University Library Service Based on Big Data

Chenyou Chen

JiLin Engineering Normal University, Changchun, China
50156506@qq.com

ABSTRACT. The rapid development of big data technology has brought huge challenges to the service innovation of university libraries, but at the same time it has also brought strong technical support, which plays a huge role in improving the service quality of university libraries. The big data analysis technology is a product of the information age widely used in the service system of university libraries, and it effectively improves the service mode and data information management method of university libraries. This article clarifies the new ways of service for university libraries in the era of big data as the basis of research, and innovates the current function transformation and service innovation, and fully researches and analyzes the new modes of university libraries. This thesis develops more personalized services and proposes effective innovation paths through the types of innovative services, creating a quality sharing platform, improving the skill level of librarians, and establishing a cloud service model for big data analysis. The experimental research results show that the big data analysis technology has raised the service form of the library to a new height, and the current situation of students who usually like to go to the library to read, the design can reflect the promotion style of the library, and let the current big data The digital reading method in the era can have a better space for development, and better play its analytical role to guide the whole school students to actively study in the library, and the sharing platform of the university library can effectively integrate big data. Analyze technology and reasonably distribute it with artificial intelligence technology, Internet technology and various front-end development software. The purpose is to enable the platform to get more student support and meet the diverse needs of students themselves, and establish a sharing Another advantage of the platform is that it can improve the current library management methods and improve the service system.

KEYWORDS: Big Data Analysis Technology, Library Management Service Innovation, Digital Library, Text Classification Algorithm

1.   Introduction

Information technology methods have been gradually introduced into library management and services, which has greatly promoted the development of academic libraries in my country, and the current self-service model has become the main service method of academic libraries[1]. However, big data technology is showing a rapid development trend, and it is still in an emerging field in my country. University libraries lack practical experience and successful cases that can be used for reference, resulting in the overall application level Lower [2]. Many colleges and universities still adhere to the traditional library service model. Although they have adopted information service methods to a certain extent, most of them are based on data foundations and have not really expanded from the field of structured data. They lack the deeper content of data information. Effective mining, the role of unstructured data is not fully highlighted [3]. Data literacy mainly includes the ability to perceive data, processing ability, collection ability, sensitivity, critical thinking and decision-making ability [4].

At this stage, digital libraries have widely appeared in colleges and universities, which not only facilitates the sorting of library information and data, but also provides convenience for readers. The relatively large scale of university libraries and the relatively large amount of books have led to challenging problems in the library management mechanism. Therefore, it is necessary to rationally use big data technology to establish and improve the data management system of university libraries, so that readers can find more information when searching for data. Efficient, improve the search effect and quality, and make full use of data [5]. At this stage, in the context of the new curriculum reform, the efficiency of the utilization of relevant academic materials has increased, and the frequency of students searching for relevant documents in the library has increased. University libraries need to use big data technology to personalize the recent reading habits and personalization of teachers and students The needs, types of books needed, the effective integration of resources and the types of related books are fully analyzed and researched, so as to effectively formulate appropriate service plans and give readers a better reading experience [6].

Generally speaking, the library of colleges and universities has a collection of up to tens of thousands of books. The work requirements of college librarians and staff under the previous library service mode are extremely high. A little carelessness may cause book placement errors, and readers cannot Finding the corresponding books according to the logo greatly reduces the service quality and efficiency of the library. Therefore, in the library reader service, it is necessary to combine the development of the times, make full use of the advantages of big data technology, and innovate the library reader service model [7]. When university libraries analyze information data, the main methods used are cluster analysis, information data mining, visual analysis, data fusion and integration, etc. [8]. These data analyses are not carried out at the same time, but after one step is completed, the next step can be carried out. It is impossible to achieve in-depth mining of information data, such as unstructured or semi-structured data, which leads to the inability to meet the needs

of every reader at the same time, such as To search for related books in the same time period, the above mode should also be adopted, so some readers often have to wait a long time to find the corresponding position of the book, which greatly reduces the service quality of university libraries [9, 10].

