Application of Artificial Intelligence System in Libraries through Data Mining and Content Filtering Methods

Zhaoxia Tian
Fujian Chuanzheng Communications College, FuZhou 350007, China

*Corresponding author: tianzhaoxia@fjcpc.edu.cn

Abstract. With the rapid advancement of artificial intelligence theory, this paper adopts a multi-intelligent agent collaboration method and derives through data mining. In combination with content filtering methods and intelligent agent learning optimization, it improves the high performance by using a personalized information service system architecture. The performance of the library system of vocational colleges. According to the difference of readers' interest, it matches the results of traditional document retrieval, effectively filtering out readers' demand information, reducing the time for readers to search for required information, improving reader retrieval efficiency, realizing information push of similar users, and realizing "information looking for people".

Keywords: Artificial intelligence, higher vocational colleges, library system, intelligent agent.

1. Introduction
Artificial intelligence is a new technology that has emerged in recent years. In China, artificial intelligence technology is also considered to be one of the top ten hot issues affecting the development of libraries, such as "disruptive technology of libraries in the next ten years" and so on. We know that service is the purpose of all the work of the library [1]. With the help of advanced artificial intelligence technology, library services will be icing on the cake, even more powerful, and will be more conducive to the Olympic spirit of "higher, faster, stronger". Increase the breadth, breadth and depth of library services, and increase readers' satisfaction with library services. The library itself is a pioneer in the application of new technologies, and it should actively pay attention to the development of artificial intelligence technology and actively practice artificial intelligence technology.

The library of higher vocational colleges is an important position to provide resource guarantee, knowledge service and decision-making reference for "building an applied technology university with Chinese characteristics" and "boosting local economic development". How to face the management mechanism brought by the new technology and new environment the disruptive changes in the service model and the acceleration of the transformation of its resource structure, space environment, service model, business focus and other aspects are topics that it needs to conduct in-depth research and dedicated exploration based on actual conditions.
2. Features of Library Smart Service

"Smart library is a new type of knowledge service system based on systematic literature resources, intelligent knowledge services and intelligent guarantee support. It is not an institution, but a form of service." What is a library smart service? Based on the opinions of many experts in the industry, the author believes that the so-called library smart service should be the integration and interaction of smart technology, smart librarians, and library business and management systems, and provide users with a new concept, new goal and new service model of the company [2]. As the core of a smart library, smart services should have the following characteristics:

2.1. Ubiquitous service
The application of technical means such as Beacon, RFID and big data enables the library to provide users with all-round and integrated resources and services without being restricted by time and space, to achieve a high degree of integration of resources and services, and to realize books and books. The interconnection between people, people, and libraries makes the service objects, service areas, and service content of smart libraries more extensive.

2.2. Integration of physical and virtual services
On the one hand, the use of virtual reality and augmented reality technologies can realize human-computer interaction and human-computer perception between the user’s visual, auditory, and tactile simulated space scenes and the real library scenes, such as wearable technology and 3D navigation, Virtual bookshelves, virtual flipping books, etc., through the establishment of a library simulation system, so that users have the same immersive feeling in the virtual space as the physical library [3]. On the other hand, the user’s good experience of the virtual library can also be transformed into the effective use of offline physical libraries, logistics and distribution, and the realization of organic integration with various colourful user activities, transforming the information flow generated online into online the offline service flow form, through online and offline interaction and integration, to the greatest extent meet the needs of users.

2.3. Personalized, efficient, precise and diversified service content
The application of artificial intelligence technology in smart libraries, with the help of big data analysis, data mining and other methods, analyses user preferences and potential information needs, and can provide users with more personalized recommendation services and reference consulting services; RFID, global positioning technology. The use of such as self-service lending, smart seat reservation, mobile library and other functions can be realized, enabling users to accurately and efficiently find the resources and services they need anytime and anywhere, saving users time, and improving user satisfaction and library resources.

