The application of web data mining in personalized learning system

Chaoyang Ji

College of Information Technology, Xuchang University, No.88, Bayi Road, Xuchang, China

_____

ABSTRACT

The paper analyzed the present situation of online learning and then described the necessity that carried out personalized online learning services firstly. It constructed a personalized online learning model based on web data mining and described the basic design idea and functions in detail. For the online learning service organization, this model will provide a certain theoretical basic and reference value to build a personalized online learning system.

Key words: Online learning; Web data mining; Personalized service; learning platform

_____

INTRODUCTION

The computer technology and network technology have been widely used in network education. Online learning is a new way of education which is based on information technology platform. It is a supplement to traditional education model and an effective way for people to use social advantage educational resources effectively. Online learning as a learning tool can be used in higher education, vocational education and adult education, and it is also a good means of providing lifelong education. The intelligent network education model not only provides a convenient way of learning and a wide selection for learners, but also provides the possibility for education institutions to understand learners' needs information and behavioral characteristics in a more in-depth way.

But the current network teaching system or learning platform is to provide fixed capabilities like browse, download, upload educational resources. It has such problems, for instance, teaching model is too unitary, dynamic interaction is not strong, lack effective guidance of individual learning problems. The current online learning is mainly centering on "information resources", dose not reflect the idea that centers on "learner". Each learner in the learning process is to see the same teaching resources, teaching platform does not have personalized and intelligent features [1]. There are great differences between learners in online education, and their learning objectives, learning abilities vary. Therefore, network education should be based on individual differences and needs of the educated person, provide learning resources relatively, which inevitably requires online learning system platform to provide individualized teaching services for learners. To achieve this goal, putting to use web data mining technology to build a personalized online learning service system can be a good choice to solve these problems.

WEB DATA MINING
1. The Concept of Web Data Mining
Web data mining is the development of data mining technology, involving Web technology, data mining, computer technology, information science and other fields. Web data mining is to discover patterns from a large number of web document collections, which are implicational, unknown, having potential value and nontrivial. Web data mining deals with objects, including static pages, web database, web structure, user records, etc.

_____

2. The classification of web data mining
According to the studying target of web data mining, web mining can be divided into three categories: web content mining, web structure mining and web usage mining [2].

3. The Process of web data mining
The process of web data mining includes four stages: data collection, data pre-processing, pattern discovery, pattern analysis [4]. Each stage of the web data mining has different output, and they are the initial data, data, patterns and knowledge.

Data collection: This stage is used to collect the valuable data for learner service and management. The Web lacks accurate data collection mechanism for usage data. Therefore, collection techniques of usage data are the important part of data mining research. The usage data on the Web is very rich, with multiple gathering places, including the client, HTTP proxy, Web server, or even the underlying network access. The web logs are the most common usage data. A variety of data are collected accurately, timely and comprehensively, and this is the important prerequisite and basis for web data mining.

Data pre-processing: The task is that the useless information is removed from the acquired original web resources and information is arranged necessarily. The information is converted into data which is available for mining methods, when necessary, storing it to a specific database. For example, the text in web pages, pictures and other files are converted into available form of data mining and little valuable information are filtered.

User pattern discovery: There are several ways of pattern discovery. In the field of Web usage mining they are mainly statistical analysis, frequent pattern mining, clustering, classification and sequence pattern. At this stage the appropriate data mining method is selected and mining knowledge from the arranged web data.

Pattern analysis and application: Through analyzing the results of web mining, the useful pattern is analyzed and filtered out from many ordinary patterns. The pattern will be applied to online education personalized information service finally.

PERSONALIZED ONLINE LEARNING SERVICES
Currently, the dynamic interactive function of most online education system is not strong, learner can only passively brow learning resources, can not actually participate in the system for discovery learning and creative learning. The intelligence of system is lower, can not provide suitable and intelligent interface according to the circumstances of learners, and can not mobilize the learners' interest in learning. The system also can not provide appropriate guidance according to the learners' learning ability and learning situation to choice suitable teaching measures for learners. The Internet test system flexibility is poor which is mostly fixed on a web page putting some fixed test questions, which can't based on learners' requirements and learning progress, couldn't choice paper topics and change topics automatically, can't insight the lack of knowledge in other related vulnerabilities according to the learners' poor knowledge. The system lacks effective guidance for learners, they browse aimlessly in network, not only waste time but can't achieve learning goal [5]. By providing personalized services, the system develops learning programs according to the specific circumstances of the learners, so as to enhance the learning effectiveness and achieve the intended purpose.

