ITESCONF

World Conference on Innovation in Technology and Engineering Sciences

03 - 05 December 2021 Athens, Greece

Comparison of different approaches for the recommending suitable products in e-shop recommender system

Bogdan Walek[1] * and Petr Fajmon[1]

[1]Department of Informatics and Computers, University of Ostrava Ostrava, Czech Republic

Abstract.

This article compares different approaches for recommending suitable products in the e-shop recommender system. The proposed recommender system consists of two main modules - a module for recommending suitable products based on viewed products and a module recommending suitable products based on the rated products. The first module uses a content-based filtering approach, and the TF-IDF algorithm is used for product recommendations. The second module uses a collaborative filtering approach, and the SVD algorithm is used for product recommendations. Furthermore, three approaches that combine the results of both modules are proposed. These approaches were experimentally verified on a group of 32 real users. The users tested the proposed recommender system and marked the relevant products from the list of recommended products. The results of the experimental verification are discussed.

Keywords: recommender system, e-shop recommender system, hybrid recommender system, content-based filtering, collaborative filtering, products

1. Introduction

Currently, various recommender systems are very popular. The main goal of recommender systems is to recommend appropriate content to the user based on his preference and behavior.

An example of a recommender system is the Amazon.com online store or the Youtube video portal. A recommender system is an information system that is used to support user decisions and recommend suitable products, information, or services in the environment of online stores, streaming services, online dating, and many other industries [1].

In the area of online stores, the main goal is to recommend relevant products that are interesting to the customer. These products have a relatively high probability that the customer will buy them.

Representatives of the largest online stores that implement recommender systems are Amazon.com, eBay.com, or Taobao.com. Products are most often recommended to users based on the popularity of the product, the customer's demographic data, and an analysis of the customer's behavior during previous purchases in the online store [2].

2. E-commerce/e-shopping recommender systems

Online stores are an important area of the e-commerce industry. In online stores, the most common element for determining the popularity of a product is its rating. The rating is very often used in the form of stars, and the user thus gives the e-shop feedback on the specific purchased product. Typically, a scale from 1 to 5 is used for this. Product rating can then be used to make recommendations and is a frequent input to recommender systems.

There are a number of implemented and published systems in the field of e-commerce/e-shopping recommender systems.

One of these systems is the Wasabi Personal Shopper (WPS). It is a domain-independent tool for browsing electronic product catalogs [3]. Cao and Li have developed a fuzzy recommendation system for products made from various components [4]. Mooney and Roy proposed a content-based book recommender system using information extraction and a machine learning algorithm for text categorization [5].

Users often want to be informed about what suitable product is available to them or why a particular product has been recommended by a recommender system. To provide a relevant explanation of why a product is better than another, McCarthy et al. developed a shopping assistant website called Qwikshop.com, on which compound criteria were used as explanations [6]. Another area is the sale of product packages (bundle of items) or promotional packages (bundle promotion). These multi-product product packages and promotional packages are often offered by online stores because they bring savings to marketers, and they are beneficial to the customers. In the system developed by Garfinkel et al. the product search method was extended to recommend suitable product packages and promotional packages [7].

A recommender system working with product-seeded and basket-seeded scenarios has been proposed for the area of small online stores [8]. During shopping in an online store, users often decide to buy the product based on their basic needs and relative needs. The authors of Tareq et al. designed a model of dynamic recommendation system (DRS) for the online market [9].

E-shopping recommender systems are suitable for various types of products, whether they are digital products (music, movies, etc.) or physical products (electronics, books, food, etc.). Many different recommender systems have been developed in this area, which has been successfully validated in online stores and can be implemented by online store developers.

3. Recommender system

This section will describe the proposed recommender system. The aim of our proposed system is to recommend online store products for the user based on the viewed and rated products. The proposed recommender system is connected to an online store with foodstuff that contains real products.

It is a hybrid recommendation system that contains two modules for the recommendation of suitable products:

- Recommendations based on viewed products (content-based filtering approach)
- Recommendations based on rated products (collaborative filtering approach)

The modules of the proposed system will be described in the following subsections. First, we describe the user's work in the online store, which is connected to the recommender system.

1.1 Registration to the system and selection of favorite product categories

First, the user creates a registration in the online store. After registration, the system will prompt the active user to select favorite categories, based on which he will select similar users and all their rated products. The system sorts the selected products according to the number of ratings, the average rating and delivers the first group of proposed products to the user.

