

Personalized Teaching Platform Based on Web Data Mining

<https://doi.org/10.3991/ijet.v11i11.6253>

Lanzhong Wang
Shandong University, China

Abstract—The purpose of this study is to develop a distance personalized teaching platform. The web data mining is used for the construction of the system and by analyzing the character of web data mining (WDM) and the essence of personalization teaching and instruction, based on WDM. The system contains knowledge base, individual database, WDM and web server four modules. The web data mining is used for the construction of the system and by analyzing the character of web data mining (WDM) and the essence of personalization teaching and instruction. Simulation results show that model has important enlightenment and pushing effect for promoting the individual service and improving teaching quality of modern distance education.

Index Terms—personalized teaching platform, modern distance education, web data mining

I. INTRODUCTION

Modern distance education is a kind of new education method combining with modern information technical development; it has already become the trend with the international education development. Modern distance education platform is the base of realizing distance education, its performance and function determine the method, means, result and quality of distance education directly. The traditional distance education platform design does not follow certain technology standards, existing the "technology for technology" phenomenon. In addition, most existing distance education resources displays in writing and planar way, what it realized is a copy to original teaching materials. In this case, the learner obtains knowledge cannot get away from the mechanical type studying[1]. The emergence of VRML technology brings the distance education a new teaching means because of its little data to transfer and with vivid 3D pictures, providing a new technology support for designing modern distance education platform. Based on the research and studying of the national information technology standard committee, who puts forward the CELTS-1 criterion.

Ye's article [2] adopts the J2EE and VRML technology to design modern distance education platform, and realizes the main parts of it. His paper used VRML technology, this article puts forward the Modern distance education platform based on the way of information deliver and the changes of the way that teachers and students communicate. Contrasts and analyzes the prevail Web 3D technology, choosing the suitable development tool and the design mode and certain suitable design process.

Wang's paper [3] designs the Modern distance education platform based on J2EE and VRML according CELTS-1 criterion. Divides the platform into five parts, (containing teaching support system, teaching resources

management system, educational administration management system, course design system and other systems), analyzing all the part of design destination and function and realizes the main parts of modern distance education platform. Yu's article [4] lays emphases on the teaching video part in teaching support system, virtual classroom system and virtual experiment system. In addition, it realizes the teaching resources management system based on the project of "The construction of computer science and professional teaching resources". In the process of building modern distance education platform, Wang's article [5] applies some key technology and strategy: (1) Applies the role-dynamic state authorization mechanism into Modern distance education platform. He realizes the clear authorization management and higher flexibility and security. (2) Improves the project of "The construction of computer science and professional teaching resources" and introduces the secondary search of and related search strategy. (3) Designs the player imbeds the web page which can adapt playing in teaching video system. It can complete the playing task in practical use and test. (4) Applies fast-modeling method according to complications in the process of VRML, heighten the efficiency of practical modeling. (5) Classifies and sums up the technology of two animations the born technology building and the strategy of interaction, discussing the method and principle of each realization, getting applied in the process of designing virtual classroom and virtual experiment. (6) Discusses the way of optimize strategy based on 3D modeling of VRML technology finally [6-7].

However, Distance Learning is the outcome of applying computer and network communication technology into education field, this teaching mode is realized through network, and it have made great and far-reaching influence on education structure, and it is a popular trend of education development in future. At present, Distance Learning based on Web because of its character of flexibility, simple, reliability, compatibility have become an important mode and development way in the Distance Learning.

Distance learning comes with the computer multimedia and network communications being applying to education. It takes a significant effect on all of the education structure. The government provides the support for the development of distance learning. Distance learning based on Internet is one way of modern education. The basis platform is indispensable for teaching, FAQ, virtual experiment, real time talking, cooperating working, online test, affair manage. The system adopts the most popular B/S system structure, namely: Browser /web server/database server; B/S divides the system into three layers: The view layer, the business logic layer and the data access layer, and realizes to separate business logic, user interface and data access, which

improves the maintains of application system. The chapter of system realization introduces how to realize the system. In general, the view layer uses html JavaScript to realize; the business logic layer apply struts frame and the data access layer uses JDBC technology. LDES applies very popular technologies such as JSP, Struts, which provides a new learning and teaching way for students and teacher and makes teaching way information-based [8-9].

