2014년 추계학술발표대회 논문집 제21권 제2호 (2014. 11)

상품 가격, 구매자 평가, 판매량에 관한 개인별 선호도에 기반한 구매 추천 기법

김병민, 사웃 알고와이자니, 한경숙
인하대학교 컴퓨터공학과
e-mail: khan@inha.ac.kr

A recommendation method based on personal preferences regarding the price, rating and selling of products

Byungmin Kim, Saud Alguwaizani, Kyungsook Han
Department of Computer Science and Engineering, Inha University

Abstract

Recently several recommender systems have been developed in a variety of applications, but providing accurate recommendations that match the preferences and constraints of various users is quite challenging. This paper presents a method of recommending digital products based on the past preference of a user on the price, rating and selling volume of a product. Experimental results of the method with actual data of Amazon showed that the average accuracy of the recommendations made by the method is 85%. Although the results are preliminary, the method is potentially capable of making more accurate personalized recommendations than existing methods.

1. Introduction

There are several studies related to the demonstration of recommendation systems in online social networks (OSNS). Discovering the exact tastes of consumers in order to recommend suitable items is the key point of the recommendation systems. There had been many efforts and researches through the interactions between the vast amounts of products and consumer information to find the best match to their preferences.

Utility-based recommendation systems which are widely used these days in online shopping mall require the consumers to type in their preferences into the system in order to compile a list of recommendations [1]. This can be very helpful for reducing the start-up problem for the cold users and, no product purchase history of the consumer is needed. However, high accuracy of preferences depends on how high effort consumers put into the preference survey. Another method of recommendation called collaborative filtering which is also popular through many researches considers consumers with similar purchase history are assumed as consumers with similar preference. For example, if the consumer *a* and consumer *b* have similar preference, the product *p* bought by *b* is recommendable for *a* [2]. However, it is not clear how collaborative filtering provides the recommendation of a specific product.

Recently, in order to overcome limitations of these recommendation systems, The Ranked Pareto-Front (RPF) was introduced [3]. This method focuses on ranking the products by using selected features from dataset and recommend to the users who prefer to purchase a product with a better utility than choice-based conjoint analysis. However, it is unable to test all conceivable decision strategies and their combinations. The model itself is designed to be adapted to further strategies.

In this study, we propose a simple method for making recommendations from the analysis of personal preferences regarding the price, rating and selling volume of a product. We analyzed the user information on purchasing for electronics at Amazon.com and developed a method for making personalized recommendations based on key features of user preferences. We tested the method on a new dataset which was not used in developing the method. Experimental results of testing the method on a new independent dataset showed that the accuracy of the recommendations made by the method is as high as 85.1%. The rest of this paper will discuss the method and its experimental results.

2. Data Acquisition and Methods

2.1. Dataset

As for the first step, we have obtained Amazon electronics purchase list from Stanford Network Analysis Platform (SNAP). The dataset consists of product ID, user ID, price, rate, review time, and review summary.

In this study, we used product ID, user ID, price, and rate information to analyze 839 digital camera buyers' information at Amazon.com. After that, we have categorized all items in Amazon.com electronics and ranked them based on our method with selected features.

2.2. Personal Preference

Reading consumer's mind is one of the most important factors in recommendation systems. Without the accurate preferences of users, it is impossible to recommend the exact product that users want, even if the recommendation system itself is good. We have made the simple method to discover the personal preferences. In order to do that, we have analyzed the purchase history to find out the relationship between consumers and their purchase products. From that information, we considered price, rate and popularity of products, which are the key features in this study to rank all existing products based on our method to discover the user preference.

First of all, it is obvious that the price is very important element when purchasing an item. As shown in Table 1, we divided price into three parts which are cheap, medium price and expensive. Customers might want to buy cheap products and some prefer expensive products because that might guarantee the quality of product. Before we set the standard for price information, we have arranged the existing product list in descending order and then considered products within top 30% as cheap, bottom 30% as expensive and middle price for the rest of the products.

Secondly, rate is another important component that consumers put big emphasis on, since the quality is already proven by users who purchased that item. The majority of earlier work in Rating Prediction and Recommendation of products (e.g. Collaborative Filtering) mainly takes into account ratings of users on products [4]. In Amazon online, users can rate their purchased items from between minimum 1 to maximum 5. Each product has its own rates given by consumers and we have used the average rate for each product as the feature.

Finally, the selling volume of a product reflects its popularity since people in general want to buy best-selling products.

With the standard setting of three features, we have ranked all of the existing products in Amazon electronics after dividing them into categories. We decided to call that, product ranked list (PRL). PRL is stored in the database so that we could easily update them. Furthermore, we have extracted purchase list of each user to see what kind of items (s)he has purchased so that it could be easier to find out the percentages of the three features. In order to do that, we compared purchased items from a user and compared with PRL to see how each item is ranked in its category. Figure 1 shows the algorithm of finding the preferences of users by using their purchase history.

<Table 1> Standard settings of features for equating user's preference

Price			Customer rating	Best-selling
Low	Medium	High		
30%	40%	30%	Top 20%	Top 30%

<Figure 1> Computing preferences of user

Algorithm 1. Personal preferences

for *i* = 0 to *n* **do**
All digital camera customers.
for *j*=0 to *m* do
Products from purchase history of each customer.
if (Low price) **then** Cheap Products ++;
else if (High price) **then** Expensive Products **Price**
else Medium Price Products ++;
if (High Rated){High Rated Products ++;} **Rating**
if (Best selling){Best selling Products ++;} **Best selling**

For example, let' s say user *x* has 70% products ranked within 30% of cheap products and 60% of good rated products ranked in 20% of good rated products in the purchase history. From this, we can realize that *x* gives considers cheap priced items and good rated products when purchasing an item. Table 2 shows the sample values of user preferences that help us easier to know what users exactly prefer. With those values, we could infer the exact preferences shown in table 3.

