Restaurant Recommender System Using User-Based Collaborative Filtering Approach: A Case Study at Bandung Raya Region

Alif Azhar Fakhri\textsuperscript{1}, Z K A Baizal\textsuperscript{2}, Erwin Budi Setiawan\textsuperscript{3}

\textsuperscript{1,2,3}School of Computing, Telkom University, Bandung, Indonesia

E-mail: \textsuperscript{1}alifazharf@student.telkomuniversity.ac.id, \textsuperscript{2}baizal@telkomuniversity.ac.id, \textsuperscript{3}erwinbudisetiawan@telkomuniversity.ac.id

Abstract. Culinary becomes one of the needs of today’s society. The large number of restaurant choices and also lack of information about the restaurant become an obstacle to people's needs in choosing a restaurant. In this paper, we build a recommender system that can recommend the restaurant in Bandung area. However, today, users want to get a restaurant with a good reputation and fit their tastes, so that restaurant ratings from other users are required in the restaurant recommendation process. We implement a user-based collaborative filtering method for recommend a restaurant personally, based on ratings given by other users. We also implement two similarities, i.e., user rating similarity and user attribute similarity to find the proximity between users. We use Mean Absolute Error (MAE) to evaluate accuracy of rating prediction. The best MAE result of each performance is 1.492 for calculation without user attributes and 2.166 for calculation with user attributes.

Keywords. Culinary, Recommender system, user-based collaborative filtering, rating

1 Introduction

Culinary has become part of people’s lives today. Culinary is a part of life that is closely related to daily food consumption. Recently, culinary is not just about food needs, but also about a community lifestyle. People are increasingly like eating in restaurant, so the number of restaurants that are growing even more, especially in the area of Bandung. With a growing number of restaurants, many websites or applications provide information about restaurants in Bandung. Some information from websites or applications about restaurants in Bandung still use filtering methods based on places or types of cuisine. However, it is still possible to display choices that are relevant to the filtering results from the user, but not according with user tastes.

The majority of people or consumers will choose restaurants with high ratings and positive reviews from other people to make choices \cite{1}. Therefore, ratings and reviews are very influential with consumers who have the same tastes, so consumers get the same recommendations with other consumers with the same taste.

In this study, we develop a recommender system using the user-based collaborative filtering method, because the system is able to produce personal recommendation using the user rating as a parameter \cite{2}. User-based collaborative filtering aims to provide recommendations to users who will choose or buy certain products based on ratings given by other users \cite{3}. In this study, we use questionnaire of users who have given a rating of the restaurant they have visited as a dataset. The dataset has 86 restaurants data taken on the \textit{zomato.com} site, 593 users data taken from the questionnaire, and 50998 ratings data.
2 Related Work

We have explored some information from previous research as reference material about the advantages and disadvantages that already exist.

Restaurant is a public place that provides one dish or various dishes. Currently, many applications contain information about restaurants, e.g., name of restaurant, menu in restaurant, and food price in restaurant. The system that is popularly used in restaurant applications is recommender system [4].

Various studies related to the recommender system have been carried out. Recommender system is an effective way to help users to get information that is useful and accordance with user interests [5]. With the increasing amount of information in the internet, recommender systems are able to solve the problems caused by increasing information rapidly [6][7]. For example, applications that use the recommender systems are Youtube, Amazon, Netflix, and Facebook.

Some studies use recommender systems to implement various domains, such as travel product, book, movie, etc [8]. On the paper [9][10][11] describes the user-based collaborative filtering method used for tourism. The recommendation process is divided into 3 stages, 1) information representation of users or tourists, 2) generating user neighbors, and 3) producing recommendations of tourist objects. On the paper [12] two collaborative filterings, user-based and item-based methods are combined to expand the capacity of available information. Combining these two methods will improve the accuracy of recommendations and reduce cold-start problems. On the paper [13] explained about knowledge-based recommender system to overcome the cold-start problem. On the paper [14] explained about user user-based collaborative filtering method by using weighting similarity calculations to improve performance and accuracy. There are two examples of the weighting similarities hierarchy, 1) one-tier weighting and 2) two-tier weighting.

3 Implementation

Overall, the system that we want to build is recommender system in the restaurant domain. In this study, we use user-based collaborative filtering method. If the user wants to find a restaurant recommended by another user, then the system will searching similarity of preferences the target user and all existing users by calculating the similarity between users and the similarity of the user attributes. Then the system search the neighbors who have biggest similarity with the target user, so that restaurants that have been given a rating by neighbors will be recommended to target users who have not rated the restaurant.

![Figure 1. Screenshot program.](image-url)
You have some recommendation

Restaurant
Martabak San Fransisco
Pizza Hut
Maja House
Warung Nasi SPG
Saboga

Figure 2. Recommendation

Restaurant ratings use a Likert scale with a range of 1 to 5. 1 which means "strongly not interested", 2 which means "not interested", 3 which means "less interested", 4 which means "interested", and 5 which means "strongly interested". The system works by receiving a rating of restaurants that have been visited by users with a value of 1-5 as in figure 1. User history will save user rating data to be used for similarity calculation. After the user gives a rating of the restaurant, the user will be given a restaurant recommendation that has never been visited as in figure 2.

