

Design and Implementation of Intelligent Teaching Platform based on Spark Big Data Analysis Technology

-- Take Big Data Courses as an Example

Yong Wang, Ting Zhang

School of Information Engineering, Wuhan Technology and Business University, China

Abstract

Currently, Spark is widely used in big data analysis, but the application of big data analysis, especially Spark, in education is relatively rare. The large-scale and massive characteristics of teaching resources make the current intelligent teaching platform have low performance and low utilization rate of teaching resources when processing and analyzing teaching quality. Therefore, this paper proposes to build an intelligent teaching platform based on Spark big data analysis technology. Based on current Spark big data analysis technology, first of all, through the analysis of the current research progress of intelligent teaching platform, design the overall structure of the intelligent teaching platform, immediately introduces the intelligent teaching platform of large data recommended modules and teaching evaluation module design and implementation, finally the intelligent teaching platform of the actual teaching situation are expounded and summarized. Years of teaching practice experience of big data courses in our school shows that the intelligent teaching platform designed in this paper based on Spark improves the utilization rate of teaching resources and can recommend personalized learning resources according to the actual situation of students, achieving good teaching effects.

Keywords

Spark; Big Data Analysis Technology; Teaching Resources; Teaching Quality Assessment; Intelligent Teaching Platform; Resources to Recommend.

1. Spark based Big Data Analysis Technology

With the development of the Internet, tens of thousands of web pages continue to emerge, and to search these pages to grab storage, and then analysis and calculation. However, due to the increasing amount of data, storage is faced with insufficient capacity of a single machine and time-consuming query. To solve these two problems, Google proposed distributed storage and distributed parallel computing solutions, and the later Hadoop products are the source code implementation of this solution.

Hadoop provides HDFS, a distributed file platform, and MapReduce, a distributed parallel computing framework. With the development of Hadoop, its ecological platform keeps emerging new project products, such as HBase, Hive, Pig, etc. However, both of them are based on the HDFS storage layer and MapReduce computing framework. MapReduce performs computation on multiple nodes in the cluster in parallel, thus greatly speeding up the query speed. However, with the increasing amount of data, MapReduce gradually becomes inadequate. Therefore, Spark computing framework based on memory computing comes into being. The query speed of Spark is 100 times faster than Hadoop, so it is the most advanced distributed parallel computing framework at present. With the development of Spark ecosystem platform,

Spark SQL, Spark Streaming, MLlib, Graphix, etc. Spark SQL is a tool developed for SQL users to analyze and query structured data using SQL language.

Currently, the mainstream big data analysis platforms are Mainly Hadoop and Spark. To be more precise, Spark is only a parallel computing framework, while Hadoop includes the computing framework MapReduce and distributed file platform HDFS. Hadoop more broadly includes other platforms on its ecosystem, such as HBase, Hive, etc. The Spark ecosystem platform has evolved into a collection of subprojects, including Spark SQL, Spark Streaming, Graphix, MLlib, etc. Spark is an alternative to MapReduce with nearly a hundred times more performance. Spark SQL is a query engine based on Spark that provides SQL query and analysis of structured data. It is an alternative to Hive based on MapReduce.

2. Overview of Intelligent Teaching Platform based on Spark Big Data Analysis Technology

In the era of big data, Statistical analysis, mining and utilization of data resources based on Spark big data analysis technology is attracting more and more attention from various industries. It is used to collect consumer preferences, guide production and product development, and greatly improve enterprise production efficiency.

However, the application of big data analysis technology, especially that based on Spark, in the field of education is relatively rare. The large-scale and massive characteristics of teaching resources make the current intelligent teaching platform have low performance and low utilization rate of teaching resources when processing and analyzing teaching quality.

At present, massive data such as test scores, homework scores, experimental scores and employment information of all previous students are stored in the university's undergraduate teaching database, which simply records the data information, but it is difficult to intuitively discover the hidden information from these data. In fact, no matter between courses, or between knowledge points in each course, as well as between students' scores and course Settings, there are numerous links, and the existing data has not played its real value at the present stage.

How to build an analysis platform based on Spark teaching big data, collect and statistically analyze teaching data in real time and efficiently, display analysis results intelligently and visually and push teaching suggestions, so as to continuously optimize and improve the teaching mode iteratively and achieve the goal of intelligent teaching?

