The Exploration of Restaurant Recommender System

Tora Fahrudin and Nelsi Wisna

School of Applied Sciences, Telkom University, Indonesia

Abstract: The exploitation of Recommender Systems (RS) is still a challenge, hence it is important to explore the three correlated attributes, such as restaurant, food, and service ratings. Therefore, this study provides an in-depth review of these attribute ratings using the Collaborative Filtering (CF) technique. Experiments were performed with k-NN, SVD, Slope One, and Co-Clustering algorithms, while RMSE, MSE, MAE, and FCP were used as evaluation metrics. The results showed that the service restaurant rating predictions produced the best average MSE and RMSE accuracy in 5 and 10-fold cross-validation. Furthermore, the best hyperparameter of algorithms using Grid Search was achieved in restaurant rating prediction. In conclusion, SVD surpasses other algorithms in MSE and RMSE for all scenarios.

Keywords: Restaurant, Recommender System, Rating, Collaborative Filtering

Introduction

Data explosion is inevitable these days and this is the reason more data are stored in computer systems in various categories and formats (Zhu et al., 2009). This explosion is called Big Data, which is a high-volume, high-velocity, and high-variety information asset requiring cost-effective and innovative forms of information processing to improve insight and decision-making (Gandomi and Haider, 2015). Implementation of Big Data is found in information retrieval systems such as Google, Devil Finder, and AltaVista (Isinkaye et al., 2015). Bellogín and Said (2019) found that Recommender Systems (RS) are closely related to information retrieval systems because they use similar models for identifying relevant data.

Furthermore, Fkih (2021) found that the difference between RS and information retrieval is that users do not need to query before getting the relevant information, as it formulates the query based on the user profile. Seo et al. (2021) identified that they are mainly divided into personalized and group recommendations since the target was built for groups and not for an individual.

Currently, RS is implemented in several fields such as E-government, E-business, E-commerce, E-library, E-learning, E-tourism, E-resource, and E-group activity (Lu et al., 2015). The restaurant recommender system is part of E-tourism that focuses on providing similar menus based on price and taste (Burke, 2000), reputation (Fakhri et al., 2019), food quality and service (Asani et al., 2021), user's preference and location information (Zeng et al., 2016) and user reviews (Hassan and Abdulwahhab, 2017).

Alhijawi and Kilani (2020) discovered that Collaborative Filtering (CF) is the most popular technique for analyzing historical user feedback information to predict recommendations. Presently, there are three types of CF, which include memory-based, model-based, and hybrid. There are also two crucial steps in CF, such as finding similar users or items by using a particular similarity measure and calculating a rating based on the similarity (Fkih, 2021).

In this study, an in-depth review of restaurant, food, and service ratings for a recommender system using the CF technique was provided. Furthermore, an experimental comparative study was conducted on restaurants using a consumer rating dataset from UCI Machine Learning Repository (Vargas-Govea et al., 2011) to compare their performances in some algorithms such as k-NN, SVD, Slope One, and Co-Clustering.

The rest of this study is arranged as follows. The Related Works section describes related works, Materials and Methods section describes the assessment models. The experimental results and evaluation is presented in the Experimental Results section and the last section is the conclusion and discussion.

Related Works

Recommendation System

Several studies about RS were conducted, but the first automated RS was established by Though Grandy (Singh et al., 2021). Afterward, RS was implemented in personalized news (Resnick et al., 1994), movies (Herlocker et al., 2000), and online jokes (Goldberg et al., 2001) that relied on a...
rating structure. It is mostly divided into three types, which include Collaborative Filtering (CF), Content-Based (CB), and Knowledge-Based (KB) (Lu et al., 2015). CF is observed to perform better than CB with low user ratings (De Campos et al., 2010). Moreover, CB recommendations have limited accuracy for users with very few historical ratings.

The CF algorithms are divided into two categories, which include memory-based and model-based approaches. The first category uses all data to find a set of users/items that are similar to the target. Meanwhile, the second category builds a model using machine learning to describe the user's behavior for predicting their choices. There is a list of m users in CF, denoted as (U), i.e., U = {u1, u2, ..., um} and a list of n items (I), i.e., I = {i1, i2, ..., in}.