The active user selects these categories:

- Fish
- Fruit
- Drinks
- Seasoning
- Alcoholic beverages

The first recommended products will contain mostly products from these categories, which have been rated by similar users to the active user.

1.2 Viewing products

Similar products are often recommended to users in various e-shops when viewing products. The system recommends to the user the most similar products to the viewed product. However, it also records the history of the products viewed, as well as whether the products have been rated by the user. In our recommender system, a content-based filtering approach will be used to recommend similar products based on the products viewed by the user.

The content-based filtering approach generally consists of several components: a) preprocessing and feature extraction, b) content-based learning of user profiles, c) filtering and recommendation [1, 10]. The TF-IDF and LDA algorithms [1] are the most widely used within the preprocessing and feature extraction component. The TF-IDF algorithm was chosen for our recommender system due to its popularity, simpler implementation and lower requirements for system resources.

As mentioned above, the system records user behavior on the e-shop and stores the browsing history of products. In our system, the CBF approach applies a modified TF-IDF algorithm to the viewed products, calculates the most similar products, and delivers the top N products to the active user. The process of the CBF approach is shown in Fig. 1.

Figure 1: Process of recommendation based on product viewing (CBF approach)

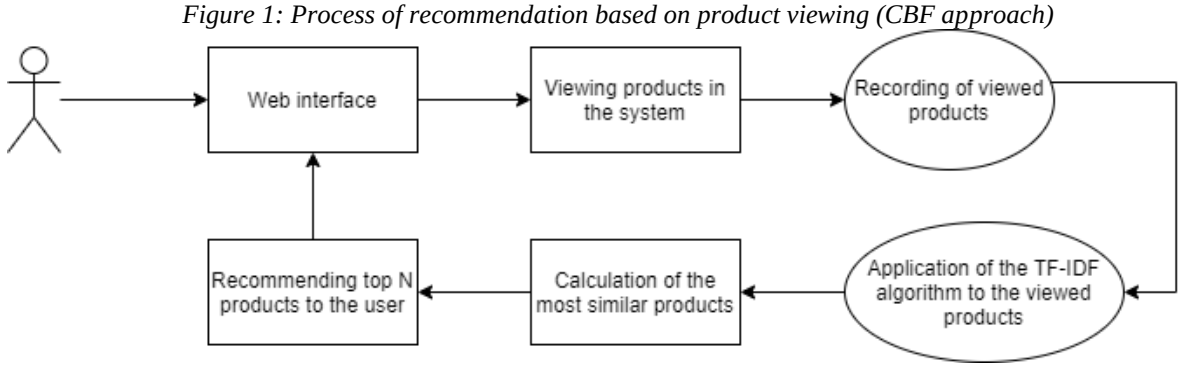

Table 1: Viewed products of the active user

Product	Active user rating	Category
Strawberries	4	Fruit
White pepper	Not rated	Vegetables
Celery	Not rated	Vegetables
Radish	Not rated	Vegetables
Pineapple	5	Fruit
Tomato	Not rated	Vegetables
Grapefruit	Not rated	Fruit
Melon	Not rated	Fruit

In Table 1, we can see the products that the user has viewed. The group of viewed products contains eight products from the categories "Fruits" and "Vegetables", the table also contains information about the rating of an active user.

The list of similar recommended products to the viewed products is shown in Table 2.

Table 2: A list of recommended products based on product browsing

Product	Similarity	Average rating	Category
Pear	1.0000	4	Fruit
Kiwi	0.8798	4.2	Fruit
Plum	0.7727	3.6	Fruit
Lemon	0.7339	3.6	Fruit
Yellow pepper	0.7127	4	Vegetables
Red pepper	0.7127	3.9	Vegetables
Carrot	0.7043	4.3	Vegetables
Grapes	0.6355	4.2	Fruit
Parsley	0.6287	4	Vegetables
Garden mushrooms	0.5457	3.9	Mushrooms
Oregano	0.5416	4	Spice
Kohlrabi	0.5344	4.1	Vegetables
Coconut	0.5278	4	Fruit
Portobello mushrooms	0.5276	3.9	Mushrooms
Lime	0.5254	3.9	Fruit
Garlic	0.5181	3.9	Vegetables
Sweet paprika	0.5130	3.5	Spice

Dill	0.5129	3.8	Spice
Celery with nati	0.5076	3.9	Vegetables
Bulgur	0.5063	4.2	Cereals

In Table 2, we can see the TOP 20 recommended products based on user's viewed products. Not rated products by the user have a lower impact on the resulting data. The user rated only two products from one category (Fruit) from the viewed products. So these products have a normal effect on product recommendations. For this reason, products from the Fruit category are the top recommender products. The result of the CBF approach is the first input parameter for subsequent product recommendation approaches, which were experimentally verified.