3. Design and Implementation of Intelligent Teaching Platform based on Spark Big Data Analysis Technology

3.1. Overall Platform Design

Based on the characteristics of teaching big data and intelligent teaching mode in the context of big data, this paper designs an intelligent teaching platform of big data analysis technology with multi-layer architecture. The overall functional structure of the platform is shown in Figure 1.

The platform structure is divided into four layers: client layer, presentation layer, service layer and resource layer.

1) Client: including students, teachers and administrators. Accessed by Browser, the server is composed of four parts: presentation layer, application service layer, public service layer and resource layer.

2) Presentation layer: its main function is to provide personalized interface and content for the client, just like navigation home page. For other function modules, access the navigation page from the navigation home page. Identity authentication is required before entering the

navigation page. After passing the authentication, access the required navigation page. The user navigation page faces different users render different pages.

3) Service layer: This layer is composed of application service layer and public service layer. The main function of application service layer is to provide the functions directly related to teaching application, including learning module and communication module.

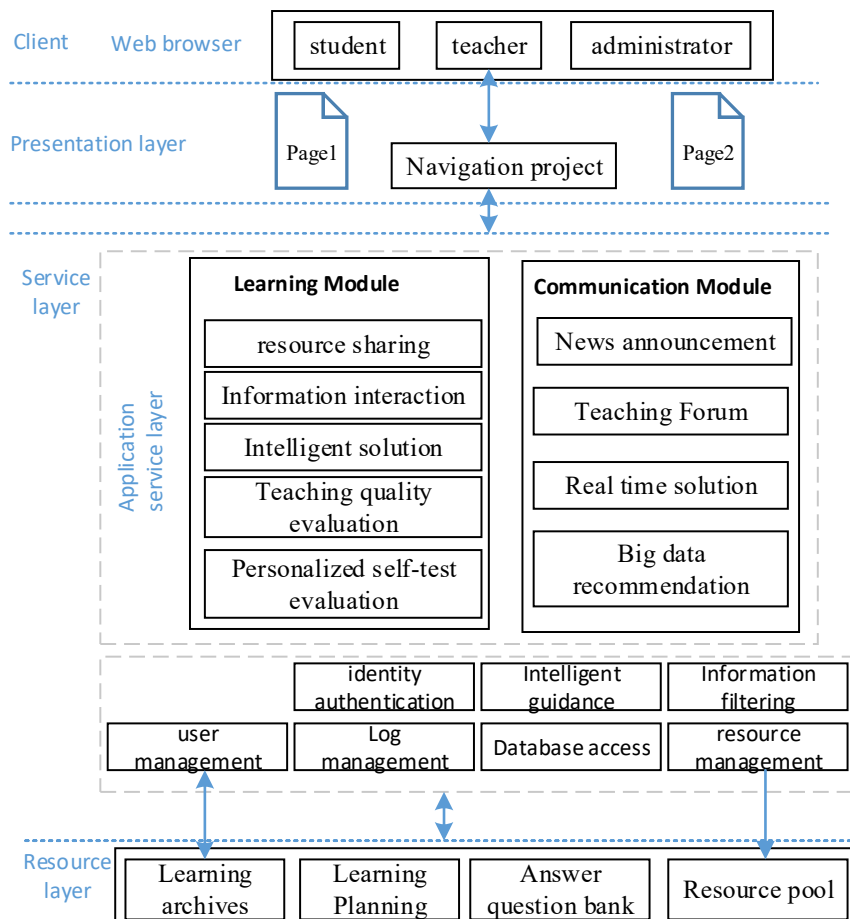


Figure 1. Overall functional structure of the platform

The main module of the platform is learning module, which can realize students' learning activities, including resource sharing, information interaction and teaching quality evaluation modules. The main way for users to communicate in the platform is the communication module, which is also the main functional module to help students learn and communicate with each other, including online real-time solutions, big data module recommendation and teaching forum modules. The module also supports real-time communication and non-real-time communication. The important teaching function layer of the platform is the application service layer. The expansion of teaching function can be completed by adding functional modules.

The common service layer and the application service layer together constitute the service layer. It is the public service layer that provides the basic public service for the application service layer, which is unrelated to the teaching application. It consists of information filter, intelligent guidance, resource management, user management and identity authentication, log management and other modules. The main functions of its module are as follows:

Information filter: through the user's learning status, learning planning, complete intelligent information screening, get personalized resources or information to the user. Prevent users from facing huge difficulties in choosing teaching resources.

