ISSN: 2664-9640

Research on the application of collaborative filtering algorithm in financial product recommendation

Qiutong Liu, Xirui Lv and Ziyun Li

North China University of Technology, Tangshan 063000, China

Abstract

With so many complicated wealth management products, how to use information technology to provide customers with suitable product recommendations has become one of the main problems faced by the general public and related financial institutions. This paper uses the current mainstream financial institutions' rating information on wealth management products, through data coding and cleaning, to remove invalid and duplicate data, and preliminarily analyze the market characteristics of wealth management products. The collaborative filtering algorithm used in the recommendation is introduced in detail, and the collaborative filtering algorithm based on the financial product rating system is used to comprehensively recommend purchasing customers, and the three indicators of recall rate, accuracy rate and coverage rate are used to test the scoring results. The collaborative filtering algorithm based on the wealth management product rating system used in this paper has an accuracy rate of 18.70%, indicating that this algorithm has certain accuracy and credibility in the future comprehensive rating and recommendation of wealth management products, and can provide customers with relevant products for purchasing related products. strong evidence. In the future, we will try to obtain more rating data, continue to improve the accuracy of the model, and at the same time combine with other types of recommendation algorithms to continue to improve the level of recommendation.

Keywords

Collaborative Filtering Algorithm, Financial product, Recommendation algorithm.

1. Introduction

With the development of big data and information age, our way of life has become intelligent. People's use of the Internet will be recorded in the form of data. The dramatic increase in the amount of data makes it difficult for us to quickly locate the information we want from the massive amount of information. Among them, the financial and wealth management industry, which is closely related to us, is also transforming into a "data-driven enterprise". According to statistics, the utilization rate of users purchasing wealth management products through the Internet has reached 13.1% [1], and the securities market has produced information dependence. People are faced with a large number of products, various types of consultation, and it is difficult to choose real-time securities. Therefore, in order to help people manage their money, how to provide customers with personalized financial product recommendation services through data analysis and data mining has become a hot issue based on information from the trading market.

Personalized recommendation can solve the problem of information overload [2]. The recommender system can mine potential information of users through user interaction records, filter redundant information, and participate in really useful information in user demand prediction. In recent years, personalized recommendation has been widely used in Internet-related industries including e-commerce, which involves recommending books on websites [3],

recommending news, movies, etc., to provide consumers with intelligent product recommendations. However, such analysis and research are rarely applied in the domestic securities industry. It can be seen that the personalized recommendation service system has broad development prospects in the securities industry.

The core of the recommendation system is the algorithm, which is of great significance to the research of the recommendation algorithm. Some scholars have made great contributions to the research and application of recommendation algorithms. By mining data, collecting user characteristics, and introducing collaborative filtering algorithm into recommendation, Zhou Ying can make recommendations based on users' real-time behavior and user-clustered popular products [4]. According to the market information purchased by users, Bian Yuning applied the improved collaborative filtering recommendation algorithm to the precise marketing of banks [5].

This paper mainly selects the collaborative filtering recommendation algorithm [6]. The algorithm calculates the similarity between users according to the user's historical interaction records, and finds nearby users through similar features between users. Recommend products that users like to the corresponding target users. In order to improve the accuracy of the recommendation system, a coverage coarse-grained computing model is added to the collaborative filtering recommendation process [7]. It not only avoids the shortcomings of attention-based Transformer models, but also enables dynamic switching. The research of this topic is mainly through the analysis of users' opinions on financial products, which can build a personalized recommendation service system that users can satisfy, and make accurate and real-time recommendations for users. Help the transformation of the financial trading market and adapt to the development of the information age.

2. Introduction to Collaborative Filtering Algorithm

2.1. The concept of collaborative filtering algorithm

Goldberg first proposed the concept of collaborative filtering algorithm. So far, many scholars have improved and designed the collaborative filtering algorithm. The current collaborative filtering algorithm is mainly divided into two algorithms: user-based collaborative filtering and commodity-based collaborative filtering. The most widely used big data recommendation algorithm, according to which the user's shopping preferences can be analyzed, so as to realize the personalized product recommendation for the user. The ultimate realization of "things gather together, people divide by groups".

2.2. User-Based Collaborative Filtering Algorithm

User-based collaborative filtering algorithms believe that some user groups with similar basic characteristics have similar product needs [4]. The main principle of this algorithm is to obtain user behavior information data according to different users' product purchases or other behaviors such as likes and collections, so as to evaluate the data [2].