In this study, we use data from a questionnaire which has been filled in by users to give the restaurants that have been visited. We also use restaurant location data in Bandung taken from zomato.com site and the user has rated the restaurant. This research process is shown in figure 3.

Figure 3. Description system.

4 User-item Matrix

In the beginning process, we build a user-item matrix that is obtained from the questionnaire given to users by giving ratings of restaurants that users have visited as in the table 1. Then we use user-item matrix to calculate similarity to determine the similarity between users.

Table 1 shows an example of a user-rating matrix that represents users who give a 1-5 rating on a restaurant. Active users are users who have not rated restaurant D. We calculate the similarity to look for active user similarity with other users, so that active-user have similarity with all users. The higher the similarity value, the more similar the preferences between the two users. Otherwise, the smaller the similarity value, the preference of the two users is not similar.
Table 1. User rating matrix.

|          | restaurant A | restaurant B | restaurant C | restaurant D |
|----------|--------------|--------------|--------------|--------------|
| active user | 3            | 4            | 1            | -            |
| user 1    | 4            | 4            | 2            | 1            |
| user 2    | 1            | 1            | 2            | 4            |
| user 3    | 2            | 4            | 3            | 1            |
| user 4    | 4            | 4            | 4            | 3            |

5 Similarity

Similarity calculations are carried out in two stages, 1) calculating the user similarity and 2) calculating the user attribute similarity. We use weighted coefficient to get result of these two similarities [13].

To get restaurant recommendations from other users, we need similarity calculation to get similarity between users that can recommend restaurants according to our taste. In this study, we use Pearson correlation formula (1) to get user rating similarity.

\[
\text{pearson}(a, u) = \frac{\sum (r_{ai} - \bar{r_a})(r_{ui} - \bar{r_u})}{\sqrt{\sum (r_{ai} - \bar{r_a})^2 \sum (r_{ui} - \bar{r_u})^2}}
\]

with \( r \) is the item rating by the user, \( \bar{r_a} \) is the average rating of user \( a \), and \( \bar{r_u} \) is the average rating of user \( u \).

Each user has attributes or characteristics, so, one user with another user has similar based on attributes. In this study, we use gender and age attributes. Similarity will produce a value of 1 if both users have the same gender. otherwise, similarity will produce a value of 0 if the two users have a different gender [16]. as in formula (2)

\[
\text{gender}(u, v) = \begin{cases} 
1, & \text{u.gender = v.gender.} \\
0, & \text{u.gender \neq v.gender.} 
\end{cases}
\]

The age attributes are divided into 3 age categories,
1 = age is less than 20 years
2 = age 20 to 40 years
3 = age is more than 40 years

Similarity calculation for age using formula (3)

\[
\text{age}(u, v) = \frac{1}{|u.age - v.age| + 1}
\]

The smaller the age difference between users, the greater the similarity value. Otherwise, the greater the age difference between users, the smaller the value of similarity. This indicates that differences in age affect appetite in a restaurant.

After getting the similarity of age and gender of the user, we combine two similarities using formula (4)

\[
\text{sim}_\text{attribut}(u, v) = \text{gender}(u, v) + \text{age}(u, v)
\]

After getting the similarity user rating and similarity user attributes, we use coefficients to improve accuracy by giving weight to each similarity (5)

\[
\text{sim}(u, v) = \alpha \text{sim}_\text{attribut}(u, v) + (1 - \alpha)\text{sim}_\text{pearson}(u, v)
\]
6 Prediction

After getting similarity between users, the next step is to find the closest neighbor based on the highest similarity value. We perform search the nearest neighbor by taking a number of \( N \) user with the highest similarity to the lowest, then calculating the prediction of the top \( N \) user results. We use prediction formula (6) to produce recommendation for user \( u \). Prediction is generated from the highest recommendation generation value.

\[
P(u, i) = r_u + \frac{\sum_{v=1}^{N} sim(u, v) \times (r_{vi} - r_v)}{\sum_{v=1}^{N} sim(u, v)}
\] (6)

System will output the top \( N \) items that have the highest predictive value. Then user can choose one restaurant that is visited.

7 Performance

We build performance measure of prediction values by calculate rating prediction error from formulas (6) and actual rating prediction using the MAE method (7)

\[
MAE = \frac{\sum_{i=1}^{N} |p_i - q_i|}{N}
\] (7)

The smaller the MAE value, the better the resulting performance. Otherwise, if the MAE value is high, then the resulting performance will be bad. MAE values range is between 0 to 5 because the smallest difference between the rating prediction and the actual rating is 0, while the biggest difference is 5.

8 Evaluation

In this section, we show the results of testing and analysis of prediction error comparison using MAE to similarity user ratings with similarity user attributes.

8.1 Result

The test results will be divided into two scenarios based on the similarity method used, i.e., similarity using user attributes and similarity without user attributes.

(i) Get MAE for user-based collaborative

The results of the test scenario using a user rating similarity are shown on figure 4. We calculate MAE value based on the closest neighbor of the user as a parameter by taking the number of the nearest neighbors as many as 50 closest neighbors with a difference of 5 to obtain 9 test results. The lowest MAE value is 1.492 with 50 closest neighbors.

