International Journal of Application on Sciences, Technology and Engineering (IJASTE)
Volume 1, Issue 4, 2023. ISSN:2987-249

MOBILE-BASED FOOD RECOMMENDATION SYSTEM USING HYBRID FILTERING METHODS

Venny Cyntia[1], Bagus Mulyawan[2], Manatap Sitorus[3]

[1] Information System Study Programs, Faculty of Information Technology, Tarumanagara University
Email : venny.535200006@stu.untar.ac.id
[2] Information System Study Programs, Faculty of Information Technology, Tarumanagara University
Email : bagus@fti.untar.ac.id
[3] Information System Study Programs, Faculty of Information Technology, Tarumanagara University
Email : manataps@fti.untar.ac.id

Submitted: 27-09-2023, Revised: 27-10-2023, Accepted: 11-12-2023

ABSTRACT

In the era of technological progress, information technology-based applications have changed the way we view and choose food. With more and more choices available, many people face the challenge of trying different foods because they are not sure whether a dish will suit their taste. Therefore, a recommendation system was developed to increase user satisfaction and ease of food selection using the Hybrid Filtering method which combines Collaborative Filtering and Content-Based Filtering with the cascade technique. The analysis carried out on this system will be based on food descriptions which include the composition of the food itself such as taste and main ingredients. Meanwhile, users will be asked to fill in preferences based on food descriptions. The operation of this technique involves making recommendations first using the Collaborative Filtering method. Next, the recommendations from this method are further filtered using the Content-Based Filtering method. To calculate the similarity value between items, the cosine similarity formula is used, and to predict missing ratings, the weighted sum formula is used. The research results show that combining these two methods produces better recommendations than using either method separately. The results obtained from the calculation data show that the value produced to find out how big the error value is between the predicted value and the actual value for ten items is 0.29. Apart from that, a survey was also conducted on respondents, and the findings showed that eight out of ten people stated that the recommendations given were in line with their preferences.

Keywords: *recommendations system, hybrid filtering, collaborative filtering, content-based filtering*

1. PREFACE

Introduction

In this era of technological progress, information technology-based applications have changed the way we interact in various aspects of daily life [8]. One significant change has occurred in the way we view and choose food. Technological advances have had a positive impact on the food and beverage industry, with food apps becoming an important tool for users to search for, order and enjoy dishes. However, the increasing number of food choices can be a challenge for many people. Observations show that this problem is felt by employees at the Mayora Indah Tbk company, where the company provides an application for ordering food with employees complaining that they find it difficult to try different foods because they are not sure whether a particular dish suits their taste or not. Additionally, research has shown that when users are faced with too many choices, they can become overwhelmed and face decision fatigue, making it more difficult for them to make a choice. In fact, having fewer options can result in more focused decision making. While food apps offer convenience and access to a variety of culinary options, they can also lead to too much choice. This is why recommendation systems, as mentioned above, are very important as they simplify the food selection process and increase user satisfaction in choosing foods.

Problem Formulation

The challenges faced by consumers and restaurants can be summarized as difficulties in choosing food and challenges in choosing types of food. Therefore, to help food companies display their variety of dishes and increase user satisfaction and ease in choosing food, a food recommendation system has been designed. This system functions to offer and recommend food items that suit consumer preferences using the Hybrid Filtering method which combines Collaborative Filtering (CF) and Content-Based Filtering (CBF). The recommendation system will recommend foods based on food descriptions that contain the composition of the food itself such as taste and main ingredients. For this part users will be asked to fill in their preferences based on the description of the food. And from there the food recommendations will be analyzed.

2. RESEARCH METHOD

Recommendation System

A recommendation system is a system designed to predict a set of items or information that users are likely to be interested in the future, and then present these items as top recommendations [4]. This system assists users in identifying suitable products that match their preferences and desires. Benefits for business while using recommender systems:

- **Revenue**

  Past years, many researchers have studied and generate many algorithms to learn increasing rate for an online customer like Amazon site. Also, These algorithms study the difference between shopping online sites with others using recommender systems for items to increase revenue by increasing the number of sales.

- **Client Satisfaction**

  Client Satisfaction Many times customers tend to expect to see near similar product recommendation from their last browsing search on thesite. Mainly because they believe they will get more serious chances for better products. When they leave the situation and get back afterward. It would assist if their browsing data from the previous shopping or viewing product list.

- **Personalization**

  We often get recommendations from our friends. They recognize what we like better than anyone else. This is the only reason they are adept at recommending things and is what recommendation systems try to model. You can utilize the data collected indirectly to improve your website's overall services and assure that they are suitable according to a user's preference.

