A Recommender System using Collaborative Filtering and K-Mean Based on Android Application

Kunyanuth Kularbphettong[1, a], Sunisa Somngam[2,b], Cholticha Tongsiri[3,c], and Pattarapan Roonrakwit[4,d]

[1,2]Suan sunandha Rajabhat Univesrsity, 1 Dusit, Bangkok, Thailand
[3,4]Silpakorn University, CAT Tower 8 floor, Bangkok, Thailand

kunyanuth.ku@ssru.ac.th, sunisa.so@ssru.ac.th,
ctongsiri@gmail.com,ajpui20@hotmail.com

Abstract. The objective of this research is to develop the diamond recommendation system by using K-Means and Collaborative Filtering techniques. The prototype system suggests users automatically in order to maximize users' satisfaction. The system design and development will be in the form of Android (android operating system). We illustrate the methodologies and experimental results of this system. In this project, it was divided the result by the research purposes into 2 parts: developing the Mobile application for diamond recommender users and evaluating and testing the system. The results showed that the experts and users are satisfied with the system at a good level. The guide to buying diamonds actually works.

Keywords: Recommendation system, K-Mean, Collaborative filtering, Android

1 Introduction

With the advance technologies, there is an overwhelming amount of information available in the world and it is very difficult for customers to find the suitable products. To provide needed information, recommender system is the application of knowledge used to make a decision to users and it is a significant component to the businesses. Many recommendation systems were developed for use in industry and education that aim to create a list of suggestions and provide information to help customers for choosing the products. For instance, a Grundy system [1] is the first recommendation system that described models of users by using stereotypes and the result shown that user modes are effective in guiding its performance.

Collaborative Filtering is one of the significant techniques used in recommender systems to suggest products and services for customers on e-Commerce systems. This approach advices user based on the preferences of similar users and it generally analyzes relationships between users and products or services to identify the user product/service associations [2]. Ratings were given by customers to catalog items/products and users who have similar tastes will have similar tastes in the future

[3]. The number of items associated with users is evaluated the accuracy of a collaborative filtering approach. There are two types of collaborative filtering: user-based and item-based. User-based collaborative filtering predicts user's preference items from rating preference of similar users in the past and item-based collaborative filtering depends on the similarity items and this approach is based on the user rating history to indicate the ratings pattern [4]-[5].

Also, recommendation system is interested to researcher because the results of this research cause to effect in many fields and it is divided in 3 categories according to the suggested method as follows: Collaborative filtering, Content-base recommendation, and Hybrid approaches. In this paper, we present the prototype of mobile application for recommendation users by using K-Means and Collaborative Filtering techniques because Collaborative Filtering approach is sensitive with sparsity rating data in small group of users. Hence, K-means and Collaborative Filtering approaches were adapted in this project to reduce the sparsity rating problem. Furthermore, user preferences were considered to enhance the quality of this prototype.

The remainder of this paper is organized as follows. Section 2 presents related works and research methodologies used in this work. Section 3 we describe the system architecture based on the purposed model and section 4 shows the results of this experiment. Finally, the conclusion and future research are presented in section 5.

2 The Methodologies

2.1 Collaborative filtering (CF)

Collaborative filtering (CF) is one of the most effective and successful techniques of recommender systems. This technique uses the relevant feedback from other similar users to predict or recommend to other users. Amazon.com [6] is the most famous recommendation system. This recommendation system incorporates a matrix of the item similarity. In order to find the similarity of the users, the Pearson's correlation coefficient is used to compute similarity between $user_a$ and $user_u$

$$C_{a,u} = \frac{covar(r_a, r_u)}{\sigma_{r_a} \sigma_{r_u}} \tag{1}$$

where $C_{a,u}$ is the Pearson's correlation coefficient between user a and user u

r_a and r_u is the received score from user a and user u

$$covar(r_a, r_u) = \frac{\sum_{i=1}^{m} (r_{a,i} - \bar{r}_a)(r_{u,i} - \bar{r}_u)}{m} \tag{2}$$

Let $r_{a,i}$ and $r_{u,i}$ are the received score of product i from user a and user u

$\bar{r}_a$ and $\bar{r}_u$ are the average of score product from user a and user u

and m is the number of the co-rated items

According to Herlocker et al [7], they suggested to weight user similarity and computed a prediction by performing a weighted average of deviations from the neighbor's mean.

$$p_{a,i} = \bar{r}_a + \frac{\sum_{u=1}^{n} (r_{u,i} - \bar{r}_u).w_{a,u}}{\sum_{u=1}^{n} w_{a,u}} \tag{3}$$

where $p_{a,i}$ is the prediction for item i of user a

n is the number of neighbors

$w_{a,u}$ is the similarity weight between user a and user u

2.2 K-means

K-means is one of the simplest unsupervised learning algorithms for clustering data. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed a priori [8]. This algorithm aims at minimizing an objective function, in this case a squared error function. The formula of Euclidean distance is as follows:

$$j = \sum_{j=1}^{x} \sum_{i=1}^{x} \left\| x_i^{(j)} - c_j \right\|^2 \tag{4}$$

where $\left\| x_i^{(j)} - c_j \right\|^2$ is a chosen distance measure between a data point $x_i^{(j)}$ and the cluster centre c_j, is an indicator of the distance of the n data points from their respective cluster centers

3 The Experimental Setup

In this section, we described an overview and detail of the proposed system. we collected data from 150 users and each user was asked to fill out his/her personal profile including age, gender, income, type of jewelry, style, diamond shape and preferences. K-means was used to cluster into 3 groups of users by measuring distance from a point centriod Euclidean. Then collaborative filtering step calculated the similarity among each user and selected a user that is similar to the current users from finding the similarities and then the rating data is processed to predict the value prediction. Finally, the system will present the results by selecting from the highest rating for the current user.