II. OVERVIEW

First, due to the specialty of Intelligent Community and Distance Education that the platforms not only provide Web service, but also provide other services such as video service, and what is more, the stability of platform is of great importance for the huge end-user accesses, the author put forward the method of adopting distributed clustering servers. After intensive study of knowledge of Distributed Systems and analysis of comparable products, the author realized the set-up of clustering servers under Linux, and developed relative software that can be used to monitor and manage clustering servers. Second, Intelligent Community needs broadband network service, which makes the author to bring about innovative idea of Content-based Search Engine and Intelligent Cache Agent. After searching and studying relative materials inland and abroad, the author finished initial research work of Cogent-based Search Engine and realized Intelligent Cache Agent. However, content-based search engine is a new idea, and is limited by difficulty of Literal Explicit of Chinese Words and irregularity of protocol of HTML; the advanced research work is still on its way to successful resolvent. Figure 1 shows the composition of data mining technology.

With the increasingly fierce competition of the telecom industry, Data Mining, as a measure of the Knowledge Discovery and the Decision Support, has already been widely used. However, the Human Turbulence and the Inhuman Turbulence on the telecom industry's Data Acquisition and Processing System which cause the Data Source's Data Quality differing largely and some problems of Data Warehouse's ETL mechanism which lead to the data missing and data mistake when the data is integrating that result in the low quality of data which affects the result of Data Mining. So solving the above-mentioned problems, improving the veracity and the efficiency of Data Mining and making use of Data Mining to improve the normal operations of services in the telecom industry have practical significance. Although the Data Mining technology has been studied for more than ten years and some technologies are suggested to try to solve kinds of problems that Data Mining technology in the telecom industry is encountered when it is applied, these technologies do not adapt to the complex and large-scale data in the telecom industry.

Figure 2 shows the basic procedure of data mining technology.

The object of web content mining and web structure mining are all primitive data in websites, while web usage mining is a kind of data mining that it mines the information of server logs after the user browse the web pages. This information include accessing date, time, user IP address, server IP address, accessed URL resources, server response state and transmitted bytes etc. Figure 3 shows the classification of websites data and web data mining [10-11].



Figure 1. The composition of data mining technology



Figure 2. The basic procedure of data mining technology

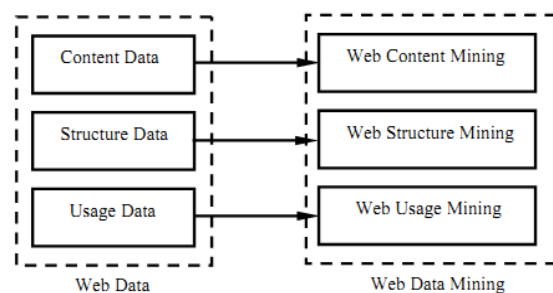


Figure 3. The classification of websites data and web data mining

Fig. 4 is the essence of personalization teaching and instruction. In Fig. 4, D denotes that different learners make use of learning resource by various ways, namely learning process. B is that the distance education system tracks learning process of different learners and builds accessing model and individual database. C is that the distance education system adjusts learning resource according to the dynamic model of users and meets personalization requirements of different learners. A is the most important aspect of personalization teaching and instruction in modern distance education.