(1) Calculate similarity

At present, there are many methods that can calculate the similarity between users, including cosine similarity, Pearson correlation, etc. The commonly used method is Pearson correlation, and the formula is:

$$sim(u,v) = \sum_{i \in I(u,v)} \frac{(r_{ui} - \overline{r_u})(r_{vi} - \overline{r_v})}{\sqrt{\sum_{i \in I(u,v)} \left(r_{ui} - \overline{r_u}\right)^2 \left(r_{vi} - \overline{r_v}\right)^2}} \tag{1}$$

Among them, $I(u,v)$ represents the rating set of user u and user v for common products, r_{ui} represents the rating value of user u for product i, r_{vi} represents the rating value of user v for

product i, and $\overline{r_u}$ represents the average rating value of user u for each product , $\overline{r_v}$ represents the average rating value of user v for each product. The closer the value of the Pearson correlation coefficient is to 1 or -1, the stronger the correlation between the two, and the closer to 0, the weaker the correlation.

(2) Get the nearest neighbor user set

After obtaining the similarity of the two users, in order to improve the accuracy of the recommendation, it is necessary to sort the similarity of the other users in the user set I and the target user u from high to low, and extract the sorted top k users as the target user's The nearest neighbor users form the set N of nearest neighbor users.

(3) Calculate the product recommendation list

It is necessary to recommend products that users may be interested in. At this time, the predicted score value of user u for product i should be calculated [2]. The formula is as follows:

$$P_{u,i} = \overline{r_u} + \frac{\sum_{v \in N} sim(u,v) \cdot r_{ui}}{\sum_{v \in N} |sim(u,v)|} \tag{2}$$

Among them, $P_{u,i}$ represents the predicted rating of product i by user u, N represents the set of k nearest neighbor users of user u, and $sim(u,v)$ represents the similarity between user u and user v.

After the calculation of the above three steps, the predicted score of the target user u for the wealth management product i is obtained. In order to make the product recommendation more accurate, it is necessary to sort the products according to the predicted score from high to low, and specify a certain benchmark score. Value x, select all financial products higher than x to recommend to target users.

2.3. Product-Based Collaborative Filtering Algorithm

Different from user-based collaborative filtering algorithms, product-based collaborative filtering algorithms recommend products that are highly similar to previously purchased financial products to users by calculating the similarity between products. At present, this recommendation model has been widely used in shopping platforms. Shopping platforms often analyze the characteristics of products according to the historical purchase records of user u, so as to push products with high similarity to the favorite products for users, so as to increase user stickiness. The product-based collaborative filtering algorithm can be divided into the following three steps:

(1) Calculate the similarity between wealth management products

Similar to the user-based collaborative filtering algorithm, the similarity between wealth management products needs to be calculated first, using the Pearson correlation coefficient method, the Jaccard coefficient method or the cosine similarity method to calculate the similarity. The Pearson correlation coefficient method has been introduced in the user-based collaborative filtering algorithm, and the Jaccard coefficient method will be introduced here.

The Jaccard coefficient is designed to measure the similarity of the product sets liked by user u and user v. The closer the Jaccard coefficient is to 1, the higher the similarity between the two product sets, and the closer to 0, the lower the similarity. The calculation formula is as follows:

$$sim(u,v) = \frac{|N(u) \cap N(v)|}{|N(u) \cup N(v)|} \tag{3}$$

(2) Obtain the nearest neighbor wealth management product

Sort the similarity between the calculated other wealth management products and the currently analyzed target wealth management products from high to low. Similar to the user-

based wealth management products, the top k products before the test are still extracted as the nearest neighbor wealth management products of the target wealth management products. , indicating that users who have purchased target wealth management products are more likely to purchase these fresh wealth management products, thus forming a collection of nearest-neighbor wealth management products, which is convenient for subsequent recommendations.

(3) Calculate the product recommendation list

According to the user's purchase situation, the corresponding score value is assigned, and the user's predicted score P_{ui} for different products is calculated. The calculation formula of the predicted score is as follows:

$$P_{ui} = \frac{\sum_{k \in S} sim(i,j) \times r_{uj}}{\sum_{k \in S} |sim(i,j)|} \tag{4}$$

Among them, a is the predicted score of product i by user u, S is the set of k nearest-neighbor financial products of financial product i, b is the similarity between product i and product j, c is the purchase of product j by user u, that is, corresponding ratings.