1. The content of Personalized Service
Personalized service is also known as customized services. Personalized service is that the online education organization use modern network technology, artificial intelligence and expert systems, etc. and take the initiative to obtain the specific information needs of individual learner and the common learning resource needs of specific learner groups. This is a comprehensive information service mechanism that can automatically retrieve information of network learning resource and push them to give learners according to online network learners' requirements [6]. It aims to satisfy learners' individual requirements through personalized service.

Through a personalized information service system, it can provide learners with information content and system functions satisfying their individual needs. According to the learners' clear or vague information resource needs, the system provide them with targeted information services, including information consultation, information retrieval, information gathering, information integration and information push, etc. In the information service process, the corresponding modules of system research and analysis the historical data of learners behaviors, habits, preferences, specific information needs and learners own information [7]. The system obtains specific information characteristics of different learners, and then loaded them into personalized online education information service system. The system actively provide learners with targeted, active, and demanded special services, and then realize the purpose that the online education system provides personalized information service to the learners.

_____

2. The classification of personalized service
Personalized services mainly include such three aspects as personalized service time, personalized service mode, personalized service content.

Individuation of service time: Under the network environment, online education web site services have broken through the limitations of time and space. According to learners own wishes, they can receive services of the online education platform anytime and anywhere.

Individuation of service mode: Personalized service mode reflects the level of service. According to different the categories and individual characteristics of learners, online education platform can properly provide services and a variety of modes of services. Thus the traditional online education services are changed and the new service modes are established, according to learners' demands.

Individuation of service content: Online education platform provides a variety of network learning resources. Due to differences in learner's demand, the demand for information resource and services also vary. Different types of services are provided for different learning level and learners can select required services according to their own needs.

THE ROLE OF WEB MINING IN THE PERSONALIZED ONLINE LEARNING
The system extracts useful and potential pattern which learners interest in from web documents and web activities based on web data mining in the online education. It will provide learners as the basis for individualized teaching services and be used to optimize the online education site structure. During the personalized services, the role of web data mining is mainly manifested in the following areas:

Improve the relevance of teaching: Through analyzing and mining the access logs of online education web server, we can deeply understand the learners, analyze their browsing habits, personal preferences, value orientation and other characteristics, fully grasp the needs of different learner, so as to carry out targeted online teaching activities and improve the learning effect of learners.

Forecast learning interest: The system collect learners' information on the behavior firstly. According to historical data of learners and access records, we can predict the research of learners, assess the change of interest in learning, and mine the potential learning demand [8]. So as to conduct targeted marketing campaigns and provide valuable and interested personalized information and services to consumers.

System optimization: This includes two aspects. The first is to improve the operating performance of the system, for example, deal with Web cache and access balance, etc. The second is to optimize the design of Web sites. Through mining learner's browsing path, the system can find the logical relationship of pages. The designers of site can modify and design site structure and appearance according to the needs of learners, and so making the structure of site more reasonable, more user-friendly browsing web pages and according with learning habits [9].

Design of Personalized needs: Personalized interface are designed to meet the information requirements of different learners, thus it can reflect the personalized information service of modern online education web site.

DESIGN OF ONLINE LEARNING MODEL
1. The Main Idea of Model Design
When learner logins the system in the first time, learners will be asked to fill out some registration information. This information will be collected through the information collection module and sent to the learner database to establish a record. The learner having been registered in the system login the online education service system after authentication. The learner makes a request for information resource through the learning interface module, and then the information collection module begin collecting information and request of the learner and tracking the behavior of learner. The collected information includes the URLs which are clicked by learner, keyword, the situation of form filled by the learner, the access type of information, the process and the results of doing homework and examinations. These data are translated to the information dispatch module, and this module gives a scheduling order to the resource database and the tool database according to the requests of learner and the personalized information. The learner is provided with the tools and the learning resource which are most necessary to him. At the same time, the learner can submit feedback to the system, according to the results of satisfaction with the information [10]. The management and teacher module are responsible for management and maintenance of the resource database, learner database and interest database. The personalized online education model is shown in Figure 1.
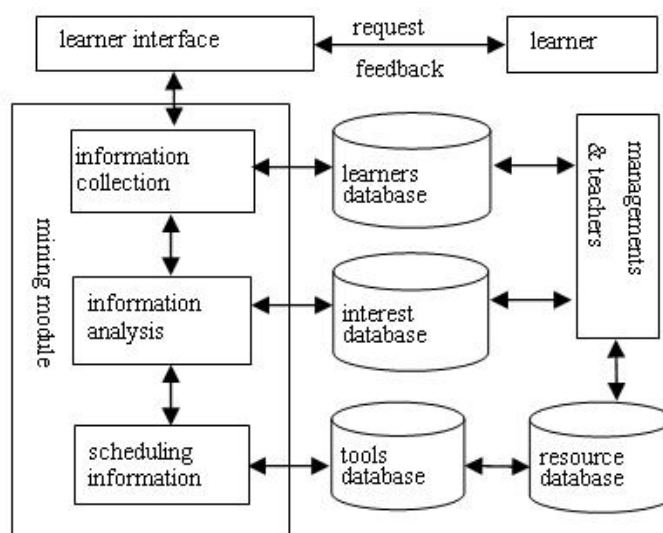
_____