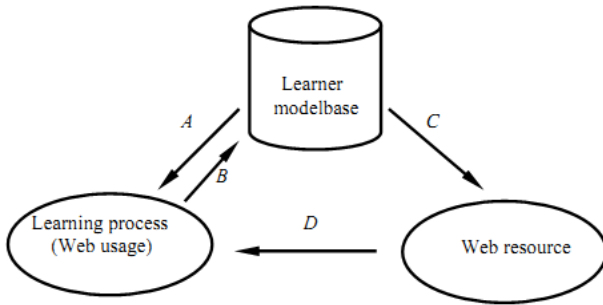


Figure 4. Essence of personalization teaching and instruction

III. METHOD AND ALGORITHM

An important component of the distance teaching platform of individualized teaching is the personalized examination system. Distance teaching has spread through all the universities of the whole country, and it has become a branch of the higher learning institutions education. Students in the strange land can listen to and watch relevant teaching course through some communication means of the Distance teaching, purpose to achieve study. But the present Distance teaching development situation can only train students in the way of spreading in one direction, and it can't response students' states of learning and can't test students' results of learning in time. Furthermore, the teachers and the students can't communicate each other promptly, thus hindered the further development of distance learning. Regardless of traditional education or the modern education including Distance teaching, it is the most important links in education to study, test and appraise. Assess and test and have an examination play a very important role in education. How to find out about students' results of learning, how to grasp student's true operative skill, how to weigh students' level reached under education and training ,is it realized through one assessment and test or examination way to need. Have an examination in the past, students use paper basically, teacher set theme, decide one, tedious and very apt to produce the mistake. Some operate type examinations, students answer the questions in a test paper inconvenient and losing contact with reality also, so it causes a lot of examinations to judge by accident resulting in grading unjustly, can't reflect students' true level objectively. Computer assesses and test is that it uses computer to replace traditional paper examination mode to have an examination, so it has just, safe, high-efficient characteristics.

Utilize network environment is it set theme group roll automatically to offer, humanized question sheet, full-automatic examination platform of marking examination papers lightening every working intensity of examination links greatly. Improved the objectivity , fairness and security of the examination greatly, dispelled both region difference and space-time difference at the same time , it accords to the nowadays open development model of education, and it is one of the most potential field of modern education. The progress of science and technology makes the society accelerate the change of demand for talent. This kind of change has promoted the reform of education. The proposition of quality-oriented education makes the concept and mode of assessing and testing and examination present new development. With the pluralism of educational form and content, assessing and testing and

having an examination present diversified forms correspondingly.

Each index can score from reviewer's subjective scoring method after obtaining. The data to be using equation (1) is normalized.

$$\bar{x}_i = \frac{x_i - b_i}{a_i - b_i} \quad (1)$$

Where X_i and X_j respectively, the i -th index and the actual value Standard value; a_i , b_i are the maximum, minimum, the i -th index. Known evaluation indexes m , n hidden layer nodes depending Problems and experimental data to determine, you can also experience the value of the formula (2) the decision.

$$n = \log_2 m \quad (2)$$

Hidden node output is calculated as follows:

$$h_j = f \left(\sum_{i=1}^m w_{ij} x_i - \theta_j \right) \quad (3)$$

Where θ_j is defined as the threshold value for hidden node.

The output of the output node is calculated as follows:

$$f \left(\sum_{i=1}^m w_{ij} x_i - \theta_j \right) = f \left(f \left(\theta_j \right) \right) \quad (4)$$

Where θ is an output node threshold.

Equation (3) and Equation (4) in the transfer function is generally expressed as (0,1) interval of S-type function:

$$f \left(\sum_{i=1}^m w_{ij} x_i - \theta_j \right) = f \left(f \left(\theta_j \right) \right) \quad (5)$$

Each index can score from reviewer's subjective scoring method after obtaining. The data to be using equation (1) is normalized.

$$\bar{x}_i = \frac{x_i - b_i}{a_i - b_i} \quad (6)$$

h_j is gaussian basis function, the basic equation of RBF neural network function is shown below :

$$h_j = \exp \left(- \frac{\|X - C_j\|^2}{2b_j^2} \right), \quad j = 1, 2, \dots, m \quad (7)$$

The output of the network is given as:

$$y_m(k) = wh = w_1 h_1 + w_2 h_2 + \dots + w_m h_m \quad (8)$$

Assuming the ideal output is $y(k)$, the performance index function is:

$$E(k) = \frac{1}{2} (y(k) - y_m(k))^2 \quad (9)$$

Based on the gradient descent method, node center and base width parameter are:

$$w_j(k) = w_j(k-1) + \eta (y(k) - y_m(k)) h_j + \alpha (w_j(k-1) - w_j(k-2)) \quad (10)$$

$$\Delta b_j = (y(k) - y_m(k)) w_j h_j \left(\frac{\|X - C_j\|^2}{b_j^3} \right) \quad (11)$$

$$b_j(k) = b_j(k-1) + \eta \Delta b_j + \alpha (b_j(k-1) - b_j(k-2)) \quad (12)$$

$$\Delta c_{j,i} = (y(k) - y_m(k)) w_j \frac{x_j - c_{j,i}}{b_j^2} \quad (13)$$

$$c_{ij}(k) = c_{ij}(k-1) + \eta \Delta c_{ij} + \alpha (c_{ij}(k-1) - c_{ij}(k-2)) \quad (14)$$

Jacobian matrix algorithm is shown as follows:

$$\frac{\partial y(k)}{\partial u(k)} \approx \frac{\partial y_m(k)}{\partial u(k)} = \sum_{j=1}^m w_j h_j \frac{c_{1j} - x_1}{b_j^2} \quad (15)$$

Where $x_1 = u(k)$.