Intelligent guidance: This module mainly provides intelligent help or guidance for users in the platform. When users use the platform, they can provide personalized and timely guidance or help for users, and help learners to realize the learning process anytime and anywhere, so as to prevent users from having no direction when using complicated platform functions.

Resource management: upload and download resources in this module, can manage courseware, test questions and learning planning and other resources. User management and authentication: Provides management of user information and roles.

Log management: Provides user operation and learning log management for the platform.

4) Resource layer (data layer): used to store and store all kinds of received data. Including user information, learning archives, resources and questions, etc. The file that records the knowledge level that each student can master through the platform is the learning file of the log, and also an important data basis for the implementation of personalized information screening and mining.

Each of the four modules works together to personalize the strategy for the entire platform. The personalization of user interface is realized by navigation. The personalization of learning resources, recommendation and activities is realized by learning module of service layer. The individuation of communication is completed by communication module of service layer. Personalized guidance assistance is accomplished through intelligent guidance.

3.2. Big Data Recommendation Module

The three parts of user feature generation, information matching and recommendation generation constitute the big data recommendation module, as shown in Figure 2. By collecting massive amounts of user interest and usage demand data, performing calculations and selections based on recommendation algorithms, counting and extracting scores to obtain the best recommendation results.

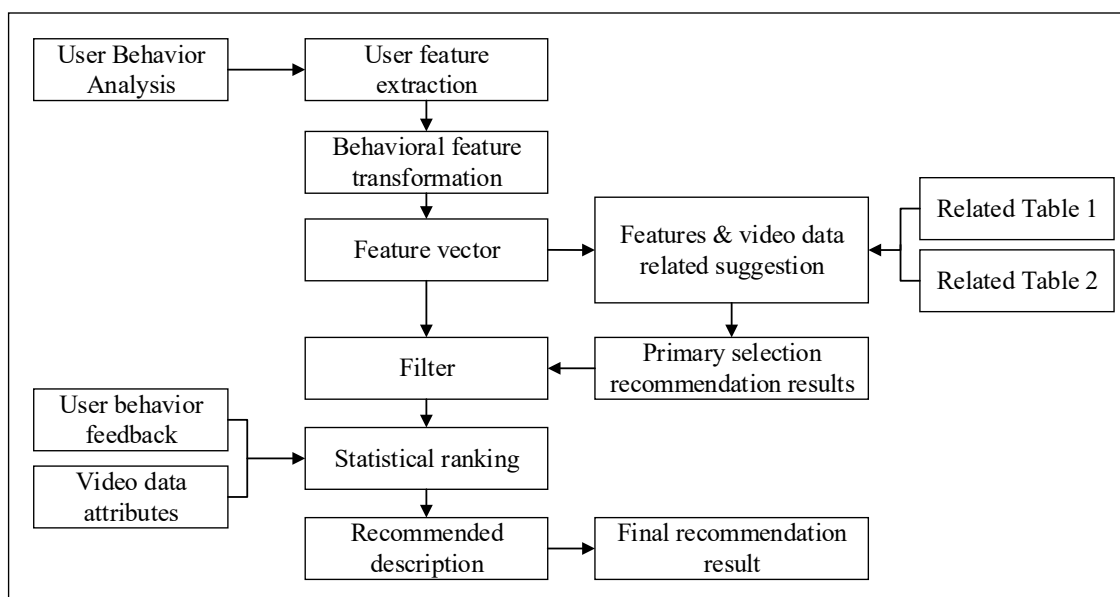


Figure 2. Big data recommendation module

The big data recommendation module includes three analysis processes:

- 1) Generate user feature vectors: user behavior features include personal hobbies and usage data, etc. The system will convert user behavior features into data and generate feature vectors;
- 2) Primary selection recommendation results: According to the data category and preliminary selection, the system selects data similar to the user feature vector and generates a data table;
- 3) Final recommendation results: The module performs statistics and arrangement on the data tables selected in step 2) to obtain the final recommendation results.

Comprehensive consideration of the performance, cost and load capacity of the system's big data recommendation module to ensure smooth operation of the teaching system and more convenient operation for users. At the same time, it is convenient to formulate the name and detailed form of each field of the log when creating the log. After the user log is successfully pushed, subsequent analysis can be performed and the data can be saved in the database. Pass points the method of layer design processes user data: the first layer is the report; the second layer is analysis of user habits; the third layer is the information that needs to be promoted. User data points after the analysis is implemented, use offline and real-time methods to complete the content push, and push content to users; real-time push means push content to user.

