Design and Research of e-Commerce Personalized Recommendation Algorithm

Jie Hu

Nanchang University of Technology, Jiangsu, Nanchang, 330044, China

Keywords: Design, Research, E-commerce, Personalized recommendation algorithm

Abstract: In the era of big data, the amount of Internet data has exploded, and how to quickly obtain valuable information from the massive data information has become a very challenging task. In this context, an e-commerce recommendation system came into being. It can dig out the different behavior preferences of users and provide personalized recommendation services for different users.

1. Introduction

At present, there are a large number of e-commerce websites, the total amount of information resources is huge, and the growth rate is rapid, which causes users to experience information trek and information overload problems in the entire commodity space, and the efficiency of information use is reduced. Existing e-commerce website search systems and search engines mainly filter information through search conditions entered by users, without personal considerations, and cannot provide convenient, efficient, and accurate services according to user needs. How to adapt to the transition of e-commerce to a user-centric development model requires the application of a personalized recommendation system. The recommendation system can provide different users with services that meet their individual needs, improve the efficiency of users seeking knowledge from information, establish long-term and stable relationships with customers, and make the website invincible. The recommendation system, by simulating store sales personnel, recommends products that meet their interests and hobbies to target users according to their hobbies. The recommendation algorithm is the most important part of the recommendation system, and it is related to the recommendation quality of the recommendation system. Collaborative filtering algorithm is currently the most widely used and most successful recommendation algorithm. However, with the rapid increase in the number of Internet users, collaborative filtering algorithms have become increasingly serious in terms of data sparsity and authenticity of user ratings, which affects the quality of the recommendation system.

E-commerce developed under the background of the era of big data, the amount of data in its field has shown a geometric growth, and the content of goods is rich and diverse. While providing users with a dazzling array of goods, it also provides users with more choices. Customization, personalization and differentiation have become the core competitiveness of enterprises. The two major trends and challenges of data analysis are: the expansion of data volume and the growth of demand for in-depth data analysis. People's lives have been closely related to and inseparable from e-commerce. The most important issue for users is how to quickly and accurately select the products they need.

Faced with massive amounts of information, personalized recommendation has become one of the most effective means to solve information overload, and it is a hot spot in the academic and e-commerce circles. In this context, e-commerce recommendation system came into being. The e-commerce recommendation system can capture key data from the rich data information, mine potential customers for merchants, expand sales scope, and provide product recommendations for old customers to expand user groups. With the continuous improvement of user needs, the recommendation quality of current e-commerce recommendations can no longer meet the requirements of users and businesses. With the continuous development and expansion of e-commerce, e-commerce models are diversified. On the one hand, the continuous increase in the amount of product data causes the data processing of the recommendation system to be untimely, and users cannot quickly and accurately search for the products they want. On the other hand, users

  DOI: 10.25236/mmmce.2021.021

the demand for diversification has also caused the recommendation system to be unable to recommend products that users are potentially interested in, and the recommended content is not diversified enough.

At present, many shopping websites at home and abroad have added e-commerce systems to their existing applications, such as Amazon, e Bay, Dangdang, Taobao, JD, etc., which can guide users in personalized shopping, and some video and music websites, such as Youku, iQiyi, Xiami Music, Netease Music, Today's Toutiao, etc. can guide users to personalize browsing, and also push personalized video and music to users. All of these use various forms to varying degrees. Recommended system. However, with the widespread use of e-commerce recommendation systems on various websites, a large number of user browsing records and purchase records have been accumulated in the database. The huge amount of data and the complex data structure are beyond the load of ordinary stand-alone programs, but with high performance The computer is expensive, which forces the original calculation and storage model to be upgraded and improved. How to make the recommended content generated by the system closer to the needs of users is the core of this article.

2. Recommendation Algorithm Design Ideas under the Platform

With the rapid development of e-commerce, people's lives are inseparable from it. People interact with e-commerce all the time, and huge user access data is generated during the interaction. In addition, the number of users is increasing every day. In this case, no matter which recommendation algorithm is used, it is necessary to perform multiple visits to the huge database system during data calculation. On the one hand, the huge amount of data processing greatly increases the running time of the recommendation algorithm. On the one hand, a large amount of heterogeneous data has brought great pressure to the calculation of recommendation algorithms, increasing the computational burden of the algorithm, and even some algorithms can no longer meet the processing of heterogeneous data. The algorithm design idea under the cloud platform is to deploy the recommendation algorithm on the Hadoop platform. First, use HBase to store and manage a large amount of information to solve the storage and calculation problem of heterogeneous data, and then use the Map Reduce programming model to decompose and merge the algorithm, and The algorithm is divided into a part that must be executed sequentially and a part that can be calculated at the same time to maximize the parallel calculation of the algorithm; finally, use HDFS to intelligently manage the information, count the number of times the product is accessed, and calculate the depth of the product being accessed, and the algorithm's comprehensive calculation result is generated Push content. The core of algorithm parallelization is to find out the elements of algorithm parallelization. These parallelized elements can be divided into two aspects, one is the parallelization of data flow, and the other is the parallelization of tasks [50]. Parallelization of data flow is to parallelize the traditional sequential execution of data flow. Parallelization and decomposition of tasks can be divided into tasks independent and non-independent. When the tasks are independent, each task is independent of each other and does not affect each other, and different tasks can be performed at the same time; when the tasks are not independent, each task depends on each other and affects each other, and different tasks cannot be performed at the same time. Traditional serial execution refers to the interdependence between tasks. The output of task A is the prerequisite for the input of task B. There are three serial executions of task A, task B and task C. The input premise of task C is that of task B. The input premise of task B is the output of task A. The three tasks restrict each other and influence each other, and the degree of mutual dependence is high.