According to Aditya et al. (2016), memory-based CF is also called neighborhood CF, which is often achieved in two ways, namely user-based and item-based techniques. In the user-based, an item's recommendation rating for a user is calculated depending on the rating of the item by other similar users. While for item-based, the rating is predicted based on how the user rates the same items. These two techniques operate on a matrix of user-item ratings.

The study conducted by Isinkaye et al. (2015) revealed that a model-based collaborative filtering algorithm provides item recommendations by first developing a user rating model. Nassar et al. (2020) further highlighted some model-based methods such as Latent Semantic Analysis (LSA), Bayesian Clustering, Support Vector Machine, Latent Dirichlet Allocation, and Singular Value Decomposition.

**Restaurant RS**

Restaurant RS such as Entree (Burke et al., 1996), R-Cube, I-m felling LoCo, REJA, Open Table, and TripBuilder is a subset of tourism recommendations (Pettersen and Tverte, 2016). In Entrée application, price and taste variables are used to build RS, but some methods utilized additional geo-references such as REJA (Martinez et al., 2009) and I’m feeling LoCo (Saiph Savage et al., 2012). Furthermore, some user preferences were also created using food, price, city area, and restaurant name variables such as in the R-Cube system (Kim and Banchs, 2014). Other restaurant recommender system such as the OpenTable application uses additional information from user interaction histories such as click and search data, the metadata of restaurants, and user reviews (Das, 2015). TripBuilder uses Spatio-temporal information to recommend personalized sightseeing tours in tourism recommendations (Brilhante et al., 2015). However, this study focused on examining restaurant, food, and service ratings for the restaurant recommender system.

**Neighbourhood CF**

The neighborhood method is focused on computing the relationships between items or users (Xie, 2019). Furthermore, it requires the relationships matrix between items i.e., item-item, or users i.e., user-user. There are two types of neighborhoods, namely item to item and user to user. Where \( N^i(u) \) denotes the \( k \) nearest neighbors of users \( u \) that have rated item \( i \) and \( N^u(i) \) represents the \( k \) nearest neighbors of item \( i \) that are rated by user \( u \). The objective of this method is to estimate the rating of user \( u \) for item \( i \) using similarity values as seen in Eq. (1):

\[
\hat{r}_{ui} = \frac{\sum_{r \in N^u(i)} \text{sim}(u,v) r_v}{\sum_{r \in N^u(i)} \text{sim}(u,v)}
\]

where, \( \text{sim}(u,v) \) is the similarity value between users \( u \) and \( v \). This similarity is either expressed as cosine in Eq. (2) or Pearson correlation in Eq. (3):

\[
\text{Cosine}(u,v) = \frac{\sum_{i \in \text{local}} (u_i - \bar{r}_u)(v_i - \bar{r}_v)}{\sqrt{\sum_{i \in \text{local}} (u_i - \bar{r}_u)^2} \sqrt{\sum_{i \in \text{local}} (v_i - \bar{r}_v)^2}}
\]

\[
\text{Pearson}(u,v) = \frac{\sum_{i \in \text{local}} (u_i - \bar{r}_u)(v_i - \bar{r}_v)}{\sqrt{\sum_{i \in \text{local}} (u_i - \bar{r}_u)^2} \sqrt{\sum_{i \in \text{local}} (v_i - \bar{r}_v)^2}}
\]

**Latent Factors Methods**

Latent or hidden features capture more relationships between users and items by transforming them into the same latent factor space, thereby making them directly comparable (Xie, 2019). When the user rating matrix is sparse, Singular Value Decomposition (SVD) is one of the latent factor models approaches for overcoming the problem (Rodpysh et al., 2021). This SVD decomposes a matrix into three more matrices i.e., user-item-rating and extracts the factors from high-level matrix factorization (Chen, 2020). The rating of user \( u \) for item \( i \) in SVD is presented in Eq. (4):

\[
\hat{r}_{ui} = USV^T
\]

where, \( U \) is a singular matrix of user latent factors, \( S \) is a diagonal matrix and the \( V \) is a singular matrix of item latent factors.