- **Discovery**

  People need to be recommended items they would like or prefer, and when they find a web page for food, shopping, movie, songs, etc. meet their hopes they bound to visit this site again [3].

Collaborative Filtering

Collaborative filtering is a recommendation system whose process involves gathering feedback in the form of item ratings collected from various other users who have similar preferences or tastes to predict the ratings a user might give to an item, determining how an item should be recommended [2]. Collaborative filtering can be divided into two types: user-based collaborative filtering and item-based collaborative filtering. This study will focus on the use of item-based collaborative filtering, which is a method that finds similar items based on a user's liked items. The way this method works is by identifying items that a user has liked, then finding other items that are similar to the previously liked items and recommending these similar items to the user [5].

CF is based on the way in which humans have made decisions throughout history. Besides on our own experiences, we also base our decisions on the experiences and knowledge that reach each of us from a relatively large group of acquaintances [1].In its implementation, the system collects explicit user information such as ratings given to items. Then, it predicts ratings for the target item by calculating similarity using the collection of ratings from items that have been rated. To calculate the similarity between items, the cosine similarity method is typically used.

Content-based Filtering

This recommendation system, using this method, recommends items to users based on the similarity between items that users have previously liked or selected. Similarity between items is calculated based on the features or characteristics of those items. In this way, the system understands user preferences and provides recommendations that share similar characteristics with items that the user has liked [6].

In this system, the first step after acquiring data is to build a binary feature matrix. This involves creating a matrix vector from the attribute information of each food item, where the matrix's entries will have a value of 1 if a specific attribute is present in the item and 0 if it is not. The system then builds item profiles and user profiles using the rated content. User profiles are formed based on explicit information provided directly by users or the use of clear and measurable data in the form of preferences or observable actions. Once all the data is encoded into vectors for calculation, the next step is to calculate item similarity with user preferences using Cosine Similarity. The smaller the angle between these vectors, the more similar they are.

$$S_C(A,B) = \frac{A \,.\, B}{|A||B|}$$

Where:

a. A.B : Product of vector A and B

b. |A| and |B| : Length of vector A and B

Cosine Similarity

Cosine similarity is an algorithm used to measure the similarity level between two vectors within an item. The result of cosine similarity is a positive value that falls within the range of 0 to 1 [2]. The fundamental concept in this calculation is to identify or separate users who have provided positive ratings for both item a and item b. After that, the similarity calculation can be performed based on the rating patterns of users who have both of these items [7].

$$Sim(i,j) = \frac{\sum_{u \in U}(Ru,i - \bar{R}i)(Ru,j - \bar{R}j)}{\sqrt{\sum_{u \in Uij}(Ru,i - \bar{R}i)^2 \sqrt{\sum_{u \in U}(Ru,j - \bar{R}j)^2}}}$$

Where:

a. Sim(i, j) : Represents the similarity value between item i and item j.
b. u ∈ Uij : Denotes the set of users who have rated both item i and item j positively.
c. Ru,i and Ru,j : Stand for the rating values given by user u for item i and item j,
d. $\bar{R}$I : Represents the average rating for item i.
e. $\bar{R}$j : Represents the average rating for item j.

Weighted Sum

The weighted sum is used to calculate the prediction or rating that a user may give to a particular item. The predicted rating is obtained from the ratings given by other users to items that have the highest similarity with the target user. Eventually, the items with the highest predicted values are recommended. The algorithm for this calculation is as follows:

$$P_{(u,j)} = \frac{\sum i \in j \ (R_{u,i} * S_{i,j})}{\sum i \in j \ |S_{i,j}}$$

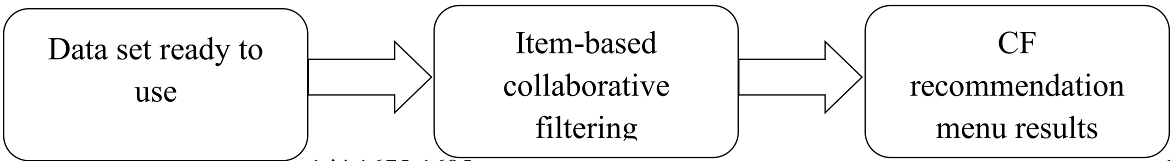
Where:

a. $P(u,j)$: Prediction of user u for product j.
b. $i \in j$: Set of similar products.
c. Ru,i : User u's rating for product i.
d. Sim (i,j) : Similarity value between product i and product j.

