Application of Collaborative Filtering Recommendation Algorithm in pharmacy system

Yihang Lv a, Jingjing Kong
Chengdu Neusoft University, Chengdu, China

a Lvyihang@nsu.edu.cn

Abstract. Recommendation algorithm has been widely used in Internet commerce. At present, the most widely used and successful algorithm is collaborative recommendation algorithm. It can be used in pharmacy marketing system based on user and project purpose. Based on the analysis of the current situation of retail drugstores, this paper constructs the marketing management system of retail drugstores, and establishes two recommendation models of the marketing system of retail drugstores and retail drugstores. The main goal of model 1 is to construct a user based collaborative filtering recommendation algorithm input, after the algorithm processing, to generate recommendation results, to achieve product recommendation to customers. The goal of model 2 is to construct a Project-based Collaborative filtering recommendation algorithm input, after the algorithm processing, produce the recommendation results, so as to find the target customers for the product. These two recommendation models are helpful to the research and improvement of pharmacy system algorithm.

1. Introduction
The competition among pharmaceutical enterprises is becoming increasingly fierce. Some pharmaceutical enterprises, especially modern retail drugstores, are developing rapidly by using modern algorithm technology. In 2021, a new retail drugstore using Internet information technology has been added to China's securities market. Through the use of advanced mathematics and algorithm technology, we can optimize the system and get through the three lines of production, supply and marketing in the enterprise. Therefore, in order to effectively solve these problems, it is necessary to effectively think and design the retail pharmacy system, and further explore the system algorithm.

When customers enter the drugstore outlets, websites or mobile applications, they should be prompted about the products they may need. User based collaborative filtering recommendation algorithm is applied to recommend products to customers, which mainly considers recommending products to the same customer group or the nearest neighbor of the customer group.

When a customer manager looks for a target customer for a product, the system displays the customers of similar products to the customer manager. The customer manager carries out in-depth marketing according to the list, which has strong pertinence. The Project-based Collaborative filtering recommendation algorithm is suitable for finding target customers for products. It mainly considers the definition of target customer group attributes during product development, or recommendation to the nearest neighbor customer group of new products.
2. The analysis of system structure optimization design

The innovation of the retail pharmacy marketing management system designed in this paper is: using JMS service message to solve the complex processing process of server-side message. The design of the system is mainly to further solve the problem of low efficiency in retail pharmacy industry, and solve practical problems through information management system. The innovation of this paper is based on the B/S structure mode, through the Internet service design to research and develop the management system, the system has a certain scalability, can meet the needs of internal and external office, on the basis of cost saving to meet the actual development needs of the company.

The structure of the point-to-point pattern is shown in Figure 1.

![Point to Point mode structure](image)

**Figure 1.** Point to Point mode structure

Before logging in to the system, the user must connect with the server and register in the client. The system will listen to the corresponding messages sent, and the server will access all functions, so as to realize the connection between the client and the server. When users transfer information, JMS will transfer data to the receiver through data queue. On the client side, the main function is to upload data to the server. In addition, when backing up the data, we only need to store the words simply, which will reduce the occurrence of security incidents and reduce the resources occupied by the server. In time also does not require synchronization, data transmission can be transferred from one to multiple, effectively improve the efficiency of the system, and maximize the simplification of data analysis in the server.

3. Marketing system model design

3.1. User based collaborative filtering recommendation model

The main goal of this model is to construct a user based collaborative filtering recommendation algorithm input, after algorithm processing, to produce recommendation results: recommend products to customers.

3.1.1 Input. In a typical algorithm of this kind, the input data can usually be expressed as an M×N user and project evaluation matrix R. m is the number of users, n is the number of projects, and Rij is the evaluation value of the i-th user on the j-th project. The evaluation value can be the user preference (like / don't like) or purchase status (purchased / not purchased) expressed by binary 0 and 1, the user preference value for the item expressed by rating, and the attribute value of the item itself.

