Original Research Article

The application of big data analysis technology in credit bank information system

Yu Song[1] *, Jiangyang Xu*[2]

1. School of Mechanical Engineering Xijing University Shaanxi, China,710123

2. Tongyu Traffic Research Co., LTD Shaanxi traffic control Shaanxi, China, 710119

Abstract: In this paper, the special case UV decomposition of SVD decomposition, a common algorithm of big data analysis technology, is used in the personalized recommendation of credit bank system. This method can use statistical data association to find data relationship and obtain useful information to make accurate recommendation.

Keywords: Big data analysis technology; Big data analysis technology; Personalized recommendations

1. Introduction

Driven by the concept of lifelong learning, many countries and international organizations are actively exploring the establishment of a new type of learning management system with learning achievement certification as the main content among different types of education, and a relatively complete system has been formed. However, countries (regions) differ in their choice of institutional models and have different emphases in practice. For example, Australia, the United Kingdom and other countries through the establishment of national qualifications framework system, so that formal education certificate and non-formal education certificate, academic certificate and vocational qualification certificate can recognize and communicate with each other; Canada, the United States and other countries have established credit certification systems by signing inter-school credit mutual recognition agreements; South Korea has established a “credit bank” system for learning achievement accumulation and certification to encourage more non-formal higher education system learners to obtain university diplomas. In recent years, the theoretical research on the learning outcome certification system in China has achieved certain results in terms of concept sorting, system design and mode selection. Provinces and cities have also begun to rely on the Open University or Radio and Television University to conduct research and exploration of the learning outcome certification system with regional credit bank pilot model. Many universities, including the Open University, are also exploring the learning results certification system on their campuses, and the practice of credit bank has also entered a key exploration period. In the information society, credit bank information system is an

carrier of learning outcome information management and an essential element for the exploration of learning outcome certification system. The establishment of a personalized recommendation system on the credit bank platform can effectively provide learning information to learners, help learners better choose appropriate courses, and ultimately promote the establishment and development of learning outcome certification system.

2. Big data analysis technology and credit bank information system

The credit bank obtains a huge amount of data about learners. Therefore, we need an effective method to process and analyze the learner-related data, so as to correctly predict the needs of learners and recommend relevant courses to them. And big data analysis technology is just such a technology to meet our needs, it can process massive data, such as classification and correlation. Classification and association technology does not care about causality, but directly uses statistical data association to find data relations, so as to extract useful information.

It is often difficult to know what subjects a learner is hoping to learn. At this time, we need to analyze the learning information of the learner and the learning information of other learners in order to make a reasonable course recommendation to the learner. This process is very similar to the product recommendation process of Taobao and other websites.

Big data analysis techniques can use data mining techniques or mathematical statistical analysis. Data mining is a specific data analysis technique that focuses on modeling and knowledge discovery for the purpose of prediction rather than pure description. Mathematical statistical analysis is the application of knowledge of mathematical statistics to analyze data, focusing on describing the properties of the data.

The commonly used algorithms of data mining mainly include: ID3 algorithm, k-means, logistic regression, degree learning technology, Decision trees, SVD decomposition, random forests, large-scale matrix operations, and more recently deep ,etc. In this paper, the UV decomposition algorithm in SVD decomposition is used to process and analyze the data, so as to extract the required data. This technology is simple to implement, the data demand is not very high, and can be effectively recommended. Therefore, we apply this technology to the personalized recommendation platform of our credit bank.

3. SVD decomposition

SVD decomposition, singular value decomposition, has been applied in many fields. It is mainly used to classify database files in data mining. This method is mainly used to reduce the dimension of data.

The principle of SVD decomposition is as follows:

For any matrix A, the singular value decomposition mathematical formula is defined as follows:

$$A=U\sum V^T \tag{1}$$

In formula (1), A represents an m ×n matrix, U represents a m ×m unitary matrix, ∑ represents a semi-positive definite m ×n diagonal matrix, and V^T represents an n ×n unitary matrix. The elements on the diagonal of the minor quadrilateral matrix are the singular values of A.

If the diagonal elements of the ∑ are arranged in order from largest to smallest,then the ∑ is uniquely determined by A. That is:

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_r & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \tag{2}$$

If the elements on the diagonal of the ∑ matrix are all 1, it is a special case of SVD decomposition, that is, UV decomposition. It is the UV decomposition used in this article.

4. The personalized recommendation of UV decomposition applied to credit bank information system

For a learner, courses of interest are often related. Therefore, we can analyze the courses that learners have learned in the past to judge the courses that learners may need to choose and make further recommendations.

For example, we can classify the courses in the credit bank platform according to Chinese,English, mathematics, computer and other course categories, and then evaluate the value of certain courses according to learners' elective situation, survey feedback and satisfaction. In our recommendation system, we divide the value of a course into 1-5 grades (the greater the value, the greater the value of the course, and the more likely it is to be taken by learners). As shown in Table 1, where A, B, C and D represent different learners, the blanks are the courses that have not been taken or have been taken without satisfaction evaluation.

Table 1. Values of different categories for different objects.

learner	Course Category 1	Course Category 2	Course Category 3	Course Category 4	Course Category 5	Course Category 6
A	4			5	1	
B	5	5	4			
C				2	4	5
D	3					

We can use an n× m matrix M with blank terms to describe the value of m different elective courses for n learners. For example, in Table 1, we can use a 4 ×6 matrix to describe, we only need to give the value of blank items, we can determine the value of unselected courses to learners. According to the value of unselected courses, we can decide which kind of courses to recommend to electives first.