3. Experiment and result analysis

3.1. Data introduction

The experimental data in this paper comes from the real data of a financial platform, with a total of 13,130 original data. According to the user's purchase situation, the unrated, neutral attitude, purchased products, and holdings of products are given scores of 0, 1, 2, and 3 respectively, the products are coded 1-64, and the data is sorted and converted into a matrix , and finally get the data set required for the experiment. The data set includes 1806 pieces of financial product purchase record data and 64 pieces of user data, some of which are shown in Table 1:

Table 1: User's purchase information form

User's ID	Choice of financial products	preferred situation
1	Product 1	neutral
1	Product 6	Bought
3	Product 1	Bought
3	Product 88	Increase
5	Product 97	Bought

3.2. Experiment process

According to the data type of the data set, it can be found that there are no relevant attributes of the characteristics of financial products in the data set, and the number of financial products is far less than the number of users. Based on the above aspects, this experiment adopts the collaborative filtering algorithm based on wealth management products to provide users with recommendations for wealth management products.

According to the introduction of the item-based collaborative filtering algorithm process described above, this paper first cleans and processes the collected data set. The outliers are deleted or hidden, and other important outliers are filled. The data set is divided into training set and test set, and the similarity matrix w is obtained according to the Pearson similarity calculation formula.

Finally, this experiment recommends other financial products based on the historical purchases of each user, and obtains a recommended list of financial products.

3.3. Result analysis

After calculating and processing the data through MATLAB software, each user's predicted score for the product is obtained. Part of the score results are shown in Table 2:

Table 2: Partial Prediction Score Sheet

User's ID	Product 1	Product 2	Product 3
1	1.75552695	2.02588884	2.20413084
2	2.07858016	1.9546312	2.09142873
3	2.22463207	2.16953069	1.96501169
4	1.82814954	2.10820039	1.92828136
5	1.75552695	2.02588884	2.20413084

The accuracy of the recommendation result can represent the probability of matching the target user's favorite product in the recommendation list, and is mainly used to measure the accuracy of the recommended product. The calculation formula is as follows:

$$precision = \frac{T}{L} \tag{5}$$

Among them, T represents the number of products that satisfy the preferences of the target user, and L represents the final number of recommended products. The higher the accuracy, the better the recommendation result obtained by the algorithm.

The final calculation results show that the accuracy rate is 18.70%, indicating that this algorithm can play a certain role in the recommendation of financial products, and the recommendation results can meet the requirements of most users.

Acknowledgements

This work is supported by Innovation and Entrepreneurship Training Project of North China University of Science and Technology "Application of Time-Based Collaborative Filtering Algorithm in Securities Recommendation" (X2021326)

References

[1] Yang Zhigao: *Research on fund personalized recommendation model based on multi-strategy* (MS., Northeast Normal University, China 2016).

[2] Shi Hongli, Wang Congyu, Xie Weiqi: Research on the recommendation of educational platform courses based on collaborative filtering algorithm . Computer Knowledge and Technology, vol. 17 (2021), No.34.

[3] Linden, G, B. Smith and J. York, Amazon.com Recommendations: Item-to-Item Collaborative Filtering. IEEE Internet Computing. 2003. 7(1): p. 76-80.

[4] Ying Zhou: *Research on personalized recommendation based on user behavior analysis of securities wealth management products* (MS., University of Electronic Science and Technology of China, China 2014).

[5] Bian Yuning, Li Yeli, Zeng Qingtao, Sun Yanxiong. Application of improved collaborative filtering recommendation algorithm in precision marketing . Journal of Beijing Institute of Graphic Design, vol. 28 (2020), No.10, p 137-142.

[6] Li Qian: *Research on the recommendation algorithm of financial products based on machine learning* (MS., Beijing University of Technology, China 2019).

[7] Xu Huifang: *Research on text representation and feature extraction based on full coverage granular computing model* (MS., Taiyuan University of Technology, China 2016).

[8] Han Xiangbo. *Research on legal issues of appropriateness of financial product sales* (MS., China University of Political Science and Law, China 2011).

[9] Wang Haolong: The current situation, risks and solutions of Internet financial wealth management products. China Management Information, vol. 24 (2021), No.22, p 140-141.

[10] Zhang Yunxiang: *Research on the construction and application of the evaluation index system of bank wealth management products* (MS., Chongqing University of Technology, China 2021).

[11] Hu Shengjie. Research on the influencing factors of the use of Internet financial wealth management products, China Collective Economy, vol. 17 (2021), p 94-95.

[12] Yan Min: *Research on the Risks and Prevention Countermeasures of Internet Financial Financial Products in my country* (MS., Harbin Normal University, China 2019).