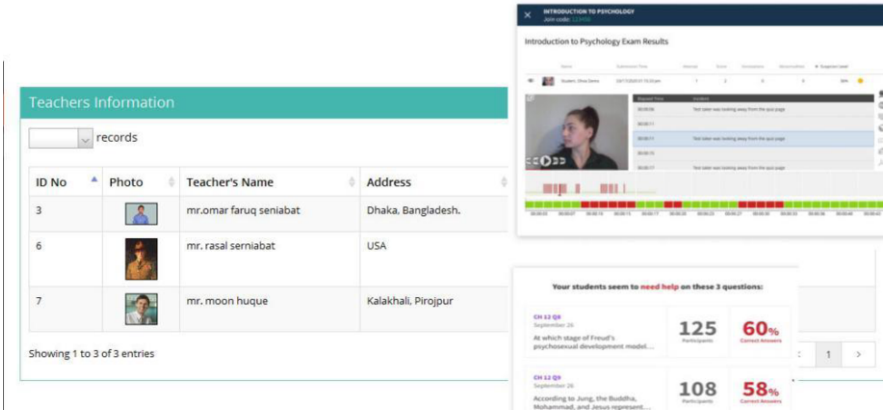

Figure 5. Administrator information statistics and viewing interface.

5. Conclusion

Through the research on the comprehensive management system of students and teaching in a university, considering the different needs of the University in different development stages, this paper puts forward a relatively complete design scheme based

on WebService and decision tree algorithm. The system selects appropriate features based on the requirements of the task and the characteristics of the data to construct a decision tree model, and uses decision tree algorithms to learn knowledge and extract rules from the data of the teaching management platform. Then an object oriented teaching information management system is designed combined with MVC and SSH, and the learning resources are effectively stored for online teaching by using MSSQL and HDFS. The core function is to interpret and understand the rules obtained through ID3 algorithm, to understand the conditions and conclusions represented by each rule, and the relationships between them, which helps top reasonably optimize the daily teaching work evaluation of teachers based on the identified key factors. Through the functional test of each functional module of the system, it is verified that the system meets the preset functional requirements and can realize the online teaching interaction based on Web.

References

[1] Zhang Hongui, Dong Wei, Xu Dongxia, et al. Design and realization of educational management information system based on web service. Journal of Shandong University of Technology (Sci & Tech), 2006(5):84-86

[2] Chen J. Design and Accomplishment to Adult Instructional Administration System Based on Web Services. Journal of Huizhou University, 2006, 26(3): 46-49

[3] Xiao-Qi Z. Analysis of Teaching Management Data Based on Decision Tree Algorithm. Journal of Qingdao University(Natural Science Edition), 2019, 32(2): 86-93.

[4] Du D J, Shao H L. Design of College Teaching Management Information System Based on J2EE Framework. Applied Mechanics and Materials, 2014, 651-653:2453-2456

[5] Lu P, Cong X, Zhou D. E-learning-Oriented Software Architecture Design and Case Study. International Journal of Emerging Technologies in Learning, 2015, 10(4): 1-6

[6] Yong M. Advanced Mathematics Teaching Evaluation Model base on Decision Tree Algorithm. Journal of Convergence Information Technology, 2013, 8(9):602-608.

[7] Hang H U, Shuang D U, Jiarou L, et al. Towards a Prediction Model of Learning Perfor-mance:Informed by Learning Behavior Big Data Analytics. Academic Abstracts of Chinese Higher Education Institutions - Pedagogy, 2022, 17(1):121-156.