2. Method

2.1 Clustering Statistical Analysis

The statistics in the pedigree clustering method mainly include:

$$R^2 = 1 - P_G/T \tag{1}$$

Among them:

$$P_G = \sum_{K=1}^{G} \sum_{i \in GK} (X_i - \overline{X_K})^T (X_i - \overline{X_K})$$

$$T = \sum_{i=1}^{n} (X_i - \overline{X_K})^T (X_i - \overline{X_K}) \tag{2}$$

$$W_{PQ} = \frac{n_q n_p}{n_r} (X_p - \overline{X_q})^T (X_p - \overline{X_q})$$

Pseudo F statistics:

$$PSF = \frac{(T - P_G)}{(G-1)} / \frac{P_G}{(n-G)} \tag{3}$$

Pseudo t^2 statistics:

$$PST2 = \frac{W_{pq}}{(s_p + s_q)/(n_p + n_q - 2)} \tag{4}$$

2.2 Classification algorithm design

The distance between the n-dimensional space is measured by the cosine angle, and the calculation formula is as follows:

$$\text{Sim}(d_1, x) = \frac{w_1 \cdot w_x}{\sqrt{\sum_{j=1}^{n} w_{ij}^2 \times \sum_{j=1}^{n} w_{xj}^2}} \tag{5}$$

$$x^2(t, c) = \frac{N \times (AC - CB)^2}{(A+C)(B+C)(A+B)(C+D)} \tag{6}$$

The following two methods can be used to process the value of the feature item:

$$x_{avg(t)}^2 = \sum_{i=1}^{n} p(c_i) \, x^2(t, c_i)$$

$$x_{max(t)}^2 = \max_{i=1} \{x^2(t, c_i)\}$$

$$DF = \prod_{(x,y)\epsilon u} \sum_{xU} a\,(x,y)$$

$$\varphi_j(\mathrm{x}) = \exp\left\{-\frac{\|x-\mu_j\|^2}{2\sigma_j^2}\right\} \tag{7}$$

The first output is:

$$y_i = \sum_{i=1}^{k} w_{ij}\varphi_j(x) \tag{8}$$

$$h_{m+1} = create\_new\_learner(D_k^{m+1}) \tag{9}$$

$$\varepsilon_{m+1}^k = \sum_{i=1}^{|D_k|} w_i^m [h_{m+1}(x_i) \neq y_i] \tag{10}$$

$$\mathrm{E(x)} = \mathrm{argmax}_{y\epsilon Y} \sum_{t=1}^{size(E)} \left(\ln\frac{1}{\beta_t}\right)[h_t(x) = y] \tag{11}$$

2.3 Establish a Smart Library Service Model to Improve Efficiency

With the development of big data at this stage, if universities want to improve their book service system, they must accelerate the construction of library information technology infrastructure. If colleges and universities want to combine the current situation and needs of the service construction of their own libraries, they need to scientifically introduce some perception technologies such as big data technology, big data analysis technology, cloud computing, etc., to build a mobile Internet platform for the library, and strengthen the readership The collection and acquisition of user data and the perception of information are conducive to improving the perception level of the library's smart services. According to the reader's demand for electricity consumption, a personalized library reading experience service project can be formulated reasonably. University library management leaders should strengthen the scientific guidance to all employees, and continuously improve their active service awareness and smart service level. Effectively construct a big data security repository, organize and analyze user-related data information, facilitate users to automatically query demand information, and improve the timeliness of library smart services. Library intelligent services can not only reveal the hidden information needs of readers and users, but also help readers and users to dig deeper into the relevant knowledge hidden in this part of the information, so as to improve the level of knowledge and understanding of readers and users.

3. Experiment

3.1 Establishment of Statistical Classification Model

$$f(\mathrm{x}) = \mathrm{sign}(\sum_i^n w_i * x_i + b) \tag{12}$$

$$\text{sign}(x) = \begin{cases} +1 & x > 0 \\ -1 & x \leq 0 \end{cases}$$

$$L(w, b) = - \sum_{x_i \in M} y_i(w_i * x_i + b)$$

$$L = \frac{1}{T} \sum_{t=1}^{T} \sum_{-c \leq j \leq c, j \neq 0} logp(w_{t+j}|w_t)$$

$$p(wo|wI) = \prod_{j=1}^{L(w)-1} \sigma\left([[n(w, j+1)]] = ch(n(w,j)) \right) * v_{n(w,j)v_{wi}}^{T} \tag{13}$$