1.3 Product rating

Another part of the proposed system is the recommendation according to the similarity of already rated products of the active user. The user can also rate products in our system. The rating of products is important for the next part of the recommendation.

This part of the proposed recommender system uses a collaborative-filtering approach to recommend similar products. For this approach, a user-item matrix of similar users is used, who rated the same products as the active user. The SVD algorithm is then applied to this user-item matrix, which is described in more detail here [1]. The system selects all data from the database, creates a user-item matrix, and applies the SVD algorithm. The system then selects the top N most suitable products.

The process of the collaborative-filtering approach is shown in Figure 2.

Figure 2: Process of recommendation based on product rating (CF approach)

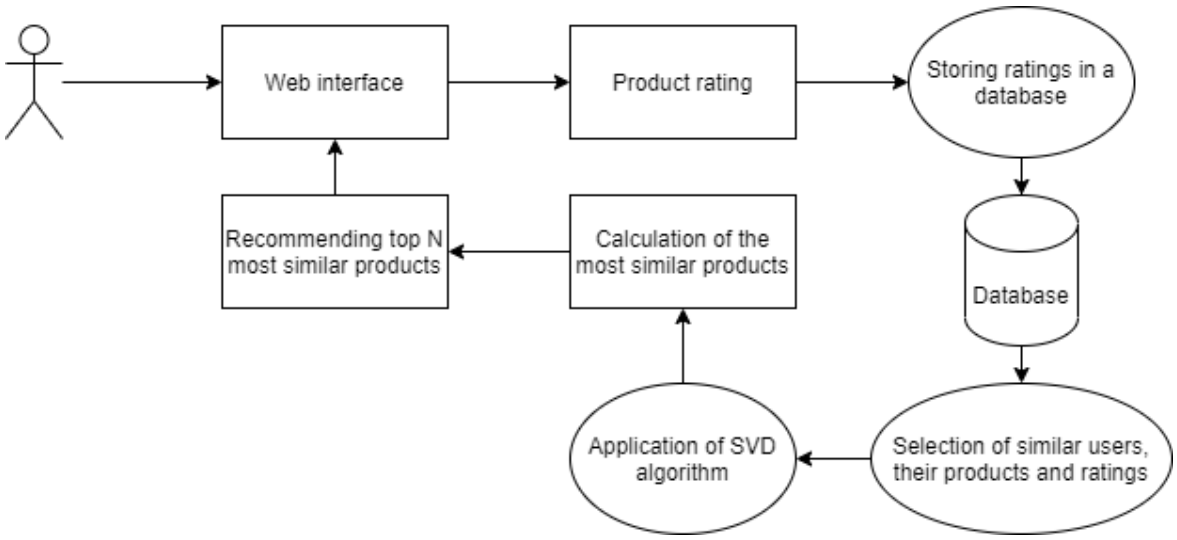

Table 3 shows the ratings of products for the active user.

Table 3: Product rating by an active user

Product	Product rating
Bio Bulgur	4
Božkov Republica	4
Bulgur	5
Sunflower bun	4
Caleo Primitivo Salento	5
Fillet with skin	2
Halibut filet	2
Bun cereal	5
Curry	2
Cheesecake	5
Grilled chicken	1
Pie with pudding filling	3
Pavolín Pálava	5
Pellegrino Marsala	5

The active user rated 14 products from various categories. The user rated products from categories Cereals, Alcoholic beverages and Pastries in a high rating. On the other hand, the user rated the products from the category Seasoning and Fish with a low rating. After performing the SVD algorithm, the system should recommend products similar to the highly-rated ones. Low-rated products should not be recommended by the system. Table 4 shows the recommended products using the SVD algorithm applied to the rated products by active user.