3. Design and Implementation of Personalized Recommendation Subsystem

The personalized recommendation subsystem is realized by establishing two parts: offline data preparation and online recommendation. The data is preprocessed through the offline state, and the data mining method is used to dig out the user's hobby model from the Web server log and the

product sales record, and calculate the user's evaluation value of the product category. Finally, the collaborative filtering algorithm based on user-commodity category is used to realize online personalized recommendation.

Offline data preparation According to the user's clear purpose when browsing the website, the user's frequent access path can represent the user's interests and hobbies in a certain period of time. Using data mining methods, the frequent access paths are extracted from user log files, the frequent access paths of users are clustered, and the user behavior model library is established. Users belonging to the same type of user behavior model library have similar access paths. When a new user visits the website, the access path generated by it can be directly compared with the user behavior model library, divided into user clusters with similar browsing paths, and recommendations are generated based on the records in the model library, which solves the recommendation Problems with new users in the system. How to build a user rating matrix is the key to the collaborative filtering algorithm. The collaborative filtering recommendation algorithm mainly finds the nearest neighbor users who have similar scores with the target user from the user rating matrix, and generates recommendations for the target user according to the purchase situation of the nearest neighbors. The authenticity of user rating values has always restricted the development of collaborative filtering recommendation algorithms. This recommendation system uses log files and user purchase databases to extract users' potential rating information for products, and establish a corresponding user rating matrix to avoid users directly. The sparsity caused by product ratings and the inconsistent scoring standards affect the accuracy of the similarity calculation results.

The establishment of an offline model is to accumulate a large amount of knowledge elements required for personalized recommendation. By applying the user-commodity collaborative filtering recommendation algorithm in the user behavior model library, that is, finding the nearest neighbors of the target user in the approximate neighbor set, thereby reducing the input size of the traditional collaborative filtering recommendation, reducing the sparsity of the matrix and making the similarity The measurement results are more accurate. The implementation of the “Tesco.com” recommendation system is that after the user completes the shopping operation, the system pops up the recommendation results obtained according to the recommendation algorithm in a visual form to realize the personalized recommendation service for the user.

The implementation process of the personalized recommendation system. The implementation process of the personalized recommendation system includes collecting data, preprocessing the data, mining the recommendation model library, and generating corresponding visual recommendations. Data collection and data preprocessing are the basis of mining recommendation models, and the mining of recommendation models is the core of achieving personalized services in e-commerce.

(1) Data collection The implementation of the recommendation function requires a lot of data preparation, so data collection is the basis of the recommendation system, which provides data support for the recommendation system to mine models that meet the user's hobbies.

(2) Data preprocessing The web server log is not specifically used for data mining. Data preprocessing must be performed to obtain “pure” data. E-commerce websites include many web pages. The recommendation system only cares about the pages related to the recommended content. In the data preprocessing stage, only log records related to the recommended products are kept, thereby effectively reducing the noise of the data.

(3) Mining the recommendation model Mining the recommendation model is the core of the e-commerce recommendation system. It is mainly based on the corresponding recommendation algorithm to find a model database that meets the individual needs of users.

(4) Generate recommendation Generate recommendation results based on the needs of target users and present them to users in a visual form.

4. Problems in the Design and Research of e-Commerce Personalized Recommendation Algorithm

E-commerce personalized recommendation technology and recommendation system have been researched and developed for many years, and once entered a low ebb, both in theory and application, there are certain limitations, and continuous research and improvement are needed. Moreover, domestic research in this area is only at the theoretical stage, and its application is still lagging behind the international level.

For the current research on personalized recommendation in e-commerce, I think there are still the following problems and shortcomings: In terms of recommendation algorithm research, the current mainstream recommendation algorithm research is not balanced, and the collaborative recommendation algorithm is sparse and new user problems and other classic problems. The solution lacks a better solution; data sparseness, the current data acquisition mainly relies on the user's explicit scoring of the product, resulting in extremely sparse data, and due to the inconsistent scoring standards and subjective factors in the scoring, it affects The authenticity of the score; research is too concentrated on improving the performance of the recommendation algorithm, lack of research on the development and application of the recommendation system, especially the research on the integration with enterprise application systems, in assisting enterprises in marketing, customer management and enterprise business intelligence Lack of research.

5. Conclusion

In recent years, with the increasing demand for machine learning, data mining, large-scale network applications and the development of high-performance computers, the development of personalized recommendation systems for e-commerce has been promoted. The application of combined strategy recommendation technology will further provide new opportunities for improving recommendation systems.

Acknowledgment

General Science and Technology Project of Jiangxi Education Department. No. (GJJ191015).

References

[1] Mi Kefei, Zhang Yong, Xing Chunxiao, etc. Analysis of open source recommendation system for big data. Computer and Digital Engineering, vol. 6, no. 20, pp.121-122, 2016.

[2] He Yao, Wang Wenqing, Xue Fei. Research on massive data mining based on cloud computing. Computer Technology and Development, ,vol. 2, no. 7, pp. 175-176, 2019.

[3] Zhai Yanlong, Luo Zhuang, Yang Kai, etc. Research on a high-performance massive data processing platform based on Hadoop. Computer Science, vol. 156, no. 40, pp. 159-161, 2018.

[4] Xu Wenlong, Yan Yongjian, Fan Bingbing. A collaborative filtering algorithm combining similarity of items. Computer Applications and Software, vol. 23, no. 40, pp. 293-295, 2018.

[5] Huang Qiao. Improvement and research of collaborative filtering algorithm in e-commerce recommendation system. North University of China, no. 1, pp. 6-7, 2015.