3. Personalized information service system model based on Agent library
Web mining technology based on intelligent agent is an effective way commonly used in library personalized information service. Combining the user's actual information application environment, mining user interest information, establishing user interest model, using library information service model can help users search for target resources and push user interest book information. The library’s personalized information service model is constructed on the basis of the traditional library management system (OPAC) platform. Its content can include information service system messages, subject librarian dialogue windows, collection resources, circulation and borrowing, user interest learning, Database management and other functions [4]. Figure 1 shows the model architecture of a personalized information service system for higher vocational libraries based on multi-intelligent agents.
3.1. User interface module
The user interface module provides an operable interface for users to interact with the system. The interface provides a variety of information service login methods, including system login, mobile reading portal login, QQ login, and WeChat login. Through the user interface, resource information can be displayed to users in multiple ways [5]. The user-friendly interface is convenient for users to query the library's huge collection resources, including physical resources and virtual resources. The information system login interface is divided into two parts: information input and output. The information input interface mainly includes the keyword input column, the setting of search conditions, and the user's feedback on the search result information.

3.2. User's personalized interest learning and information query recommendation module
User interest learning and information query recommendation modules mainly include user personalized interest information database, user learning interests, reading habits, information filtering, information retrieval, etc. (Figure 2). According to the composition diagram of the user's interest learning, reading habits and information recommendation module shown in Figure 2, the content of each component of the module includes readers' personalized characteristics, learning interests and habits, information search and filtering, etc.
3.3. **Personalized application function module**

The application function module is the general term for the realization of the main functions of the library information service platform. The main application function modules include the cataloguing of book information, entry, book search, system management, etc. (Figure 3).

![Figure 3. Functional structure of library personalized information service application.](image)

In the application function of the library information service system shown in Figure 3, the circulation module realizes the interaction between the reader and the system. Readers submit required information to the system and evaluate the results of the information provided by the library. The user uses the public query module to query the required book information. By selecting the query method, the user sets the query restriction conditions, which can improve the accuracy of the query. The public query module provides a personalized retrieval function [6]. The system performs fuzzy query of information in the keyword-related fields according to the keywords entered by the user, and obtains the results of user interest information through the mutual cooperation between intelligent agents. The digital information module can store multiple types of information resources such as audio, video, Flash, etc., to provide users with multimedia demand information. The key of the circulation statistics module is to count the interaction between users and the system, to count the browsing records of users in the system and the records of user behaviours, and to provide a data basis for the system to update user interests.

3.4. **Information Resource Management Module**

The main function of the information resource management module is to realize the management of system metadata and information resources. Among them, the management of metadata is to classify, organize and store the metadata of information resources; the management of information resources is to classify and organize existing data resources in the database, as shown in Figure 4.

![Figure 4. The composition of information resource management.](image)
4. Digital representation of user search interest and visit behaviour

Suppose that a certain keyword is used to search in a database, and all the searched information constitutes a search queue. The content of the search queue remains unchanged during the search process. It is numbered sequentially, assuming that it is numbered from 1 to N. In most cases, users will only access a small part of the search queue. The basis for selecting this small part of information is mainly based on manual identification. The general method is based on the title or opening the trial reading, and then making a choice. For a long queue, some really useful information may be missed, or a lot of effort may be spent trying to read a lot of unnecessary information. A certain number of interest libraries need to be saved in the system in advance, which can be generated through manual settings or historical collection records. The interest database is composed of a group of interest records. An interest record represents information worth reading for readers of a particular interest. It is a subset of the search result queue and records the search information related to the interest. The current mainstream interest expression method is based on vector representation. The expression of an interest in this article is as follows:

\[ \text{Interest} = (I_1, I_2, \ldots, I_N) \]  

Among them, \( I_i \) corresponds to the i piece of information, and its value is 0 or 1. A value of 0 means that the i piece of information is not in the interest, and 1 means that the i piece of information is in the interest. Here, the information of different information in the same interest is ignored. The degree of importance difference, that is, it is assumed that all information with 1 is equally important in the interest. From an interest record, what readers of the interest need most is what information in the search result queue. Multiple interests can form an interest matrix, which represents the interest library, expressed as follows:

\[ \begin{pmatrix}
I_{1} \\
I_{2} \\
\vdots \\
I_{K}
\end{pmatrix}
= 
\begin{pmatrix}
I_{11}, I_{12}, \ldots, I_{1N} \\
I_{21}, I_{22}, \ldots, I_{2N} \\
\vdots \\
I_{K1}, I_{K2}, \ldots, I_{KN}
\end{pmatrix} \]