The goal output vector quantity is defined as:

$$E(net) = \sum_{k=1}^p \sum_{i=1}^m (t_i^k - y_i^k)^2 \quad (16)$$

The following Sigmoid transmit function is used.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (17)$$

The fitness is defined as follows.

$$f(net_i, t) = C - E(net_i, t) \quad (18)$$

We use roulette wheel selection operation. Set colony's size as N, the fitness function of the individual i is f_i , its selection probability is p_{si} .

$$p_{si} = \frac{f_i}{\sum_{i=1}^N f_i} \quad (19)$$

The fitness in each evolving population is defined as equation (20):

$$p_c = \begin{cases} p_{c1} - \frac{(p_{c1} - 0.6)(f_{\max} - f)}{f_{\max} - f} & f > \bar{f} \\ p_{c1}, f < \bar{f} \end{cases} \quad (20)$$

IV. EXPERIMENT RESULT

With the launching of modern distance education, E-learning learning format is gradually accepted by people, and we have gradually entered the era of network-oriented education. Around us there are a large number of digital educational resources: lesson plans, handouts, images, audio, video and so on, which have greatly enriched people's life and improve the quality of learning. The development of internet technology has promoted the large-scale and rapid spread of digitized educational resources. Modern distance education is inseparable from learning platform. If users use a platform together at the same time, there will be a greater amount of data volume, higher bandwidth and real-time audio and video resources. All of these make a higher demand of a server and network per-

formance for strong distance learning platform. A large amount of data storage and the simultaneous access of the same data by a large number of users will result in many difficulties in the server's storage capacity, I/O bandwidth, processing load and network transmission, which will greatly affect the distance education learning and its development. If the teaching and learning platform and resources storage are set up in all schools, in the practical implementation many problems will be created because of student number, economic development, and network conditions. Particularly in the less developed western regions, there will be difficulties even greater difficulties in establishing and applying a complete platform for distance education and teaching. In recent years, the time series data mining technology has made great progress. With the network technology, sensors and other data sensing technology continues to develop, on the one hand, the structure of time series data is becoming more complex, the massive of data is also increasing, on the other hand, there is a growing need for people to found more useful information and knowledge from these more complex data. At the same time, computing power continues strong also make the study of these more complex structures and characteristics of the data out of discovering which information and knowledge possible. For such as time series flow, uncertainty time series, multi time series and more complex structures such as time series data mining techniques, the conventional time-series data mining technology is relatively for most of the simple structure time series data mining.

The traditional distance education platform design does not follow certain technology standards, existing the "technology for technology" phenomenon. In addition, most existing distance education resources displays in writing and planar way, what it realized is a copy to original teaching materials. In this case, the learner obtains knowledge cannot get away from the mechanical type studying. The emergence of VRML technology brings the distance education a new teaching means because of its little data to transfer and with vivid 3D pictures, providing a new technology support for deigning modern distance education platform. Fig. 5 is a kind of model of personalization teaching and instruction of modern distance education, it contains knowledge base, individual database, WDM and web server four modules.

Different professional or different field need different examination mode and even the same specialty and the same course need different examination type and examination way too, which has offered wide development space for computer examination. With the development of information technology, the popularization of computer network and intellectual technology of calculation are becoming riper, furthermore, all kinds of examinations will adapt to this kind of change too. Achieving a kind of examination platform intelligently based on network multi-functional will become first-selected targets and development trends in all kinds of examinations in the future. Combining perfect network application with the traditional test and training mode, advanced software engineering, but that becomes that Adopting reeducation platform constructed by latest .net Framework 1.1, professional large-scale database MS SQL Server 2000, XML technology of American Microsoft corp., We have realized a professional distance teaching way based on web way. The whole platform is divided into three sub platforms of students LTEA platform, system management platform and teaching management platform.