**Slope One**

Slope one algorithm is based on differences in popularity of items for the user project scoring matrix introduced by Daniel Lemire and Anna Machlahan (Lemire and Maclachlan, 2005). This approach basically uses a unitary linear model \( y = x + b \), where \( y \) is the score of the predicted target user, \( x \) is the user's target score and \( b \) is the deviation value (Song and Wu, 2020). Therefore, with the use of slope one, the rating of user \( u \) for item \( i \) is presented in Eq. (5):

\[
\hat{r}_{ui} = \sum_{r \in N^u(i)} \text{sim}(u,v) r_v + b
\]
\[
\hat{y}_u = \mu_u + \frac{1}{|R(u)|} \sum_{j \in R(u)} \text{dev}(i, j)
\]

where, \( \mu_u \) denotes the mean of all ratings given to item \( i \), \( R(u) \) represents the set of relevant items of user \( u \), and also \( \text{dev}(i, j) \) represents the average difference between the rating of \( i \) and \( j \).

**Co-Clustering Algorithm**

The co-clustering algorithm is a CF method that uses co-clustering to generate predictions based on the average ratings of the co-clusters i.e., user-item neighborhoods, and takes into account the individual biases of the users and items (George and Merugu, 2005). In this approach, some clusters are assigned to users and items, which include \( C_u \) denoting user cluster, \( C_i \) representing item cluster, and \( C_{ui} \) indicating co-cluster of user and item. The prediction is expressed in Eq. (6) below:

\[
\hat{y}_{ui} = \overline{C_u} + (\mu_u - \overline{C_u}) + (\mu_i - \overline{C_i})
\]

where, \( \overline{C_u} \) represents the rating of the co-cluster \( C_u \), \( \overline{C_i} \) is the average rating of \( u \) ’s cluster, \( \overline{C_i} \) denotes the average rating of \( i \) ’s cluster, \( \mu_u \) represents the mean of all ratings given by user \( u \) and \( \mu_i \) denotes the mean of all ratings given to item \( i \).

**Materials and Methods**

**Dataset**

The dataset of restaurants and consumers used in this study is obtained from UCI Machine Learning Repository (Dua and Casey, 2017). Furthermore, a user-item-rating dataset containing 1161 rows of data with five attributes is used, such as userId, placeId, restaurant rating, food rating, and service rating. Table 1 shows a detailed description of the attributes.

Figures 1, 2, and 3 show the distribution data for each rating category. It was observed that most users are satisfied with the restaurants and their foods, but not with the services. In Fig. 3, most users rated the service as a medium, but the difference is small.

**Tools and Library**

A Python scikit known as surpriselib was used for RS (Hug, 2019) to assess three restaurant rating attributes. Furthermore, Kaggle is used with Python version 3.7.12, scikit learn 0.23.2, and surprise 1.1.1 to process the recommendation assessment.

**Evaluation Metrics**

In the evaluation of the three performance attributes of the restaurant recommender system, four metrics were utilized, which include Mean Absolute Error (MAE), Mean Square Error (MSE), Root Mean Square Error (RMSE), and Fraction of Concordant Pairs (FCP) (Al-Ghamdi et al., 2021).

Based on Eq. (7), MAE was used to compute the average magnitude of the errors between the observed and predicted ratings without considering their direction:

\[
\text{MAE} = \frac{1}{|R|} \sum_{u \in R} |r_u - \hat{y}_u|
\]

In Eq. (8), MSE is the average of the squared errors between the observed and predicted ratings. MSE is a measure of the quality of an estimator:

\[
\text{MSE} = \frac{1}{|R|} \sum_{u \in R} |r_u - \hat{y}_u|^2
\]

RMSE expressed in Eq. (9) is used to calculate the residual i.e., the difference between predicted and actual values for rating the data. It was observed that RMSE was heavily affected by a few worse predictions compared to others when the errors were squared and the mean was calculated:

\[
\text{RMSE} = \sqrt{\frac{1}{|R|} \sum_{u \in R} (r_u - \hat{y}_u)^2}
\]

