Food Recommendation Systems Based On Content-based and Collaborative Filtering Techniques

Conference Paper · October 2023

DOI: 10.1109/ICCCNT56998.2023.10307080

2 authors:

Reetu Singh
Motilal Nehru National Institute of Technology

Pragya Dwivedi
Motilal Nehru National Institute of Technology

Food Recommendation Systems Based On Content-based and Collaborative Filtering Techniques

1st Reetu Singh
Research Scholar
Computer Science and Engineering Department
Motilal Nehru National Institute of Technology
Allahabad, Prayagraj,(Uttar Pradesh), India
reetu.2021rcs55@mnnit.ac.in

2nd Pragya Dwivedi
Assistant Professor
Computer Science and Engineering Department
Motilal Nehru National Institute of Technology
Allahabad, Prayagraj,(Uttar Pradesh), India
pragyadwi86@mnnit.ac.in

***Abstract*—On the internet, numerous options are available for a specific type of product. It is tough to manually go through every product in a particular type when a user is trying to choose the best one. Because of this, manual searching is not very efficient. The recommendation system is crucial in recommending the best product in that situation. A food recommendation system is developed in this research paper using K-nearest neighbor's methods. The food data set is taken from Kaggle. We used Python programming language to implement the system. Our proposed recommendation system recommends food based on food name, food id, cuisine type, diet type like veg or non-veg in the case of content based filtering recommendation. For recommending the food with help of collaborative filtering we have used user id, food id and rating as an attributes.**

***Index Terms*—recommender systems, k-nearest neighbors, food recommendation systems, content-based filtering, collaborative filtering**

I. Introduction

By selecting user-related data from a wide range of data, recommender systems are simple algorithms that give users recommendations for products relevant to their needs. This system identifies data patterns in the data set by learning user preferences. It provides them with results relevant to their needs and interests. The internet has now become an irreplaceable part of modern life. Nowadays, users are flooded with options, and much information is available for anything from finding a hotel to making intelligent financial decisions [1]. Hence, companies have created recommendation systems to help users deal with the information explosion. Research regarding these has been going on for many decades because they can be applied in different areas, which is being used everywhere. As a result, many companies are interested in recommender systems to offer personalized options that meet users' preferences. The users accept the recommendations at their own risk and provide, immediately or later, implicit and explicit feedback. The users' activities and feedback can be stored in the recommender database and later used to provide new recommendations in the subsequent user-system interactions. In order to improve the user experience, personalization is a crucial method to facilitate. Recommendation Systems are prevalent in many web domains, including e-commerce or media sites, and have significantly enhanced business and decision-making in various information access systems.

Based on the types of input data, recommendation models generally fall into collaborative filtering, content-based recommender systems, and hybrid recommender systems. Recommendation systems have historically been developed across many fields for various machine-learning applications. Social networking sites like Facebook illustrate this, where user connections are suggested via recommendation systems; for music and media services like iTunes and Spotify, the same machine and recommendation logic are used to recommend various tunes to users based on their previous selection and preferences [2]. In keeping with this basic concept, the proposed system recommends food to users which puts it one step ahead of current recommendation applications.

Food recommender systems are software applications that provide personalized recommendations and suggestions for food items based on users' preferences, dietary needs, and other relevant factors.With the increasing availability of diverse food options, it can be overwhelming for users to navigate through the vast selection and make informed decisions. Food recommender systems are one of the best solutions in this environment by analyzing user preferences, dietary needs, and other relevant factors to generate tailored recommendations.

The proposed food recommender systems recommend food to users based on food name, food id, cuisine type, diet type like veg or non veg in the case of content based filtering recommendation. For recommending the food with help of collaborative filtering we have used user id, food id and rating as an attributes.

The paper is organized as follows: Section 2 provides related work that is relevant to our proposed approach. In Section 3, we introduce our proposed approach for generating recommendations. Section 4 details the computational experiments conducted and presents the corresponding results. Lastly, in Section 5, we present the conclusion of our work along with potential future research directions.

