Taylor & Francis
Taylor & Francis Group

Journal of Interdisciplinary Mathematics

ISSN: 0972-0502 (Print) 2169-012X (Online) Journal homepage: https://www.tandfonline.com/loi/tjim20

Personalized recommendation system of e-commerce based on big data analysis

Hua Chen

To cite this article: Hua Chen (2018) Personalized recommendation system of e-commerce based on big data analysis, Journal of Interdisciplinary Mathematics, 21:5, 1243-1247, DOI: 10.1080/09720502.2018.1495599

To link to this article: https://doi.org/10.1080/09720502.2018.1495599

Journal of Interdisciplinary Mathematics
ISSN: 0972-0502 (Print), ISSN: 2169-012X (Online)
Vol. 21 (2018), No. 5, pp. 1243–1247
DOI : 10.1080/09720502.2018.1495599

Taylor & Francis
Taylor & Francis Group

Personalized recommendation system of e-commerce based on big data analysis

Hua Chen
Department of Agricultural Economics and Management
Zhejiang Agricultural Business College
Shaoxing 312000
China

Abstract:

In the traditional e-commerce system, the keyword matching algorithm is used to implement the function of commodity search, and only the goods containing the customer input keyword can be obtained. In view of this limitation, a personalized recommendation system for e-commerce based on big data analysis is studied. The use of improving the search algorithm for goods, personalization system rules of e-commerce and text matching algorithm are introduced, so in the search of goods, it not only displays the exact matching of goods, but provides the goods similar to their requirements, for reference. The research shows that the personalized recommendation system of e-commerce based on big data analysis can not only increase the transaction opportunities, explore potential customers, but also improve the level of personalized service. This is very important for the enterprise economy and the personalized development of e-commerce.

Keywords: *Big data analysis, e-commerce, Personalized recommendation, Algorithm research*

1. Introduction

Compared with the traditional e-commerce activities, the personalized way of e-commerce based on big data analysis in the form of transaction has many different, and therefore, it also brings about changes in the mode of commodity marketing[1-2]. E-commerce sites are faced with a serious problem: how to recommend the right goods to users when they browse the site, to overcome the adverse impact of information overload[3-4].

According to the above situation, a personalized recommendation system of e-commerce based on big data analysis is proposed to solve the information overload problem. The massive data are analyzed through big data tools, to realize users' business trend prediction, customer value mining and evaluation analysis of user perception, so as to design of personalized e-commerce with the user as the center, facing the specific business scene.

E-mail: 415113721@qq.com

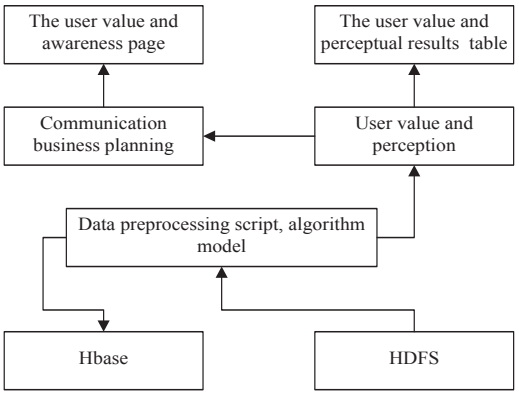

Figure 1

The framework of accurate e-commerce system based on big data analysis

2. Personalized recommendation system of e-commerce based on big data analysis

2.1 *Analysis of personalized recommendation system framework of e-commerce*

The proposed personalized recommendation system of e-commerce based on big data analysis can be divided into four levels, which are data layer, management layer, business layer and presentation layer, and each layer is closely related to big data. According to the precision of the electronic commerce system[5], the e-commerce framework based on big data is constructed as shown in Figure 1.

It can be understand from Figure 1, the framework of precision electronic commerce system based on the big data analysis is: user value, perception page and perception result table in the first layer represent the big data display layer; communication business rules and user value and perception of the second layer, represent a big data business layer; data preprocessing script and algorithm model in the third layer is the data management layer; and the fourth layer contains detailed data, user information, complaint data HDFS and the big data layer represented by Hbase[6-7].

2.2 *The main functions of personalized recommendation system for e-commerce*

The main functions of the personalized recommendation system for e-commerce based on the big data analysis are as follows:

(1) web page extraction

The component extracts the relevant web pages from the network and downloads them through the web crawler. After downloading, it is stored on the server. These pages are pre-processed and used to build the recommendation library.

(2) feature analysis

The component is used to analyze pages downloaded from the Internet, to extract key features, and to categorize (cluster) documents using the two classification model (SVM) to store web pages and classification results in the recommendation library[8].

(3) behavior record

The component records the user's access date and behavior path.

(4) interest modeling

The user interest model is created by collecting the user information, and the user interest model is generated.

(5) information recommendation

This component acts as a recommendation engine, looking for the next web document most likely to be accessed according to the access path and the user interest model, which will appear in the library according to the keyword recommendation associated with the web page to the user[11-12].