Moreover, FCP is a method used for overcoming the drawback of MAE, MSE, and RMSE because it does not consider the different rating scales from one user to another. A higher FCP means more accuracy than a lower FCP and it is calculated by using Eq. (10)-(12) as follows:

\[
\text{FCP} = \frac{n_1}{n_s + n_d}
\]

where:

\[
n_1 = \sum \left| \left( i, j \right) | \hat{y}_{ui} > \hat{y}_{ui and r_u > r_i} \right| \right| \right|
\]

\[
n_d = \sum \left| \left( i, j \right) | \hat{y}_{ui} > \hat{y}_{ui and r_u < r_i} \right| \right| \right|
\]

**Methods**

This study aims to evaluate three correlated attributes for a restaurant recommender system. The experimental method is shown in Fig. 4, in which the three rating attributes were the primary sources. Afterward, k-fold cross-validation was conducted by using 5 and 10 as the k parameter.
Table 1: Attributes description

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Type</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>UserId</td>
<td>Nominal</td>
<td>-</td>
</tr>
<tr>
<td>PlaceId</td>
<td>Nominal</td>
<td>-</td>
</tr>
<tr>
<td>Restaurant rating</td>
<td>Numeric</td>
<td>[0, 1, 2]</td>
</tr>
<tr>
<td>Food rating</td>
<td>Numeric</td>
<td>[0, 1, 2]</td>
</tr>
<tr>
<td>Service rating</td>
<td>Numeric</td>
<td>[0, 1, 2]</td>
</tr>
</tbody>
</table>

Table 2: Hyper-parameter setting for GS

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>k-NN</td>
<td>Number of neighbors</td>
<td>[5, 10]</td>
</tr>
<tr>
<td></td>
<td>Similarity measures</td>
<td>[msd, cosine, pearson]</td>
</tr>
<tr>
<td>SVD</td>
<td>Number of iterations</td>
<td>[5, 10]</td>
</tr>
<tr>
<td></td>
<td>Learning Rate</td>
<td>[0.002, 0.005]</td>
</tr>
<tr>
<td></td>
<td>Regularization Term</td>
<td>[0.4, 0.6]</td>
</tr>
<tr>
<td>Slope</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Co-Clustering</td>
<td>Number of user clusters</td>
<td>[5, 10]</td>
</tr>
<tr>
<td></td>
<td>Number of item clusters</td>
<td>[5, 10]</td>
</tr>
<tr>
<td></td>
<td>Number of iterations</td>
<td>[5, 10]</td>
</tr>
</tbody>
</table>

Fig. 1: Distribution of restaurant rating

Fig. 2: Distribution of food rating

Fig. 3: Distribution of service rating

Fig. 4: Experiment methods

The optimal values of the algorithm’s parameters were also compared using Grid Search (GS). Furthermore, it is a tuning technique for computing optimal hyperparameter values using an exhaustive search method. The objective of GS is to compare the best accuracy of three restaurant rating attributes. Table 2 shows a hyperparameter used in this experiment.

Experimental Result and Evaluation

To evaluate the performance of three rating attributes in restaurant RS, the results of four algorithms were compared under two scenarios, which include surprise usage of the default algorithm and exploring the best accuracy for each algorithm by using GS.

Figure 5 shows that service and restaurant ratings have better accuracy prediction compared to food. It was observed that the best average result in 5-fold cross-validation was achieved by service ratings in the MSE, RMSE, and FCP metrics, but food rating prediction was
the worst. This achievement is consistent with the 10-fold cross-validation scenario’s results shown in Fig. 6.

Tables 3 and 4 provide the best value for each algorithm in 5- and 10-fold cross-validations. It was observed that the best values achievement was dominated by restaurant rating followed by service and food.

Figure 7-9 shows the average MAE, MSE, RMSE, and FCP for 5-fold cross-validation in each algorithm and rating attribute. It was observed that SVD showed the best accuracy of MSE and RMSE in the three rating attributes. Meanwhile, the worst accuracy was recorded in k-NN for restaurant and service ratings, as well as Co-Clustering for food ratings.