(ii) Get MAE for user-based collaborative using user attribute

The results of the test scenario using a user attributes similarity are shown on figure 5. We calculate MAE value based on the closest neighbor of the user as a parameter by taking the number of the nearest neighbors as many as 50 closest neighbors with a difference of 5 to obtain 9 test results. The lowest MAE value is 2.166 with 50 closest neighbors.

8.2 Analysis

In this section, we analyze the result based on the testing scenario.

Figure 4 show the lowest MAE value is in the number of closest neighbors 50 with MAE value 1.492. The more the numbers of neighbors, then the information from user preferences is good enough to be able to produce a small MAE value. Otherwise, when the number of neighbors 20, the MAE value is 1.981 This is because the amount of information from user preferences is not enough to produce a low MAE.

Figure 5 show the lowest MAE value is in the number of closest neighbors 50 with MAE value 2.166. The more the numbers of neighbors, then the information from user preferences is good enough to be
Figure 4. MAE results without user attributes.

Figure 5. MAE results with user attributes.

able to produce a small MAE value. Otherwise, when the number of neighbors 35, the MAE value is 2.531. This is because the amount of information from user preferences is not enough to produce a low MAE.

9 Conclusion

In summary, we studied how to calculate similarity between user-based collaborative filtering with user-based collaborative filtering using user attribute. Calculation on user-based collaborative filtering only using a Pearson Similarity while calculation with user attribute using similarity by adding the calculation of attribute age and gender. The selection of the number of neighbors affects the MAE value. the more the number of neighbors, the better the performance will be. Performance results are better in calculations without user attributes because preference information from user attributes is less so the resulting calculation becomes less accurate. MAE results with a value of 1.492 are good because the range of the MAE value is 0 to 5 so that if the MAE value is less than 2.5 the MAE produced is good. Otherwise, if the MAE value is more than 2.5 then the MAE produced is not good. So, recommender system using user-based collaborative filtering is better than using user attribute.

References

[1] Farooque U, Khan B, Junaid A B, and Gupta A 2014 Collaborative filtering based simple restaurant recommender Computing for Sustainable Global Development pp. 495-499 IEEE.

[2] Mu X and Li T 2010 User-based collaborative filtering based on improved similarity algorithm Computer Science and Information Technology 8 pp. 76-80 IEEE.

[3] Aggarwal C C 2016 An introduction to recommender systems (Cham: Springer)

[4] Gupta A and Singh K 2013 Location based personalized restaurant recommendation system for mobile environments Advances in Computing, Communications and Informatics pp. 507-511 IEEE
[5] Zeng J, Li F, Wen J, and Hiroka S 2016 A restaurant recommender system based on user preference and location in mobile environment Advanced Applied Informatics pp. 55-60 IEEE.

[6] Jonnalagedda N, Gauch S, Labille K, and Alfarhood S 2016 Incorporating popularity in a personalized news recommender system PeerJ Computer Science 2 e63

[7] Baizal Z K A, Widyantoro D H, and Maulidevi N U 2016 Factors Influencing Users Adoption of Conversational Recommender System Based on Product Functional Requirements 1575-1585 TELKOMNIKA

[8] Baizal Z K A, Widyantoro D H, and Maulidevi N U 2016 Query refinement in recommender system based on product functional requirements Advanced Computer Science and Information Systems pp. 309-314 IEEE.

[9] Jia Z, Yang Y, Gao W, and Chen X 2015 User-based collaborative filtering for tourist attraction recommendations Computational Intelligence Communication Technology pp. 22-25 IEEE.

[10] Arigi L R H, Baizal Z K A, and Herdiani 2018 A Context-aware recommender system based on ontology for recommending tourist destinations at Bandung Journal of Physics: Conference Series 971 IOP Publishing.

[11] Baizal Z K A, Rahmawati A A, Lhaksmana K M, Mubarok M Z, and Qadrian M 2018 Generating Travel Itinerary Using Ant Collony Optimization 16(3) Telkomnika.

[12] Wang B, Huang J, Ou L, and Wang R 2015 A collaborative filtering algorithm fusing user-based, item-based and social networks pp. 2337-2343 IEEE.

[13] Baizal Z K A, Widyantoro D H, and Maulidevi N U 2016 Design of knowledge for conversational recommender system based on product functional requirements Data and Software Engineering pp. 1-6 IEEE.

[14] Li W, Xu H, Ji M, Xu Z, and Fang H 2016 A hierarchy weighting similarity measure to improve user-based collaborative filtering algorithm Computer and Communications pp. 843-846 IEEE.

[15] Guo Y, Huang M, and Lou T 2015 A Collaborative Filtering Algorithm of Selecting Neighbors Based on User Profiles and Target Item Web Information System and Application Conference pp. 9-14 IEEE.

[16] Zeng Y, Bi Y, Wang J, and Lin Y 2015 Collaborative Filtering Recommendation Algorithm Optimization Based on User Attributes Computational Intelligence and Design 1 pp. 580-583 IEEE.