<Table 2> Sample of user preferences

User	Price (%)			Rating (%)	Selling (%)
	Low	Medium	High		
A	0	42.9	57.1	35.7	64.3
B	0	0	100	100	100
C	66.7	0	33.3	33.3	100
D	75	25	0	25	75
E	60	40	0	20	20
F	14.3	85.7	0	57.1	85.7
G	25	50	25	75	25

<Table 3> Inferred user preferences from Table 2

User	Price	Rating	Selling
A	Cheap	None	Yes
B	Expensive	Yes	Yes
C	Cheap	None	Yes
D	Cheap	None	Yes
E	Cheap	None	None
F	Middle	None	Yes
G	Middle	Yes	None

Since the rate information has subjective assessment from consumers, it might have fake information. Users might give bad rate to a product because of the late shipping even though the product itself is good. On the other hand, other two features which are price and popularity are less subjective compared to the rate information. With the combination of these three features, the disadvantage of rate can be ignorable. Also, our way becomes very strict in getting the user preference and product rank, making recommendation much more accurate.

2.3. Recommendation

Recommendation systems aim to select products for a particular user from a list shared by all (available products for instance) according to known previous preferences of users [5]. If the preference of user is not clear, the recommendation systems become useless. Our purpose is that, when a user clicks on an item, the system should automatically recommend the products that match the user' s preference from that category. Once we found the accurate inferred preference of users, it is very easy to recommend the appropriate items to that user. In this study, we have used digital camera consumers' data to test our method but since all of the products' rank lists for electronics categories are included in PRL database, we can find the user preferences and use the recommendation for other types of items too. There are steps for our recommendation method:

First step, we need to know the exact preference of a user to see what the most important thing that (s)he is considering. In table 2, we can see that user *A* considers popularity and expensive price the most since the percentages are over 50%.

Second step, we need to choose items that match the preference from PRL, since it contains the whole ranking list of products. Thus, we need to make sure we must choose only one (cheap, middle, or expensive) from the price feature.

Third step, we have to consider unusual events. For example, a user has equal number of items that are both cheap and expensive in purchase history. We have to recommend cheap and expensive products at the same time, which doesn' t make sense. In this case, we decided to think that user does not care about the price.

Finally, with the user preferences that were made previously, we constructed the recommendation list for each user by checking the PRL database.

3. Result and Discussion

In this study, we used digital camera consumers' purchase history from Amazon electronics. The total number of digital camera consumers is 839 and there are 59 types of digital cameras in the dataset. To test how accurate our method was, we got the preferences of digital camera users by using our method.

First of all, In order to evaluate the performance of the recommendation, we have deleted all digital camera data from the purchase lists of consumers. Secondly, we computed the preferences of those users and recommended appropriate digital cameras depending on their preferences. Thirdly, we compared the recommended list of digital cameras with the actual purchased list and if the same camera existed in the list, we considered as correctly recommended (1) and if not, we considered as incorrectly recommended (0). Finally, we have calculated the accuracy by dividing the total number of users by the number of correct recommendations.

Table 4 shows the results of the recommendation for inferred preferences of some users shown in Table 3. We have considered price information all the time unless when the price preference had cheap and expensive at the same time Table 5 shows the total result for the recommendation. We have recommended 59 types of cameras to 839 users. As a result, we have correctly recommended the camera lists to 714 users out of 839, which is 85.1% of accuracy.

<Table 4> Result of recommendation system

User ID	Price	Rating	Selling	Result
A	High	None	Popular	1
B	High	High Rated	Popular	1
C	Low	None	None	0
D	High	None	Popular	0
E	Medium	None	None	1
F	High	None	Popular	1
G	High	None	Popular	1
H	Low	None	None	0

<Table 5> Accuracy of the recommendation

Total users	Number of camera types	Correctly Recommended	Incorrectly Recommended
839	59	714 (85.1%)	125 (14.9%)

4. Conclusion

In this study, we have emphasized the importance of getting user preference by comparing with the PRL database. We have used price, rate, and popularity of the product as the key features. Recommending suitable items to each user after getting the preferences using our way, we could get 85.1% of accuracy, even though we have tested on the small dataset. Our way of recommendation method will be very powerful and widely used on various types of products, large amount of dataset and also on active consumers. However, our method is not that effective on users who only purchase one or two items. We will improve the performance of our method in next research.

References

[1] Cao Y and Li Y, " An intelligent fuzzy-based recommendation system for consumer electronic products" , Expert Syst Appl., Vol. 33, no.1, pp. 230- 240, 2007.

[2] Adomavicius G and Tuzhilin A, " Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions" , IEEE Trans Knowl Data Eng., vol. 17, no.6, pp. 734- 749, 2005.

[3] Jella Pfeiffer and Michael Scholz, " A Low-Effort Recommendation System with High Accuracy: A New Approach with Ranked Pareto-Fronts" , Business & Information Systems Engineering, Vol. 5, no. 6, pp. 397-408, 2013.

[4] Panagiotis Symeonidis, Eleftherios Tiakas, and Yannis Manolopoulos, " Product Recommendation and Rating Prediction based on Multi-modal Social Networks" , RecSys, ACM, pp 61-68, 2011.

[5] Sebastien Hemon, Thomas Largillier, and Sylvain Peyronnet, " Partial Ranking of Products for Recommendation Systems" , E-Commerce and Web Technologies Lecture Notes in Business Information Processing, Vol. 61, pp. 265-277, 2010.