Hybrid Filtering

Hybrid filtering is a recommendation system that combines several different methods. In this research, hybrid filtering combines the content-based filtering and collaborative filtering approaches. The aim of this approach is to leverage the strengths of each method and address the weaknesses of each method to generate better recommendations [3]. For example, CBF is used to handle the cold start problem (the inability to provide recommendations due to a lack of preference information) for new users and items that are present in the CF method, while CF is used to address the over-specialization problem (where users are limited to recommendations similar to their known profiles).

The steps in the Hybrid Filtering CF-CBF method can be seen in the following flowchart:



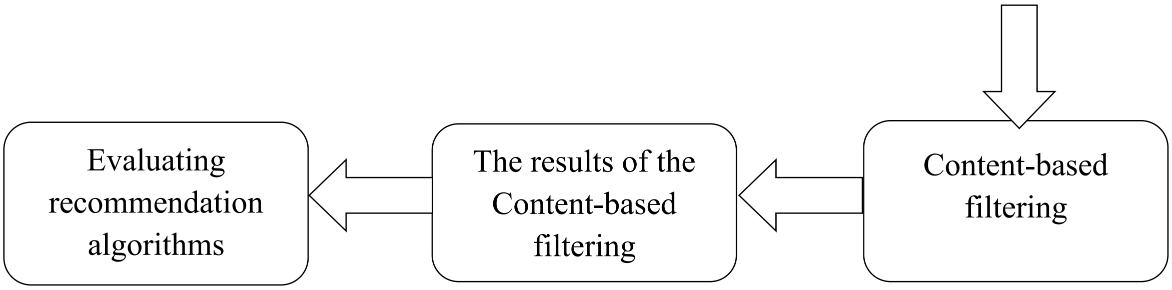


Figure 1. Stage of Hybrid Filtering

In this study, the focus is on creating a recommendation system using a cascade approach to the CF and CBF methods. In this hybrid approach, the recommendation system is designed to generate a list of recommended items using item-based CF, which is then re-ranked using the CBF method to select the top items as the final recommendations. This approach involves refining the results of previous recommendations, where the first recommendation method is tasked with generating a rough list of candidate recommended items, and the second and subsequent recommendation methods filter this initial candidate list to create the final list of recommendations (Hilmi & Agung, 2010).

Mean Absolute Error

Afterwards, use Mean Absolute Error (MAE) to measure the evaluation metric used to quantify how much the predictions differ from the actual values of a variable. In the context of recommendation systems, MAE assesses how closely the predicted ratings generated by the system match the actual ratings provided by users. A lower MAE value indicates better prediction accuracy.

$$MAE_i = \frac{1}{n}\sum\nolimits_{i \in E}(|pi - ri|)$$

Where:

a. MAEi : Model evaluation algorithm.

b. ri : Actual rating value.

c. pi : Predicted rating value.

d. n : Number of pairs of actual and predicted ratings.

RESULT AND DISCUSSION

Recommendation giving experiment

In the results and discussion section, we will try implementing a food recommendation system using Hybrid Filtering CF-CBF. First, we will generate some recommendations using the CF method, and from those results, we will generate recommendations again using the CBF method. The data used comes from the collected data, which consists of food ratings and user preferences. There are two aspects tested in this data: the accuracy of the prediction results compared to the actual ratings and how well users match the recommendations given. In this experiment, we are using 10 food items and 5 users. You can see an example of the data in Table 1.

Tabel 1. User data and ratings for each item.

	User 1	User 2	User 3	User 4	User 5
Sapi lada hitam	4	0	0	3	4
Ayam cabe ijo	0	4	2	0	2
Udang goreng tepung	5	4	5	4	2
Soto betawi	4	0	5	3	5
Soto medan	4	5	5	0	5
Bakso ikan	0	4	2	3	4
Tomyam	5	0	5	3	0
Sate taichan	5	4	0	3	5
Nasi padang	5	0	5	4	0
Rice bowl chicken katsu	5	4	4	0	4

The first step in Collaborative Filtering is normalization, which involves subtracting the rating values given by the user from the overall average rating value given by all users. The purpose of normalization is to improve efficiency by eliminating potential biases that may arise in the ratings given by users. These biases can occur due to differences in the rating scales provided by users. The calculation results can be seen in Table 2.

Tabel 2. The value of subtracting the average rating for each item.