Table 4: List of recommended products based on the product rating

Product	Number of ratings	Average rating	Category
Quinoa	29	4	Cereals
Cheese snail	31	4	Pastry
Poppy seed cake	29	3.9	Pastry
Moët & Chandon Imperial	24	4.1	Alcoholic beverages
Chardonnay Pays	25	4.1	Alcoholic beverages
Barley groats	27	4.1	Cereals
Multigrain bread	31	3.8	Pastry
Dark lump	28	3.8	Pastry
Organic wheat	29	3.5	Cereals
Becherovka	29	4	Alcoholic beverages
Organic Millet	30	4.1	Cereals
Straw wine	24	3.9	Alcoholic beverages
Plait	31	3.7	Pastry
Wheat bulgur	27	4	Cereals
Malibu	30	4.1	Alcoholic beverages
Jack Daniels Honey	37	4.2	Alcoholic beverages
Hulled buckwheat	31	3.4	Cereals
Tullamore Dew	32	3.9	Alcoholic beverages
Whiskey	28	3.9	Alcoholic beverages
Bun natural	30	3.8	Pastry

Table 4 shows the TOP 20 recommended products based on user's rated products. We can see that all products have a high rating and a lot of products are from the categories that the user rated with a high rating. Products from the category Fish and Seasoning, which the user rated with a low rating, are not in the resulting list. The result of the CF approach is the second input parameter for subsequent product recommendation approaches, which were experimentally verified.

4. Results

For the system verification, we used a group of users who performed all the necessary steps in the proposed (registration, selection of favorite categories, viewing products, product rating). Also, three approaches were proposed for system verification.

4.1 Approaches for the system verification

All approaches use a combination of outputs of the CF approach (similar to rated products) and the CBF approach (similar to viewed products). The resulting list shown to the user consists of a maximum of 40 products, and the top products of the given outputs are always recommended and sorted by similarity. Each approach represents a different ratio of share to individual outputs and always the corresponding number of products. If we take the ratio 30% / 70%, the resulting list will contain 30% of the most suitable products from the results similar to those viewed, applying the TF-IDF algorithm to the viewed products, and 70% of the most suitable products from the results of similar already rated products, applying the SVD algorithm to already rated products by an active user. The approaches are proposed as:

- Approach A: 30% TF-IDF / 70% SVD
- Approach B: 50% TF-IDF / 50% SVD
- Approach C: 70% TF-IDF / 30% SVD

For verification of our proposed system, a group of 32 users was selected. Users performed these steps:

- Registration to the system
- Selection of 5 favorite product categories
- Viewing 40 products
- Rating 40 products

Then the system shows all recommended products in all approaches. The task for users was to mark as “relevant” these products which are relevant to their preferences and behavior in an e-shop. And also mark as “no relevant” these products which are not relevant to their preferences and behavior. Results of the ratio of relevant products for each user as shown in Table 5.

Table 5: Results of verification

User ID	The ratio of relevant products Approach A	The ratio of relevant products Approach B	The ratio of relevant products Approach C
1	100%	100%	100%
2	100%	100%	100%
3	100%	100%	100%
4	100%	100%	100%
5	100%	100%	100%
6	100%	97%	97%
7	100%	100%	94%
8	97%	97%	97%
9	97%	97%	87%
10	97%	91%	92%
11	100%	97%	83%
12	93%	91%	94%
13	100%	89%	84%
14	90%	90%	90%
15	92%	90%	82%
16	81%	89%	92%
17	89%	85%	88%
18	83%	76%	97%
19	88%	88%	78%
20	86%	85%	82%
21	79%	85%	87%
22	95%	79%	76%
23	81%	83%	81%
24	89%	78%	69%
25	82%	77%	74%
26	77%	79%	77%
27	75%	75%	75%
28	62%	71%	76%
29	65%	71%	71%
30	38%	64%	80%
31	56%	54%	61%
32	53%	50%	44%
Average	**86%**	**85%**	**85%**

Table 5 shows in the first column the ID of the user who tested our proposed system. The next columns show the ratio of products that the user marked as relevant to all products that were recommended to the user by the system. It is the ratio of relevant products to all recommended products. The results are sorted in the table according to the share of relevant products for individual users (first, the table includes the users who marked the most products as relevant). The table shows that the best results are achieved by Approach A (13 users identified 95% and more relevant products). This is followed by Approach B (10 users identified 95% or more of relevant products) and then Approach C (8 users identified 95% or

more of relevant products). On another side, only a few users marked less than 60% of products as relevant. In Approach A are three users with less than 60% (38%, 56% a 53%), in Approach B two users (54% a 50%), and in Group C one user (44%). So we can conclude that this approach has promising results – the average ratio of relevant products is 86% in Approach A, 85% in Approach B, and 85% in Approach C.