In this model, the customer group of the pharmacy is the user, and the customer attributes (including whether to buy a product) are regarded as items, which constitute an input matrix (see Table 1). In order to facilitate the calculation, the values of the representative range in the input matrix must be single valued. For example, the age attribute sets 0-5 years old as 1, 6-12 years old as 2, and so on.
3.1.2 Handling. Generally speaking, there are two different ways to determine the nearest neighbor set: first, according to the preset similarity threshold, select neighbors whose neighbors are not less than the threshold value; and the second one selects the top \( N \) users with the most similarity as the neighbor users according to the number of neighbors determined by \( N \). Therefore, the core task of this model is to calculate the similarity between customer groups.

Table 1. User / project evaluation matrix

| Customer group | Age group | Income range | Purchase \( j \) products (yes / no) | Purchase \( n \) products (yes / no) |
|----------------|-----------|-------------|-------------------------------------|-------------------------------------|
| C1             | R11       | R12         | R1j                                 | R1n                                 |
| C2             | R21       | R22         | R2j                                 | R2n                                 |
| ...            | ...       | ...         | ...                                 | ...                                 |
| Ci             | Ri1       | Ri2         | Rij                                 | Rin                                 |
| ...            | ...       | ...         | ...                                 | ...                                 |
| Cm             | Rm1       | Rm2         | Rmj                                 | Rmn                                 |

The similarity measure of this model adopts the distance measure method. The customer group \( C_i \) is regarded as a point in the \( n \)-dimensional space. The similarity between the other customer group \( C_j \) and \( C_i \) depends on the distance between the two points in the \( n \)-dimensional space. The smaller the distance, the higher the similarity. According to the values provided in Table 1 of the input matrix, considering that different attributes of customers have different effects on customers' choice of products, the corresponding weight can be set for each attribute. If the weight corresponding to attribute \( \{ P_1, P_2, \ldots, P_n \} \) is set as \( \{ W_1, W_2, \ldots, W_n \} \), the distance calculation formula of \( C_i \) and \( C_j \) is as follows.

\[
d = \sqrt{W_1(R_{ij} - R_{ii})^2 + W_2(R_{jj} - R_{ii})^2 + \ldots + W_n(R_{nn} - R_{ii})^2}
\]

The distance between customer group \( C_i \) and \( M \) customer groups is calculated by the above formula, and a distance set \( D = \{ D_1, D_2, \ldots, D_M \} \) is obtained. According to the data in set \( D \), the customer group whose distance is less than a certain threshold value or the first \( n \) customer groups whose distance is the smallest is selected to form the nearest neighbor set \( n = \{ C_1, C_2, \ldots, C_n \} \) of customer group \( C_i \).

3.1.3 Generate recommended results. The products recently purchased by neighbors and not purchased by customer group \( C_i \) are recommended to customer group \( C_i \). Let \( P_1 \) be the product set purchased by the customer group, \( P_2 \) be the product set purchased by the nearest neighbor, and the product set recommended to the customer group is the difference set between \( P_2 \) and \( P_1 \): \( P = p_2 - p_1 \).

3.2 Project based collaborative filtering recommendation model

The goal of this model is to construct a project based collaborative filtering recommendation algorithm and generate the recommendation result through algorithm processing, which is to find the target customers for the product.
3.2.1 Input. The input of the Project-based Collaborative filtering recommendation algorithm is divided into two parts. One part is the \( m \times n \) matrix representing the characteristics of the project, which is composed of the characteristics of \( M \) projects, and each project has \( n \) characteristics; the other part is the score value of a user to the project. According to the similarity of the item features, the user's rating value of the ungraded item is estimated.

In the retail pharmacy marketing system, the product attributes are regarded as the characteristics of the project, and whether the user purchases or not depends on the user's score of the project. Therefore, the input of this model includes an \( M \times n \)-dimensional product attribute matrix and an \( n \times m \)-dimensional matrix of whether the user purchases the product (1 for purchase and 0 for non purchase). In order to facilitate calculation, the project properties must be single valued, and the inapplicable properties must be assigned 0.