	User 1	User 2	User 3	User 4	User 5
Sapi lada hitam	0.333333	0.000000	0.000000	-0.666667	0.333333
Ayam cabe ijo	0.000000	1.333333	-0.666667	0.000000	-0.666667
Udang goreng tepung	1.000000	0.000000	1.000000	0.000000	-2.000000
Soto betawi	-0.250000	0.000000	0.750000	-1.250000	0.750000
Soto medan	-0.750000	0.250000	0.250000	0.000000	0.250000
Bakso ikan	0.000000	0.750000	-1.250000	-0.250000	0.750000
Tomyam	0.666667	0.000000	0.666667	-1.333333	0.000000
Sate taichan	0.750000	-0.250000	0.000000	-1.250000	0.750000
Nasi padang	0.333333	00.000000	0.333333	-0.666667	0.750000
Rice bowl chicken katsu	0.750000	-0.250000	-0.250000	0.000000	-0.250000

Next, calculate the similarity for each item using the cosine similarity formula. The calculation results can be seen in Table 3.

Tabel 3. The results of the similarity for each item.

	Sapi lada hitam	Ayam cabe ijo	Udang goreng tepung	Soto betawi	Soto medan	Bakso ikan	Tomyam	Sate taichan	Nasi padang	Rice bowl

Sapi lada hitam	1.000	0.000	0.1889	0.866	0.000	1.000	1.000	1.000	1.000	0.000
Ayam cabe ijo	0.000	1.000	0.1889	0.000	0.000	0.500	0.0000	1.000	0.0000	0.000
Udang goreng tepung	-0.188	0.188	1.0000	-0.246	-0.471	-0.760	1.000	-0.207	1.000	0.471
Soto betawi	0.8660	0.0000	0.2461	1.0000	1.0000	0.000	0.8660	0.8660	0.8660	-1.000
Soto medan	0.000	0.000	-0.760	1.000	1.000	0.000	0.000	-0.500	0.000	-1.000
Bakso ikan	1.000	0.500	1.0000	0.000	0.000	1.000	-1.000	0.8660	-1.000	0.000
Tomyam	1.000	0.000	0.2075	0.866	0.000	-1.000	1.000	1.000	1.000	0.000

After obtaining the similarity scores for each item, the next step is to predict the empty items using the Weighted Sum formula. The prediction results can be seen in Table 4.

Tabel 4. Predictions for items that haven't been rated yet.

	User 1	User 2	User 3	User 4	User 5
Sapi lada hitam	4	3.9116667	3.766667	3	4
Ayam cabe ijo	3.6666667	4	2	2.486667	2
Udang goreng tepung	5	4	5	4	2
Soto betawi	4	4.270000	5	3	5
Soto medan	4	5	5	3.500000	5
Bakso ikan	3.780000	4	2	3	4
Tomyam	5	4.213333	5	3	4.263333
Sate taichan	5	4	4.350000	3	5
Nasi padang	5	4.546667	5	4	4.596667
Rice bowl chicken katsu	5	4	4	3.610000	4

Once the prediction results are obtained, error calculation is performed to determine how accurate the predicted values are compared to the actual values using the Mean Absolute Error formula. The MAE results can be seen in Table 5.

Tabel 5. The Mean Absolute Error (MAE) calculation results

Item ID	MAE

Item 0	0.16083315
Item 1	-0.576667
Item 3	0.7300000
Item 4	0.5000000
Item 5	1.2200000
Item 6	-0.2383333
Item 7	0.350000
Item 8	-0.0716667
Item 9	0.610000

Recommendation results were obtained using Collaborative filtering. The recommendation results can be seen in Table 6.

Tabel 6. The recommendation results using Collaborative Filtering.

Title	Rating prediksi	Kategori
Nasi padang	4.546667	Kuah, Nasi, Ayam, Sapi, Asin
Soto betawi	4.270000	Kuah, Sapi, Asin
Tomyam	4.213333	Kuah, Seafood, Asam
Sapi lada hitam	3.916667	Kering, Sapi, Pedas

Next, from the results of Collaborative filtering, calculations will be made to provide recommendations using Content-based filtering. The first step is to build a binary feature matrix. The result of the binary feature matrix can be seen in Table 7.

Tabel 7. The Binary Feature Matrix results.

	Asam	Asin	Ayam	Kering	Kuah	Nasi	Pedas	Sapi	Seafood
Nasi padang	0	1	1	0	1	1	0	1	0
Soto betawi	0	0	0	1	0	0	1	1	0
Tomyam	0	1	0	0	1	0	0	1	0
Sapi lada hitam	1	0	0	0	1	0	0	0	1

Once the matrix is obtained, similarity calculations can be performed for each item to determine how similar one item is to another using the Cosine Similarity formula, as shown in Table 8.