A user's visit in the search result queue must be a visit to certain information, which can be recorded as an action. When the user has made one or more visits, there must be several actions, abbreviated as a. These several visits constitute one visit and are recorded as Behaviour. By tracking an ongoing access behaviour, we can get a real-time access behaviour record every time. The user's access behaviour record can be expressed as a one-dimensional digital matrix as shown below:

\[ \text{Behavior} = (a_1, a_2, \ldots, a_N), \sum_{i=1}^{N} a_i = 1 \]  

Behaviour changes dynamically during the user's access process, and is recalculated every time an operation occurs. It records the user's latest access information on the search results. Now we can perform a fuzzy matching operation between a Behaviour and an Interest, the purpose is to determine to what extent the detected access behaviour belongs to the Interest. The method of judgment uses fuzzy AND operation, and the result of the operation is between 0 and 1, indicating the degree to which the behaviour of the visit is subordinate to the interest, denoted as \( d \).
\[
d = \text{Interest}\_\text{Behavior} = (I_1, I_2, \ldots, I_N)(a_1, a_2, a_3, \ldots, a_N)^T = \sum_{i=1}^{N} (a_i \times I_i)
\]  \hspace{1cm} (4)

D represents the extent to which the currently tracked behaviour is similar to the standard Interest, and its value is between 0 and 1. After getting d, we need to compare with the threshold to determine whether the user has the interest, and then give priority to recommending the follow-up information in the Interest record. For example, if the threshold value is set to 0.8, when the value of d reaches or exceeds 0.8, we believe that the visit behaviour belongs to the Interest of Interest, and the follow-up information in the Interest of Interest shall be presented to the user in a timely manner.

5. Conclusion

Smart service is a new direction for the development of higher vocational libraries, and the construction of "double first-class" has become a new opportunity for the transformation and upgrading of higher vocational libraries. With the rapid development of information technologies such as the Internet of Things, big data, and cloud computing, the development prospects of smart libraries have become broader, and smart service models have become the general trend. Some achievements have been made in the development of smart services in higher vocational libraries, but there are still many shortcomings. University libraries should firm up the concept of smart service, and constantly open up new prospects for smart service through the use of various new types of smart technology.

References

[1] Lu, L., Meng, X., Mao, Z., & Karniadakis, G. E. DeepXDE: A deep learning library for solving differential equations. SIAM Review, 63 (1) (2021) 208-228.

[2] Yu, S. Y., Chhetri, S. R., Canedo, A., Goyal, P., & Al Faruque, M. A. Pykg2vec: A python library for knowledge graph embedding. Journal of Machine Learning Research, 22 (16) (2021) 1-6.

[3] Bancerek, G., Byliński, C., Grabowski, A., Korniłowicz, A., Matuszewski, R., Naumowicz, A., & Pąk, K. The role of the Mizar Mathematical Library for interactive proof development in Mizar. Journal of Automated Reasoning, 61 (1) (2018) 9-32.

[4] Umar, A. A., Ahmad, M., & Batcha, M. S. Library and Culture: A Scientometric Analysis and Visualization of Research Trends. Journal of Cultural and Social Anthropology, 2 (2) (2020) 1-08.

[5] Lewis, S. J., Gandomkar, Z., & Brennan, P. C. Artificial Intelligence in medical imaging practice: looking to the future. Journal of Medical radiation sciences, 66 (4) (2019) 292-295.

[6] Geis, J. R., Brady, A. P., Wu, C. C., Spencer, J., Ranschaert, E., Jaremko, J. L., ... & Kohli, M. Ethics of artificial intelligence in radiology: summary of the joint European and North American multisociety statement. Canadian Association of Radiologists Journal, 70 (4) (2019) 329-334.