We estimate the value of the blank term of the matrix M by UV decomposition.

The algorithm is shown as follows:

For an n ×m matrix M, we can assume that U is an n ×d matrix, and the V is an d ×m matrix (d can be any positive integer), and for each of the U, V matrices element is assigned an initial value (in general, the initial value is assigned 1), so that P=UV. Then by adjusting the way of each element in U and V many times, in order to make P as close to M as possible, we can use the sum of squares of the difference between M and the non-empty element of P as the metric standard, and the sum of squares is recorded as the root-mean-square error, that is, the smaller the root-mean-square error, the closer P is to M. When all the elements in U and V are adjusted only once, it can be remembered as an iterative process. The values of all elements in U and V are constantly changed in an iterative way. When the absolute value of the root-mean-square error of an iteration process and the root-mean-square error of the previous iteration process is less than a certain threshold, the iterative process is terminated and UV is determined. In this case, the root-mean-square error can reach a local minimum, but the local minimum may not be the global minimum. We can obtain different local minima by assigning different initial values to UV, estimate the global minimum with the minimum value of all local minima, and save the value of UV in this case, and find out the P-value through P=UV. The M blank value is estimated by the corresponding element of the P value. In this way, it is possible to decide which type of data is selected first.

The elements are adjusted as follows:

If we need to adjust an element in row r, column s of the matrix U, we first *replace* u_{rs} with x, then:

$$p_{rj} = \sum_{k=1}^{d} u_{rk}v_{kj} = \sum_{k\neq s} u_{rk}v_{kj} + xv_{sj} \tag{3}$$

Where p_{rj} is an element in row r, column j of matrix P, and u_{rk} *is row r, column S of matrix* U

An element of column k, v_{kj} is an element of row r, column j of matrix V.

And the square of the difference between an element of column j in row r of an M matrix and an element of column j in row r of a P matrix

can be expressed as:

$$\left(m_{rj}-p_{rj}\right)^2=\left(m_{rj}-\sum_{k\neq s}u_{rk}v_{kj}-xv_{sj}\right)^2 \quad (4)$$

From equation (4), we can see that when we *replace* u_{rs} with x, the *difference between* m_{rj} *and* p_{rj} will be affected for all j. We want to ensure that the value of the root-mean-square error is the minimum, that s, we find the minimum value of equation (5).

$$\sum_j\left(m_{rj}-\sum_{k\neq s}u_{rk}v_{kj}-xv_{sj}\right)^2 \quad (5)$$

When we take the derivative of x, we can find the minimum value of the above equation when the derivative is 0. That is:

$$\sum_j -2v_{sj}\left(m_{rj}-\sum_{k\neq s}u_{rk}v_{kj}-xv_{sj}\right)=0 \quad (6)$$

Then x can be obtained as:

$$x=\frac{\sum_j v_{sj}\left(m_{rj}-\sum_{k\neq s}u_{rk}v_{kj}\right)}{\sum_j {v_{sj}}^2} \quad (7)$$

That is, we can adjust the elements of U by formula (7) in turn, and similarly, when we set any element of V to y, that is, $vrs = y$. We can find:

$$y=\frac{\sum_j u_{ir}\left(m_{is}-\sum_{k\neq r}u_{ik}v_{ks}\right)}{\sum_i {u_{ir}}^2} \quad (8)$$

That is, we can adjust the elements of V in turn by equation (8).

For example, if M is:

$$\begin{bmatrix} 5 & 2 & 4 & 4 & 3 \\ 3 & 1 & 2 & 4 & 1 \\ 2 & & 3 & 1 & 4 \\ 2 & 5 & 4 & 3 & 5 \\ 4 & 4 & 5 & 4 & \end{bmatrix} \quad (9)$$

We can assign the initial value UV to:

$$\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad (10)$$

We first set u_{11} to x, and by equation (8), we can find x=2.6.

By iterating over and over again, we can get the value of UV, and then we can calculate the value of M, including the value of the position of the M blank. That is, we can estimate the value of the course the learner did not choose, and then decide which course to recommend to the learner.

5. Conclusion

In this paper, UV decomposition is used in the personalized recommendation of credit bank information system, which can provide intelligent and accurate recommendation for learners' personalized learning needs, guide learners' learning direction, and improve learners' satisfaction.

Funded Project

Shaanxi Province education science "14th Five-Year" plan 2021 annual topic. project number: SGH21Y0280.

References

[1] Lian Yun, LV Youwei, WANG Qianwen. Construction and implementation path of lifelong Education Credit Bank 2.0 under the background of common prosperity [J]. Chinese Adult Education,2023(09):16-21

[2] Shao Wensha. Application of Big Data Analysis Technology in Credit Bank Information System [J]. Computer Knowledge and Technology,2018,14(26):245-246.

[3] Hake Ming. Credit Certification, transfer system and lifelong learning - the final construction in 2016 Speech at the Academic Forum on Overpass and Credit Banking System (Nanjing) [J]. Lifelong Education Research,2017(2):6-10.

[4] Rajaraman A, Ullman J D. Mining of massive datasets[M]. Cambridge University Press, 2012.

[5] He Liangliang. The application of SVD in text classification [D]. Guangzhou: South China University of Technology, 2012.

[6] Yu Yanfang, Han Shimei. Knowledge aggregation and personalized recommendation of credit bank platform China Distance Education,2017(3):45-51. (in Chinese)

[7] TANG Shubo, Chen Meiyan, Li Zhiping. Probe into the design scheme of open education credit banking system [J]. Audio-visual Education Research,2011(8):78-87.s