Figure 10-12 show the average MAE, MSE, RMSE, and FCP for 10-fold cross-validation for each algorithm and rating attribute. It was discovered that the accuracy in 10-fold cross-validation scenarios is better than 5-fold cross-validation in all rating attributes and all metrics.

In the aspect of algorithm performance, SVD also achieved the best accuracy of MSE and RMSE in restaurant and service rating. Furthermore, Co-Clustering showed the worst performance in the food attribute, and k-NN produced the worst performance in both restaurant and service attributes.

This means that the algorithm performances were generally identical for the 5 and 10-fold cross-validation scenarios. SVD outperforms all other algorithms in MSE and RMSE metrics and the best accuracy for all approaches is recorded in a restaurant, followed by service and food attributes.

**Table 3:** Best scores in GS for 5-fold cv

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Restaurant</th>
<th>Food</th>
<th>Service</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>0.518(^1)</td>
<td>0.541(^3)</td>
<td>0.540(^2)</td>
</tr>
<tr>
<td>MSE</td>
<td>0.490(^1)</td>
<td>0.513(^2)</td>
<td>0.513(^2)</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.700(^1)</td>
<td>0.716(^2)</td>
<td>0.716(^2)</td>
</tr>
<tr>
<td>FCP</td>
<td>0.553(^1)</td>
<td>0.553(^1)</td>
<td>0.518(^2)</td>
</tr>
</tbody>
</table>

**Table 4:** Best scores in GS for 10-fold cv

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Restaurant</th>
<th>Food</th>
<th>Service</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>0.517(^1)</td>
<td>0.541(^2)</td>
<td>0.541(^2)</td>
</tr>
<tr>
<td>MSE</td>
<td>0.485(^1)</td>
<td>0.513(^2)</td>
<td>0.513(^2)</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.695(^1)</td>
<td>0.716(^2)</td>
<td>0.716(^2)</td>
</tr>
<tr>
<td>FCP</td>
<td>0.545(^1)</td>
<td>0.529(^3)</td>
<td>0.532(^2)</td>
</tr>
</tbody>
</table>

Fig. 5: 5-fold cross-validation average results

Fig. 6: 10-fold cross-validation average results

Fig. 7: Restaurant’s performance for each algorithm in 5-fold

Fig. 8: Food’s performance for each algorithm in 5-fold
Discussion

An in-depth comparative study of three ranking attributes of the restaurant recommender system, namely restaurant, food, and service ratings has been conducted. The four most common RS algorithms were utilized to examine the attribute that gives the best rating prediction performance result.

Furthermore, four well-known evaluation metrics were used in 5 and 10-fold cross-validation to evaluate the accuracy of three restaurant attribute ratings. The grid search method was also performed to explore the hyper-parameters of each algorithm for three rating attribute accuracy. It was observed that the best average rating prediction accuracy achieved by all the algorithms is service, followed by restaurant and food rating.

Conclusion

From the experimental result, evaluation, and also discussion section, it can be concluded that food rating prediction is the most difficult in relation to the restaurant recommender system. Meanwhile, for the best hyperparameter using GS, restaurant rating prediction accuracy beat other attributes in MAE, MSE, RMSE, and FCP. In conclusion, SVD suppresses other algorithms in MSE and RMSE in all scenarios.

Acknowledgment

The funding of this study was supported by Telkom University Research and Community Service's Directorate via Applied Basic Research Scheme.

Author's Contributions

Tora Fahrudin: Responsible for designing the method, arranging the paper’s contribution, and conducting experiments.
Nelsi Wisna: Helped to analyze the experimental results and draw conclusions.

Ethics

This article is original and contains unpublished materials. All authors have read and approved the manuscript and no ethical issues are involved. Also, there is no conflict of interest between authors.

References


George, T., & Merugu, S. (2005, November). A scalable collaborative filtering framework based on co-clustering. In Fifth IEEE International Conference on Data Mining (ICDM’05) (pp. 4-pp). IEEE. https://doi.org/10.1109/ICDM.2005.14