3. Research on the algorithm for the personalized recommendation system architecture of e-commerce based on big data analysis

3.1 *establishment of the rules of personalized recommendation algorithm for e-commerce*

Here we can measure for goods from different angles: there are M items, N characteristics which are representative but not related of the commodity are founded out, the membership degree of commodity m in feature n is umn, where, m=1,2,... M, n=, 1,2,... N. Thus, if the feature is selected properly, for the two commodities of X and Y, uxn, uyn, or uxn are approximately equal to uyn, n=, 1,2,... N. Then, we think that the two goods are similar. In some cases, they are substitutes. When the customer is browsing the goods X, we can give the goods Y to him to compare. The more features are selected, the smaller the classification is, the more accurate the similarity description is.Specific methods are:

(1) to establish a Thesaurus, list the words that are commonly used by customers, and each word represents certain features, some vocabulary may represent several characteristics, therefore, value of the Pkn of each word in each feature is "1" or "-1";

(2) to count the number of times of each word "k" appears in the commodity m, and then the degree of membership of the commodity m on the feature n can be represented as:

$$U_{mn} = \frac{\sum_{i=1}^{k} w_{mi} \times p_{kn}}{\sum_{i=1}^{N} |\sum_{i=1}^{k} w_{mi} \times p_{ki}|} \tag{1}$$

Where: m is the goods; n is the representative and unrelated feature; umn is the degree of membership of commodity m on feature n; k is each word; w is the number of occurrences; i is the vector, Pkn is the value of each word in each feature.

3.2 *Personalized recommendation algorithm of e-commerce*

For both the commodity m and the keyword k in e-commerce, it is available:

$$\begin{aligned} &\left|u_{\min}\right| \le 1 \\ &\left|p_{kn}\right| \le 1, n = 1, 2, \ldots, N \\ &\left|d_{mk}\right| \le 1 \end{aligned} \tag{2}$$

Where: dmk is the similarity of goods. If m and k are similar in N characteristics, the membership of each feature is the same "+" or same "-", and the corresponding product is "+"",

$$Ud_{mk} \geq MAX\left(u_{mn}\right) \times MAX\left(p_{kn}\right) \tag{3}$$

Usually the dmk is larger. And if m and k are not similar, it indicates that they are opposite in some characteristic, and the membership of them is opposite sign, the corresponding product is "-", then mkd is smaller. Considering two kinds of extreme circumstances, when the membership of goods m and keyword k is 1 only in a characteristic, the other is 0, then dmk =1, it is thought that m and k are completely similar; while the membership of m and k is 1 and -1 respectively in a characteristic, the other is 0, then dmk =-1, it is thought that m and k are not completely similar. Therefore, for keyword k, it can recommend the appropriate personalized recommendation algorithm by selecting the larger items of dmk.

4. Conclusion

A text matching algorithm is used to define the membership of goods for a set of features in this paper, and the search keywords input by customers are also defined as the same degree of membership, so as to find out the similar goods to the input characteristics of customer. At the same time of accurate search, it can display the goods which are similar to the search intention to the customers, and increase the sales opportunities, and also reflect the personalized business intelligence services. By testing in a small range, the effect is remarkable. When the method is applied to comprehensive commodity search, due to the variety of goods, the proposed method is difficult to grasp when selecting the representative features and computing membership, so it needs further study.

5. Acknowledgement

Fund Project: Special Funding project of high level talent research of Zhejiang Agriculture and Business College (No. GCCKY201602).

References

[1] R. J. Hathaway, J. W. Davenport, J. C. Bezdek. Relational Dual of the C-means Algorithms. *Pattern Recognition,* 22(2015) 205-212.

[2] G. Y. Wang, J. S. Xu. TCBLSA A new method of Chinese text clustering. *The computer Simulation*, 30(2014) 21-22.

[3] P. Resnick, H.R. Varian. Recommender systems. *Science and technology,* 40(2016) 56-58.

[4] X. L. Shao. The user the design and implementation of personalized recommendation system. *Computer engineering and design,* 30(2009) 4681-4685.

[5] B. C. Kwon, S. H. Kim, T. Duket. Do People Really Experience Information Overload While Reading Online Reviews. *International Journal of Human-Computer Interaction*, 31(2015) 959-973.

[6] Y. Sasaki, D. Kawai, S. Kitamura. Unfriend or ignore tweets?: A time series analysis on Japanese Twitter users suffering from information overload. *Computers in Human Behavior*, 64(2016) 914-922.

[7] C. F. Liu, K. M. Kuo. Does information overload prevent chronic patients from reading self-management educational materials. *International Journal of Medical Informatics*, 89(2016) 1-8.

[8] P. Li, Y. Sun. Modeling and performance analysis of information diffusion under information overload in Facebook-like social networks. *International Journal of Communication Systems*, 27(2015) 1268-1288.

[9] Al-Hassan, Q.M. 2017. On some interesting properties of a special type of tridiagonal matrices. *Journal of Discrete Mathematical Sciences and Cryptography,* 20(2), 493-502.

[10] Yuksel, G., Cetin, M. 2016. Outlier detection in a preliminary test estimator of the mean. *Journal of Statistics and Management Systems*, 19(4), 605-615.

[11] T. Yuan, J. Cheng, X. Zhang. Enriching one-class collaborative filtering with content information from social media. *Multimedia Systems,* 22(2016) 51-62.

[12] Y. D. Seo, Y. G. Kim, E. Lee. Personalized recommender system based on friendship strength in social network services. *Expert Systems with Applications*, 69(2017) 135-148.

Received November, 2017

Revised January, 2018