Data loss function model:

$$L = \sum_{i,j=1}^{v} f(x_{ij})(w^T w_j + b_i + b_j - \log(x_{ij})^2)$$

$$f(x) = \begin{cases} (x/x_{max})^a & x < x_{max} \\ 1 & otherwise \end{cases} \tag{14}$$

3.2 *Research on Collaborative Filtering Model in Book Recommendation System*

By introducing the collaborative filtering algorithm and applying it to the problems raised in this article, he can provide readers with the service requirements that users need.

$$r(X, Y) = \frac{\sum_k (X_k - \overline{X})(Y_k - \overline{Y})}{\sqrt{\sum_k (X_k - \overline{X})^2 \sum_k (Y_k - \overline{Y})^2}} \tag{15}$$

$$p(X_i) = \frac{\sum_k Y_i - r(X,Y)}{n} \tag{16}$$

Where Y is composed of all n persons who rated book i. The calculation formula is as follows:

$$p(X_i) = \overline{X} + \frac{\sum_k (X_k - \overline{X})^2 \cdot r(X,Y)}{\sum_k r(X,Y)} \tag{17}$$

4. Results

4.1 *Data Analysis*

From the data shown in Figure 1, we can see that the error after selecting 30 averages reaches a relatively stable state, and because the current university library books are much larger than university students, the effect of choosing the number of items is more important. Forcing the effect based on user algorithm is better. In order to improve the service quality of university libraries, under the background of this era of big data, universities need to vigorously develop the construction of information platforms, a total service system that combines teaching, daily life, and

learning services, precisely because of this The purpose is to improve the quality of informatization services and provide very convenient service efficiency for the management, teachers and students of colleges and universities
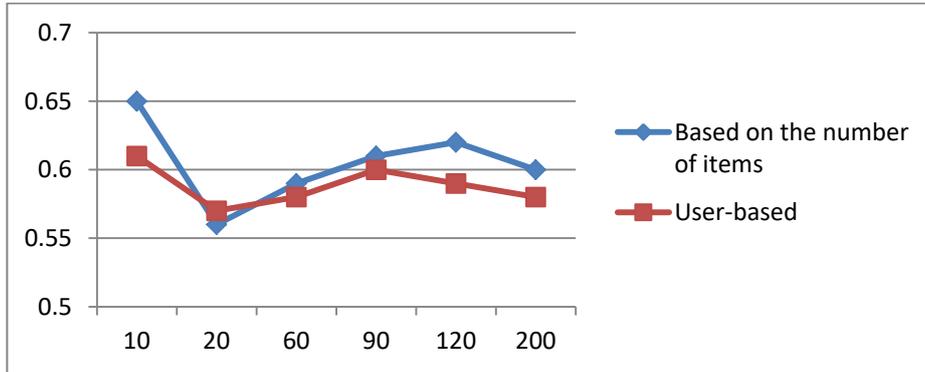

Figure 1. Comparison of user-based and item-based synergy effects in different neighborhoods

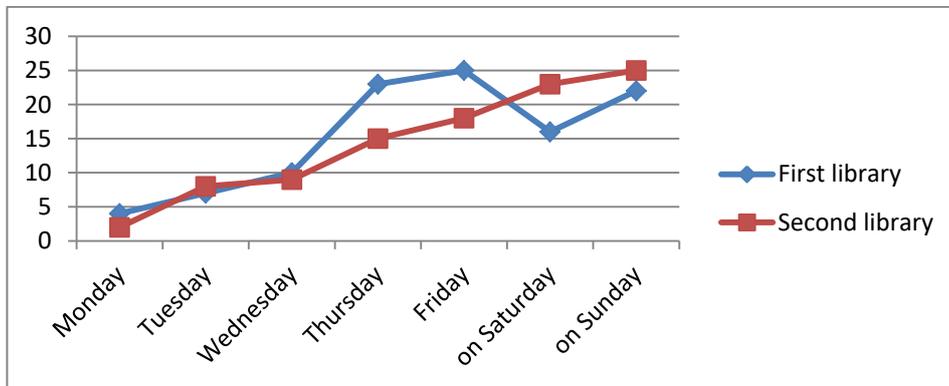### 4.2 Realization of shared book analysis function

Figure 2. Functional diagram of shared books between two libraries

It can be seen from Figure 2 that the two libraries of H University were investigated and compared. The book resource sharing function includes three parts: book selection, reader information selection, and submission and borrowing.