4 The Results

In this project, it was divided the results by the research objectives into 2 parts: developing the mobile based on K-Means and Collaborative Filtering techniques and evaluating and testing the application.

4.1 Developing the mobile application

In this section, to develop the mobile application, Fig 1 and Fig 2 were shown the results of mobile application.

Fig. 1. The main page of application

Fig. 2. The results page of application

4.2 Evaluating and testing the application

Black box Testing and Questionnaires by 5 experts and 150 users were used to evaluate and test the qualities of this application. Respondents were asked to rate the recommendation results and the rating score was from 1 to 5. Black Box testing is the testing approach that focuses only on the outputs generated in response to selected inputs and execution conditions and the internal mechanism of a system or component is ignored [9]. Black box testing was assessed in the error of the project as following:

ISBN: 978-1-61804-246-0

functional requirement test, Function test, Usability test, Performance test and Security test.

Table 1. The results of Black box testing

	Experts		Users	
	$\bar{x}$	SD	$\bar{x}$	SD
1.Function Requirement Test	4.46	0.51	4.22	0.56
2. Functional Test	4.2	0.57	4.27	0.64
3. Usability Test	4.16	0.55	4.25	0.59
4. Performance Test	4.00	0.50	4.33	0.63
5. Security Test	4.25	0.47	4.19	0.54

Functional Requirement Test is evaluated the satisfaction on the ability of the system so as to meet the needs of users and functional test was used to evaluate the accuracy of the system. Usability test is a measurement of the suitability of the system. The performance of the system is assessed the processing speed of the system in Performance Test. Finally, Security test was evaluated the security of the system and Table 1 and Fig 3 were shown the results of Black box testing.

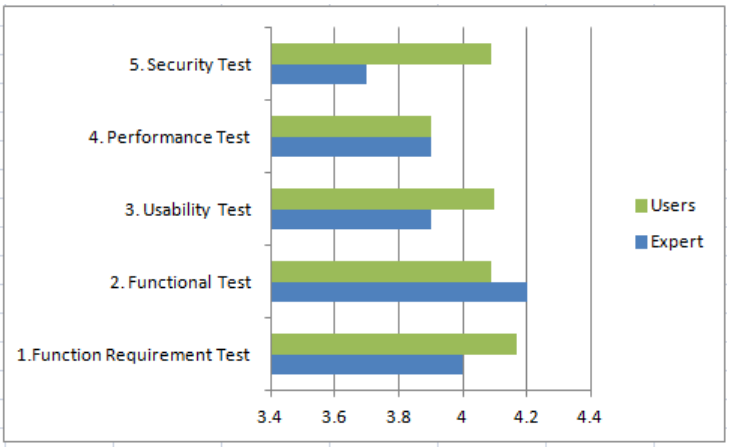

Fig. 3. . The results of Black box testing

The results showed that the Diamond recommender system based on Mobile application was satisfied the requirements of users. Means for 5 experts and 150 users were 4.2 and 4.25 respectively.

5 Conclusion

In this work, we proposed the diamond recommendation system by using K-Means and Collaborative Filtering techniques based on Mobile Application. This

system provides more suitable recommendation information to users. K-means was used to cluster optimal groups and Collaborative filtering produced recommendation results based on user's voting and preferences. The initial results showed that our approach is successfully generated the recommendation results matching with the group of users. As for the future work, we need to explore more reasonable other technologies to apply in this project to enhance the quality and quantity of services to users.

6 Acknowledgements

The authors would like to thank Suan Sunandha Rajabhat University for scholarship support this project.

References

1. Rich, "User Modeling via Stereotypes," Cognitive Science, vol. 3,no. 4, pp. 329-354, 1979.
2. Y. Hu, Y. Koren, and C. Volinsky, "Collaborative filtering for implicit feedback datasets", ICDM'08. Eighth IEEE International Conference on, pp. 263–272, 2008
3. Dietmar Jannach, Markus Zanker and Gerhard Friedrich, "Tutorial: Recommender Systems", International Joint Conference on Artificial Intelligence Beijing, August 4, 2013.
4. Ricci F, "Mobile Recommender Systems", International Journal of Information Technology and Tourism, Vol. 12, No. 3, pp. 205-231, 2011.
5. Su, X., Khoshgoftaar, T.M., X. Zhu, and R. Greiner, "Imputation-Boosted Collaborative Filtering Using Machine Learning Classifiers", Proceedings of the ACM symposium on Applied Computing, pp. 949-950, 2008.
6. Greg Linden, Brent Smith, Jeremy York, "Amazon.com Recommendations Item-to-Item Collaborative Filtering", "IEEE Internet Computing", IEEE Computer Society, vol. 7, no 1, pp 76-80, January 2003.
7. MacQueen, R., "Some Methods for classification and Analysis of Multivariate Observations" Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press, vol.1, pp.281-297, 1967.
8. Herlocker JL, Konstan JA, Borchers A and Riedl J, "An algorithmic framework for performing collaborative filtering", In: Hearst MA, Gey F and Tong R, eds., Proceedings of the 22nd International Conference on Research and Development in Information Retrieval (SIGIR'99). ACM Press, New York, pp. 230–237, 1999.
9. Laurie Williams. "Testing Overview and Black-Box Testing Technique