3.2.2 Handling. There are two tasks in this model: the first task is to calculate the nearest neighbor set of a product \( I_i \), and the calculation method is similar to the first model in this paper; the second task is to find out the customers of the nearest neighbor set of product \( I_i \).

The nearest neighbor set of product \( I_i \) is also determined by similarity measurement, which selects the nearest neighbor set whose correlation is greater than a certain threshold or the first \( n \) nearest neighbors. The similarity measure also adopts the distance measure method. According to the above formula, the distance between product \( I_i \) and other products is calculated, and a distance set \( D = \{D_1, D_2, ..., D_m\} \) is obtained. The smaller the distance is, the greater the similarity is. The nearest neighbor set \( n = \{N_1, N_2, ..., N_n\} \) of product \( I_i \) is formed by selecting \( n \) products with the smallest distance or neighbors whose distance is less than a certain threshold.

According to the matrix of products purchased by customers, the nearest neighbor set \( n \) of products is projected on the matrix to form a matrix of products purchased by customers (as shown in Table 2). In the matrix, 1 means that customers have purchased products, 0 means that they have not purchased products, all rows and columns of "0" in the matrix are removed, and the remaining customers in the matrix are customers of nearest neighbor set.

| Customer / neighbor products | N1 | N2 | ... | Nj | ... | Nn |
|------------------------------|----|----|-----|----|-----|----|
| C1                           | 1  | 0  |     | 1  |     | 0  |
| C2                           | 0  | 0  |     | 0  |     | 0  |
| ...                          |    |    |     |    |     |    |
| C_i                          | 0  | 0  |     | 1  |     | 1  |
| ...                          |    |    |     |    |     |    |
| C_m                          | 1  | 0  |     | 0  |     | 1  |

3.2.3 Make recommendations. In the matrix of customers purchasing nearest neighbor products after removing all "0" rows and columns, all customers can be the target customers of product \( I_i \).

3.3. Some problems of system algorithm

In general, collaborative filtering recommendation algorithm has matrix sparsity problem, cold start problem and special preference problem. In the pharmacy marketing system, each attribute exists objectively, so the probability of these three problems is small. However, each attribute has different influence on the recommendation result, so each attribute is given weight. The initial weight depends on the subjective judgment of the model operator, which may lead to large error in the recommendation result. If we use the correlation between customer or product attributes to determine the weight of each
attribute, the greater the correlation, the greater the weight of the attribute should be, then the recommendation algorithm will achieve good results.

4. Conclusion
With the continuous emergence and application of various kinds of computing science and technology, the informatization level of various industries is also improving. Based on the analysis of the current situation and demand of retail drugstores, this paper constructs a retail drugstore marketing management system, which corresponds the two requirements of the drugstore marketing system to the algorithm respectively, and establishes two recommendation models, which is helpful for retail drugstores to realize precision marketing and cross marketing.

Acknowledgments
Mixed curriculum reform project of Chengdu Neusoft University -- "marketing"

References
[1] Fang Gao, Gao Fang. Optimization Layout of Marketing Management Information System of Hotel and Restaurant Based on B / S Mode[J]. Journal of Physics: Conference Series, 2020, 1578(1).
[2] Gao Jing, Li Sihui. Design and implementation of video forum website based on B / S mode [J]. Southern agricultural machinery, 2020, 51 (23): 162 + 169
[3] Shi Shengyuan, Wang Lin. design and development of experimental training management system of production operation management based on B / S mode [J]. Public standardization, 2020 (21): 222-223
[4] Li Yimeng. Design and implementation of enterprise marketing management system [D]. Shandong University, 2019
[5] He Qingchun. Application of Collaborative Filtering Recommendation Algorithm in bank marketing system [J]. Electronic finance, 2014 (03): 80-81