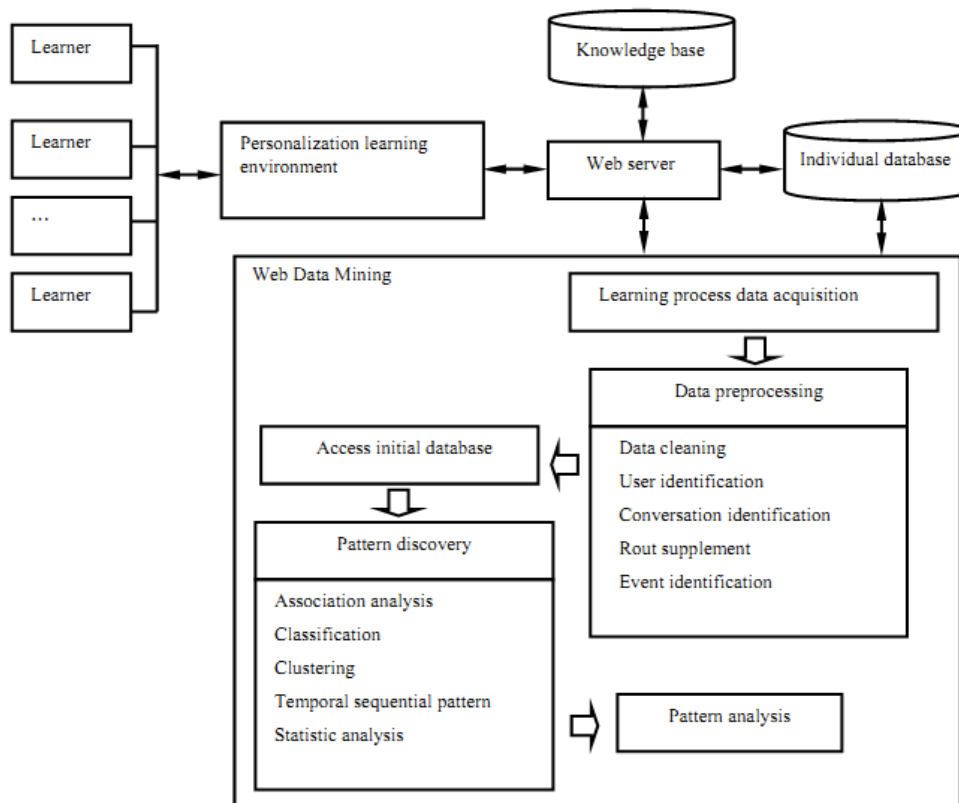


Figure 5. The framework of personalization teaching and instruction

1) Knowledge base: The distance education system provides abundant learning resources to learners, which include courseware base, job library, and answer-question system and case database etc. These resources are organized according to knowledge points and are collectively called as “Knowledge base”. The learner takes advantage of learning resource in knowledge base by invoking web server.

2) Individual database: Individual database is used to save static information and dynamic information of different learners. Static information includes name, ID, password, email address etc., it is not often change and offers the log-in inlet. Dynamic information consists of knowledge structure, learning history, learning schedule etc., it changes with the learning process.

3) WDM module: This module achieves dynamic information of different learners from web server. By data preprocessing, pattern discovery and pattern analysis three process, it perfects continuously individual database and returns appropriate advices and strategy to different learners.

4) Web server: Web server is the intermediate module of realizing personalization teaching and instruction of modern distance education. The learner gains knowledge by accessing web server, and the result of WDM acts on web servers. Finally the distance education system chooses suitable learning content from knowledge base and put forward appropriate learning advices for the learner.

V. DISCUSSION

It takes a significant effect on all of the education structure. The government provides the support for the development of distance learning. Distance learning based on

Internet is one way of modern education. The basis platform is indispensable for teaching, FAQ, virtual experiment, real time talking, cooperating working, online test, affair manage. The system adopts the most popular B/S system structure, namely: Browser /web server/database server; B/S divides the system into three layers: The view layer, the business logic layer and the data access layer, and realizes to separate business logic, user interface and data access, which improves the maintains of application system. Distance teaching has spread through all the universities of the whole country, and it has become a branch of the higher learning institutions education. Students in the strange land can listen to and watch relevant teaching course through some communication means of the Distance teaching, purpose to achieve study. But the present Distance teaching development situation can only train students in the way of spreading in one direction, and it can't response students' states of learning and can't test students' results of learning in time. Furthermore, the teachers and the students can't communicate each other promptly, thus hindered the further development of distance learning. Regardless of traditional education or the modern education including Distance teaching, it is the most important links in education to study, test and appraise. Assess and test and have an examination play a very important role in education.