3.3. Multi-dimensional Teaching Quality Assessment Module

- 1) Establish a Multi-dimensional assessment and evaluation system:

The teaching quality assessment module has good interactivity and usability, and can provide functions such as teaching feedback, classroom discipline supervision and intelligent attendance. The student terminal, classroom terminal, educational administration terminal, teacher terminal and cloud server constitute the teaching quality evaluation module. Each terminal and cloud server use mobile network communication to connect and utilize the development of client and classroom terminal. Through the combination with cloud server and the construction of background data, a teaching quality evaluation module is formed.

The assessment and evaluation standards of big data courses should correspond to the teaching analysis system, and it should be carried out in stages. Within the specified time, students complete the phased goals of the task, using a combination of student self-evaluation, student mutual evaluation, teacher comments, etc. The method summarizes and evaluates the students' learning situation, and gives appropriate scores, which are converted in a certain ratio with the final exam results. The final result can reflect the students' mastery of theoretical knowledge and the students' practical hands-on the final comprehensive score of ability.

The teaching quality evaluation module has good interactivity and usability, which can be provides functions such as teaching feedback, classroom discipline supervision and intelligent attendance. Teaching the structure diagram of the quality assessment module is shown in Figure 3.

From student terminal, classroom the terminal, educational administration terminal, teacher terminal and cloud server constitute the teaching quality evaluation module. Each terminal and cloud server use mobile network communication for connect. Utilize the development of client and classroom terminal, through and cloud server combined with the construction of background data, the teaching quality evaluation module is formed. Should the module collect data from the classroom terminal through image processing technology. In addition, the module is also used for data transmission among clients, classroom terminals and databases. The classroom terminal is set in the classroom, including network communication sub-modules, data collection sub-module, terminal management sub-module and frame processing sub-module. Cloud service the server transmits the management signal to the network communication sub-module, and the network communication sub-module the module transmits the management signal to the terminal management sub-module; the data collection sub-module collects classroom data according to the management signal, which includes

images and video frequency; the frame processing sub-module intercepts the collected videos into frames according to time sequence image, judge the current frame image, if the current frame image and the previous frame the similarity of the image is less than the set similarity threshold, delete the current frame like, complete the frame acceleration and processing of the video stream [9]; network communication sub-module the video and image processed by the frame processing sub-module and the collected image and video the frequency is transmitted to the cloud server for storage.

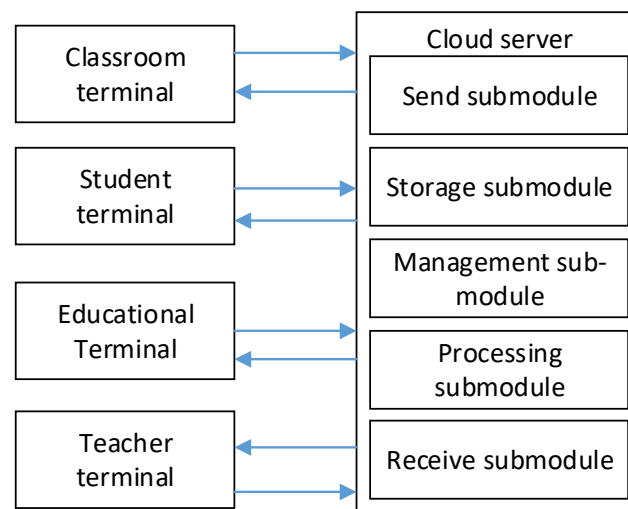


Figure 3. Teaching quality evaluation module structure

The client terminal includes teachers, educational administration and student terminals, and the main functions include send, receive and display, use the cloud server to receive and send from each terminal image and text information, and the number returned by the cloud server passes through each terminal receive and display [10-11]. The main function of the cloud server is to carry out the analysis of attendance, attendance rate and classroom behavior, and finally analyze the result is fed back to the client through the feedback unit.

Big data evaluation standard classes should correspond to teaching analysis platform, take the stage type, within the prescribed period of time, students' progress goal to complete the task, using the self-evaluation of students, students mutual and teacher comments on a combination of the summary and evaluation on the students' learning situation, and give appropriate scores, It is converted to the final exam result in a certain proportion, and the final comprehensive score is obtained, which can reflect students' mastery of theoretical knowledge and practical ability.