From this discussion, we can conclude that the best results are in Approach A:

- The highest number of users with marked relevant products with the ratio of 95% and higher (13 users)
- The highest number of users with marked relevant products with the ratio of 100% (9 users)
- The highest average ratio – 86%

Based on experimental verification by a group of users, Approach A brings the best results. In this approach 30% of products were recommended using the CBF approach (TF-IDF algorithm), and 70% of products were recommended using the CF approach (SVD algorithm). So we can see that a higher ratio of recommended products based on rated products (SVD algorithm) recommends more relevant products to the users (in our verification). In other approaches, the ratio was lower (Approach B – 50% TF-IDF and 50% SVD, Approach C – 70% TF-IDF and 30% SVD).

5. Conclusion

In this paper, the recommender system for recommending relevant products in e-shop was proposed. The proposed recommender system consists of two main modules - a module for recommending suitable products based on viewed products and a module recommending suitable products based on the rated products. The first module uses a content-based filtering approach, and the TF-IDF algorithm is used for product recommendations. The second module uses a collaborative filtering approach, and the SVD algorithm is used for product recommendations.

Also, three approaches for the verification of our system were proposed. The approaches use a different ratio of products recommended by the TF-IDF algorithm and SVD algorithm. These approaches were verified by a group of 32 users. The task for users was to mark as “relevant” these products which are relevant to their preferences and behavior in an e-shop. And also mark as “no relevant” these products which are not relevant to their preferences and behavior. The results of verification showed that the best results had been achieved in Approach A (30% products recommended by TF-IDF algorithm and 70% products recommended by SVD algorithm). The results of verification were also shown in detail and discussed.

In future work, we would like to focus on the verification of our proposed system on other groups of users. Also, we would like to focus on improving modules for the recommendation of products (adding possibilities of marking not favorite categories, not favorite products, etc.).

Acknowledgment

This work was created during the completion of a Student Grant SGS20/PřF-MF/2021 with student participation, supported by the Czech Ministry of Education, Youth and Sports

References

[1] Falk, K. (2019). Practical recommender systems (1st ed.). Shelter Island: Manning ISBN 1-61729-270-2 .

[2] Schafer, J. B., Konstan, J., & Riedl, J. (1999, November). Recommender systems in e-commerce. In Proceedings of the 1st ACM conference on Electronic commerce (pp. 158-166).

[3] Burke, R. (1999, July). The wasabi personal shopper: A case-based recommender system. In AAAI/IAAI (pp. 844-849).

[4] Cao, Y., & Li, Y. (2007). An intelligent fuzzy-based recommendation system for consumer electronic products. Expert Systems with Applications, 33(1), 230-240.

[5] Mooney, R. J., & Roy, L. (2000, June). Content-based book recommending using learning for text categorization. In Proceedings of the fifth ACM conference on Digital libraries (pp. 195-204).

[6] McCarthy, K., Reilly, J., McGinty, L., & Smyth, B. (2004, August). Thinking positively-explanatory feedback for conversational recommender systems. In Proceedings of the European Conference on Case-Based Reasoning (ECCBR-04) Explanation Workshop (pp. 115-124).

[7] Garfinkel, R., Gopal, R., Tripathi, A., & Yin, F. (2006). Design of a shopbot and recommender system for bundle purchases. Decision Support Systems, 42(3), 1974-1986.

[8] Kaminskas, M., Bridge, D., Foping, F., & Roche, D. (2017). Product-seeded and basket-seeded recommendations for small-scale retailers. Journal on Data Semantics, 6(1), 3-14.

[9] Tareq, S. U., Noor, M. H., & Bepery, C. (2020). Framework of dynamic recommendation system for e-shopping. International Journal of Information Technology, 12(1), 135-140.

[10] Aggarwal, C. C. (2016). Recommender systems : 1. Cham: Springer International Publishing.