VI. CONCLUSION

Simulation results show that model has important enlightenment and pushing effect for promoting the individual service and improving teaching quality of modern distance education.

The purpose of this study is to develop a distance personalized teaching platform. The web data mining is used

for the construction of the system and by analyzing the character of web data mining (WDM) and the essence of personalization teaching and instruction, based on WDM, this paper build a kind of model of personalization teaching and instruction of modern distance education. The author adopted technology of self-adaptive multimedia stream for the system, synchronized and unsynchronized information communications and stream-based compression.

Moreover, the author implemented automatic transmissions of different language Char-codes and authenticity and protection of network resource. Finally, the author studied how to make development of Platform more transplantable and more modularized. Based on realistic requirement and needs, the author resolved several key problems in the development of Distance Education Platform. Distance learning comes with the computer multimedia and network communications being applying to education. It takes a significant effect on all of the education structure. The government provides the support for the development of distance learning. Distance learning based on Internet is one way of modern education. The basis platform is indispensable for teaching, FAQ, virtual experiment, real time talking, cooperating working, online test, affair manage. The author adopted the more and more popular XML technology and discussed how to develop plat for with the joint of XML and JSP, especially on how to resolve the data-sharing among hetero-structured DBMSs. Through research work and realization of above key technologies, the author finished the development of Intelligent Community platform and Distance Education platform.

REFERENCES

- [1] H. Jing, "The Study on the Impact of Data Storage from Accounting Information Processing Procedure," *International Journal of Database Theory and Application*, vol. 8, no.3, pp. 323-332, June 2015. <https://doi.org/10.14257/ijdt.2015.8.3.28>
- [2] H. Ye, "Discussion about the personalized distance education service system," *Journal of Jiamusi Education Institute*, 2013.
- [3] J. W. Wang, X. Wang, L. B. Zhao, et al., "Applied Research of the Decision Tree Method in Personalized Distance Education System," *Advanced Materials Research*, pp.881-884, 2013. <https://doi.org/10.4028/www.scientific.net/amr.756-759.881>
- [4] X. Yu, and M. Ma, "On Design and Implementation of Personalized Distance Education System," *Journal of Ningbo Radio & Tv University*, 2007.
- [5] X. Wang, J.W. Wang, Y. F. Zhong, et al., "Research of data mining technology application in personalized distance education platform," *Journal of Heilongjiang Institute of Technology*, 2010.
- [6] C. Ji, and M. Fan, "The application of web data mining in personalized modern distance education," *Energy Procedia*, no.13, pp.714-720, 2011.
- [7] S. Yao, "Research on Web-based Autonomous English Learning of Engineering Students," *International Journal of Emerging Technologies in Learning*, vol. 11, no.6, pp. 4-9, June 2016. <https://doi.org/10.3991/ijet.v11i06.5802>
- [8] Y. Li, J. Sun, and W. Qiang, "Application of Data Mining in Personalized Remote Distance Education Web System," *Open Cybernetics & Systemics Journal*, 2015, 9(1):1769-1775. <https://doi.org/10.2174/1874110X01509011769>
- [9] J. Yang, R. Xu, Z. Lv, et al., "Analysis of Camera Arrays Applicable to the Internet of Things," *Sensors*, vol. 16, no.3, March 2016. <https://doi.org/10.3390/s16030421>
- [10] P. Li, "Based on Data Mining Technology in Distance Education System Applied Research," *Applied Mechanics & Materials*, pp.2260-2264, 2014.
- [11] Z. Wang, "Why Are There "Digital Ruins" in the Network Education Resources—The Difficulty of China's Network Education Resources Construction," *Modern Distance Education Research*, 2015.

AUTHOR

Lanzhong Wang is with Shandong University, China (75152966@qq.com).

Submitted 09 September 2016. Published as resubmitted by the author 23 October 2016.