Tabel 8. The similarity of each item results.

	Nasi padang	Soto betawi	Tomyam	Sapi lada hitam
Nasi padang	1.000000	0.258199	0.774597	0.258199
Soto betawi	0.258199	1.000000	0.333333	0.000000

Tomyam	0.774597	0.333333	1.000000	0.333333
Sapi lada hitam	0.258199	0.000000	0.333333	1.000000

Once the equations are obtained, the similarity of each item with the user's preferences is calculated, resulting in recommendations using Content-based filtering, as shown in Table 9.

Tabel 9. The CBF recommendations results.

Item ID	Title
0	Nasi Padang
2	Soto Betawi

Respondent Results

In the example data that has been collected, recommendations are provided to users, and the results regarding how well the recommendations match their preferences can be seen in Figure 2.

Figure 2. Respondent Results

In Figure 2, it can be seen that 8 out of 10 people say that the recommendations provided match their food preference.

Conclusion and recommendations

Each paper ends with a conclusion, which summarizes the results of the paper written, as well as suggestions as recommendations resulting from the research.

From the results and discussions conducted, the following conclusions can be drawn:

- The number of items rated by users affects how accurate the recommendations are.
- Using the cascade technique, approaches with CF-CBF and CBF-CF will produce different recommendations because the first method used will influence the second method.

- Combining these two methods produces better recommendations than using only one of the methods.
- It cannot address the problem of data sparsity because the first method used is Collaborative Filtering.

Suggestions that can be considered for further research include the use of hybrid filtering methods with other techniques that can address the issue of data sparsity and produce even better recommendations

Acknowlegdement
Researchers would like to thank Tarumanagara University for supporting some of the funding for this research. Researchers also thank everyone organizations and participants taking part in this research study.

REFERENCE

[1] Arfisko, H. H., & Wibowo, A. T. (2022). Sistem Rekomendasi Film Menggunakan Metode Hybrid Collaborative Filtering Dan Content-based Filtering. *EProceedings of Engineering*. Retrieved from https://openlibrarypublications.telkomuniversity.ac.id/index.php/engineering/article/view/18066

[2] B. Walek and V. Fojtik. (2020, November). A hybrid recommender system for recommending relevant movies using an expert system. *Expert Systems with Applications*, 113452–113452. doi:https://doi.org/10.1016/j.eswa.2020.113452

[3] G. Geetha, M. Safa, C. Fancy, and D. Saranya. (2018). A Hybrid Approach using Collaborative filtering and Content based Filtering for Recommender System,. *Journal of physics*, 012101–012101. doi:https://doi.org/10.1088/1742-6596/1000/1/012101

[4] H.-Q. Do, T.-H. Le, and B.-S. Yoon,. (2020, February). Dynamic Weighted Hybrid Recommender Systems,. doi:https://doi.org/10.23919/icact48636.2020.9061465

[5] Mohamed, M. H., & I. M. (2019). Recommender Systems Challenges and Solutions Survey. *International Conference on Innovative Trends in Computer Engineering (ITCE)*. doi:10.1109/itce.2019.8646645

[6] . Haubner and T. Setzer,. (2020). *Applying Optimal Weight Combination in Hybrid Recommender Systems,*. Retrieved from ResearchGate:

https://www.researchgate.net/publication/339021993_Applying_Optimal_Weight_Combination_in_Hybrid_Recommender_Systems

[7] Prasetio, N. (2019, July 11). *Recommendation System Dengan Python : Definisi (Part 1). Medium*. (Data Folks Indonesia) Retrieved from https://medium.com/data-folks-indonesia/recommendation-system-dengan-python-definisi-part-1-71154dc3f700

[8] qutbuddin, . (2020, March 7). *Comprehensive Guide on Item Based Collaborative Filtering. Medium*. Retrieved from Towards Data Science: https://towardsdatascience.com/comprehensive-guide-on-item-based-recommendation-systems-d67e40e2b75d

[9] Utami, S. (2018). Kuliner Sebagai Identitas Budaya: Perspektif Komunikasi Lintas Budaya. *Journal of Strategic Communication*, 36–44.

[10] Y. Koren, S. Rendle, and R. M. Bell,. (2021). *Advances in Collaborative Filtering,*. Springer eBooks. doi:https://doi.org/10.1007/978-1-0716-2197-4_3