4.3 Student Data Analysis

Table 1. Book analysis results of some students

Reader barcode	Types of	Unit	Total borrowing	K-Means
0421	2	25	12	Cluster-1
0422	3	25	12	Cluster-3
0423	3	25	18	Cluster-2
0424	3	25	11	Cluster-1
0425	2	25	19	Cluster-1

Table 1 shows the clusters to which some student readers belong.

From Table 1, we can see some clusters of student readers. According to the clustering results, college student readers can be divided into two types, which are characterized by the largest number of borrowed books, with an average of 34 books borrowed per book. People; The second category has 317 students, which is characterized by a larger number of books, with an average of 12 books per person. Data analysis and its application terminal are characterized by wide coverage, high penetration rate and convenience. Mobile devices allow users to carry and connect to the Internet anytime, anywhere. If it is necessary to reserve a shared service space, students can use mobile devices such as mobile phones or tablets to log in to the system to make reservations, which is convenient, fast and easy to operate.

4.4 Use Big Data Technology to Promote Intelligent Conversion of Library Functions

Big data technology is widely used in the service mode of university libraries, which promotes certain changes in library functions. First, the library's information management model has changed from manual to intelligent management. Under the background of big data, the traditional manual information service model has been unable to effectively manage books, and can not meet the reading needs of readers. Therefore, big data technology can be used to Data are analyzed and researched more accurately and effectively, providing readers with accurate and comprehensive data information through intelligent services. Readers can quickly and accurately query book information through intelligent retrieval equipment, thereby effectively reducing book search time and improving readers' satisfaction with library services. Second, the information service mode of university libraries has changed from passive to active. Passive information services lead to poor readers' satisfaction with library services and fail to reflect readers' voluntary wishes. Although big data technology is applied to the service mode of university libraries to effectively improve the quality and efficiency of library services, it also faces huge challenges.

Big data technology is essentially a modern emerging technology and has certain unknowns. Therefore, there are many problems in the actual application of university library services. In order to ensure the effective application of big data technology in university library services, it is necessary to Reasonable analysis of the problem and a relative solution.

5. Conclusion

In summary, in order to ensure the smooth development of library management and provide readers and users with the best personalized services, modern universities must strengthen the innovative application of big data technology. From the perspective of big data, university libraries should establish a scientific and complete intelligent service system, innovate the management mode and content of intelligent services, organize a professional and complete technical talent team, strengthen the collection and processing and analysis of user data, and play a role

References

[1] Valverde A, Barquero V, Soler C. The application of computer-assisted semen analysis (CASA) technology to optimise semen evaluation. A review[J]. Journal of Animal and Feed Sciences, 2020, 29(3):189-198.

[2] Danish M S, Ranjan P, Sharma R. Valuation of patents in emerging economies: a renewal model-based study of Indian patents[J]. Technology analysis & strategic management, 2020, 32(4):457-473.

[3] Nie Y, Wang J, Guo S, et al. POD Scale Analysis of Turbulent Premixed Flame Structure at Elevated Pressures[J]. Combustion Science and Technology, 2019(10):1-23.

[4] Nivedhitha G, Rupavathy N. Data mining in personalized service of digital library[J]. International Journal of Engineering & Technology, 2018, 7(1-7):51.

[5] Li S, Hao Z, Ding L, et al. Research on the application of information technology of Big Data in Chinese digital library[J]. Library Management, 2019, 40(8/9):518-531.

[6] Liu J, Chen C. New Measures of Chinese University Library Knowledge Service under the Cloudization Background[J]. Intelligent Information Management, 2018, 10(3):79-86.

[7] Jia. A Comparative Study of Resource and Service Needs of International Students at an Academic Library[J]. The Southeastern Librarian, 2019, 67(1):3-3.

[8] Schneider B I, Miller B R, Saunders B V. NIST's Digital Library of Mathematical Functions[J]. Physics Today, 2018, 71(2):48-53.

[9] Wenige L, Ruhland J. Retrieval by recommendation: using LOD technologies to improve digital library search[J]. International Journal on Digital Libraries, 2018(1):1-17.

[10] Shiri A, Stobbs R. Community ˇriven user evaluation of the Inuvialuit cultural heritage digital library[J]. Proceedings of the Association for Information Science and Technology, 2018, 55(1):440-449.