2) Establish an iterative teaching improvement mechanism

The processing sub-module is composed of the attendance analysis unit, the listening rate analysis unit, the classroom behavior analysis unit and the feedback unit. These four units are mainly used for the analysis and feedback of classroom data.

Discover the laws and causes behind the fluctuations in performance, form effective teaching feedback, and promote continuous iterative optimization and improvement of teaching methods, methods, teaching resources and other aspects of teaching to achieve the goal of intelligent teaching.

4. Conclusion

The intelligent teaching platform designed in this paper based on Spark big data analysis technology has achieved good teaching results in the teaching practice of big data courses (computer network, cloud computing technology, etc.) in the School of Artificial Intelligence of Wuhan Institute of Technology. The platform can not only be customized for different courses and teaching requirements on the teacher side, but also intelligently push pre-school resources, in-school teaching tests and post-school analysis results on the student side. At the same time, the massive teaching data can be automatically carding, analysis, mining the rules and causes behind the performance, the results of the analysis can be targeted for teachers teaching reflection to provide data support, so as to improve the teaching quality, the results of the analysis and provide data to support decision making for the teaching reform, make the decision more accurate and efficient teaching reform. In short, the intelligent teaching platform based on Spark big data analysis technology plays a crucial role in the research and development of intelligent teaching mode of big data courses.

Acknowledgments

Teaching and Research Project: Education and Teaching Research Project of Wuhan Institute of Business and Technology (Project Number: 2019Y07).

References

- [1] Liu ping. Research on campus big data platform based on Hadoop and Spark[J]. Software engineering, 2018,21 (5): 15-18.
- [2] Xu Chengjie, Xiao Xirong, Zhang Jingyi, et al. big data analysis based on Spark Design and Application of Platform [J]. Chinese Journal of Health Information Management, 2019,16 (5): 633-637.
- [3] Zhang Yi. Artificial intelligence precision teaching research based on teaching precision interactive behavior analysis [J]. Adult Education, 2019, 39(9): 83-88.
- [4] Ge Fuhong, Wang Yun. Research on the construction and application of hybrid teaching mode in colleges and universities based on intelligent teaching platform [J]. Modern Distance Education, 2020 (3): 24-31.
- [5] Zhu Shiming, Guo Yan. New engineering teaching reform in the context of smart education: framework design and implementation path [J]. Higher Engineering Education Research, 2019 (6): 155-161.
- [6] Ding Caicheng, YUAN Hongzhi, DING Jingmin. Curriculum diagnosis and reform in higher vocational colleges vocational and technical education, 2018(35): 34-38.
- [7] Huang Danxia, Liu Xinxin. Design of learner model based on large-scale online open courses Design and Realization[J]. Computer Application, 2018, 38(z2): 327-330.
- [8] Yao Lin, Shi Zhiguo. Research on Artificial Intelligence Curriculum System and Teaching Methods [J]. China University Learning and Teaching, 2019 (10): 19-22.
- [9] Song Ying, SONG Sean. Practice of distributed virtual reality based on personal intelligent terminal The design of teaching system[J]. Laboratory research and exploration, 2020, 39(2): 227-232.
- [10] Zheng Yueyuan, Chen Zehao, Liang Jiazhi, etc. Electron microscope simulation with intelligent feedback Teaching system[J]. Journal of Chinese Electron Microscopy Society, 2019, 38(2): 182-188.
- [11] Xu Peng, Wang Yining. Big data perspective analysis of learning revolution [J]. The remote teaching Journal of Education, 2013 (06): 11-17.
- [12] Wang Zhiguo, Li Yanxing. Teaching diagnosis reform in higher vocational colleges in the era of big data exploring the construction of informatization [J]. Information Technology and Informatization, 2017 (12): 119-121.

- [13] Feng Meixian, Lu Chun. Research on Teaching and Diagnosis Reform Based on Classroom Teaching Behavior--Take the five-year higher vocational "electronic technology foundation" course as an example [J]. Jiangsu Teaching Education, 2019(84): 39-42.
- [14] Ma Xing, Wang Nan. Construction of university teaching Quality Evaluation System based on big Data [J]. Tsinghua university education research, 2018,39 (02): 38-43.