Fig.1 The structure of personalized service model

2. The main function modules of model
The model consists of personalized mining module, management and teacher module, learner database, interest database, resource database, tools database and personalized learning interface module.

Learner database: It is mainly used to store learner personal information which is submitted for registration and learners can update the registration information real-timely. When a learner login the system, it can obtain registration information from the information database and thus authenticate the identity of learner. This database can also be expanded at any time according to actual needs.

Resource database: It includes question and answer database, exercise database, test database and other resources related to teaching.

Tool database: It contains question and answer tools, working tools, test tools, communication tools and so on.

Personalized mining module: It is the core of the online education system, and is also the key reflecting the personalization. This module includes information gathering, information analysis, and information scheduling three sub modules. Information gathering module uses the key words learning of intelligent agent and dictionary learning to collect learners' current information automatically. This information include personal circumstances, interested topic, frequently search keywords and the result of test, and these are translated to the module of information analysis. The analysis module gets these data and makes personalized analysis using data mining technologies, and then sends commands to the information scheduling module. The scheduling module uses learning tool access learning resource database according to commands and returns corresponding information to learner interface. At the same time, the information is returned to analysis module, the analysis results are stored in learner interest database. Thus, the learner makes a request and receives corresponding personalized content. With the changes in the learning process, interest database is also continuously updated, preserves and reflects the learner's personalized information in time.

Learner interface module: This module is the interactive interface of learners and the system and it can accept learner registry, requesting information and feedback information, as well as the system can study and record the learner's interests and habits during the process. The module also has the function of outputting the final result to corresponding learners and obtaining the content of local learner information database.

The module of push and feedback information: Using intelligent push technology, it pushes the learner needed information to learner's computer, E-mail and so on. In order to improve service quality, the module may require learners to provide feedback to the information provided (For example: whether to continue providing such information, cancel customization, etc.), and then the module analyses the learner's feedback again to adjust and improve the learner interest database.

The learner interest database: It includes basic information of learners, learner interests, hobbies, field of study, knowledge structure, and behaviors and so on.

_____

CONCLUSION

Web data mining is a new technology that is the combination of Web and data mining, and it is a research hotspot of domestic and foreign scholars. There is a considerable difference of required learning information between learners in online education, and so the personalized information service has become a new service mode. Through personalized service, it can shorten the distance between learners and the education service organization and provide learners with more targeted services to improve learners' service quality. The paper first analyzes the status of online education, and then put forward a personalized online education system model based on web data mining technology, and last introduces the basic designed idea of the model and the basic functions and implementation techniques of each module. This will play a positive role in promoting the development of personalized online learning.

REFERENCES

[1] Ruixian Liu, Ling Gao. *Computer Applications and Software*. **2009**.26(10):143-145.
[2] Bunchner A G, Mulvenna M D. *Sigmod Record*, **2006**, 27 (4):145-156.
[3] Xue Chan. *Lantai World* **2008**, 24:73-74.
[4] Jiawei Han, Micheline Kamber. Data Mining: Concepts and Techniques. China Machine Press, **2006**.
[5] Tong Liu. *Science   & Technology   information*. **2008**.21:62-63.
[6] Guangxing Song, Chunfang   Cao. *Forecasting*.**2009**, 28(5):55-58.
[7] Haiyang Wang, Yan Zhang, Jing Xu. *Microcomputer Information*. **2010**, 26(2-3):114-115.
[8] Lanjuan Yang. *Computer Knowledge and Technology*. **2010**, 06(7):1548-1550.
[9] Chaoyang Ji, Yanjie Zhao. *Science and Technology Management Research*. **2009**, 06:278-279.
[10] Lili Wu, Bo Zhang,Tao LI. Analysis and Implementation of Distance Education Individuation Based on Web Data Mini. *Microcomputer Applications*. **2010**, 26(2):59-61.