II. Related Work

Any system that generates personalized recommendations as an output or directs the user in a tailored manner towards interesting or practical things from a wide range of available possibilities is referred to as a recommender system [1]. Systems that provide recommendations analyze user preferences for objects and suggest products that users might proactively enjoy. Collaborative filtration, content-based, and hybrid systems of recommendation are the three primary categories under which models of recommendations fall. Collaborative filtering suggests learning from past user interactions, whether explicit (like user ratings) or implicit (like browsing history). Most content-based recommendations are based on comparisons between the auxiliary data of products and users. Text, images, and videos are just a few examples of additional information that could be analyzed. Numerous systems have been proposed in the literature, some of which are detailed below:

Using sentimental analysis, Petrusel ,Renata and Sergiu-George [3] developed a model to propose a restaurant based on positive and negative customer feedback. They combined the outputs of a recommendation system and a sentimental analysis system to create a model to recommend items effectively and efficiently. The recommendation system uses the results of the sentimental analysis to recommend restaurants to other customers using a collaborative filtering method. Sumedh proposed a Yelp food recommendation system based on customer restaurant reviews [4]. They retrieved collaborative and content-based features from the Yelp dataset to identify customer and restaurant profiles. The algorithm that recommends restaurants based on customer preferences and the location was developed using K-nearest neighbor clustering. A food recommendation system based on nutrition and calories was introduced by R. Yera et al. [5]. The research focused primarily on the users' diets rather than the restaurant's famous cuisine, flavor, or quality.

Rui Maia and Joao et al. [6] conducted a study to recommend restaurants and food based on reviews from previous users. When a new user moves by a nearby location, the context-aware method is used to identify the restaurants that are close by and perform collaborative filtering on food items to predict which items are most popular based on collaborative ratings on food items. Li Chen et al. proposed restaurant recommendation system [7] that was developed based on the opinions of the visited users. This study developed a review-based restaurant recommendation algorithm using collaborative filtering, preference-based product rating, and content-based approaches. When customers give a preferred rating, all restaurants with that rating are presented. This technique is implemented by using text analysis of customer evaluations.

However, this cannot simultaneously sort restaurants based on location and rating. Saito, Asada, Ysohitomi, et. al. [8] proposed a system based on hybrid filtering that recommends recipes based on last night's dinner recipes, user-selected impression words (sweet, warm, spicy), past successful recommendations, and same-rated recipes.M. Gupta et.al. [9] tried to discover the user's interests and made recommendations for personalizing recommendations to match the user's mood using collaborative filtering. The algorithm can also be enhanced to identify facial expressions, predict a user's mood, and make food recommendations based on their preferences. A restaurant recommendation system based on opinion mining of unstructured customer feedback, such as emojis and gifs, is proposed in work by Khan et al. [10] .The study's findings were peculiar because they converted each unstructured opinion into information that can be utilized for rating restaurants.

Much research was carried out, and many models were built for recommending food and restaurants based on factors like past orders (content-based filtering), ratings (collaborative filtering), location, nutrition, etc. However, the research still needs to develop a model for recommending food based on the content based and collaborative filtering. The existing models have some gaps because they cannot include different features during the prediction.

III. Proposed Systems

In our proposed work, we have included three features C_type (Cuisine type), Veg_Non (Veg and Non-Veg category of food), and Describe (Description related to that food). By including these features, our recommender model will become better compared to the existing one, these features C_Type', 'Veg_Non,' and 'Describe, to make our recommendation better. On the other hand, for recommending food with the help of collaborative filtering, we have used nearest neighbors algorithms, and for calculating the nearest neighbors, we have used cosine similarity.

A. Dataset

For evaluating the performance of recommender model we have used two datasets. The utilization of these datasets is illustrated in Figure 1. For evaluating the performance of recommender model we have used two datasets. The utilization of these datasets is illustrated in Figure 1. In this research study,

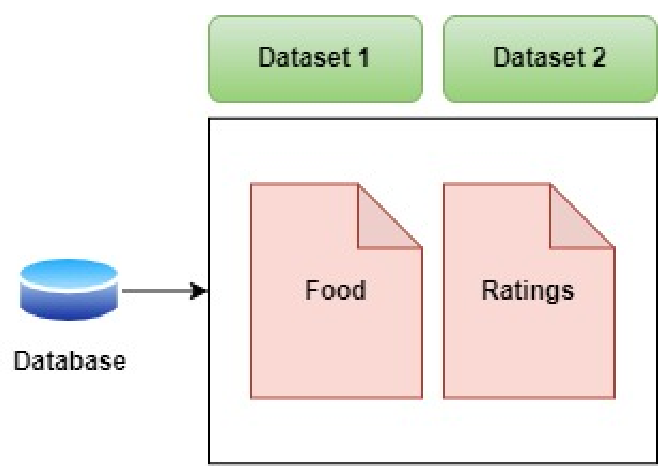

Fig. 1. Food dataset

datasets obtained from Kaggle were utilized to construct a model and develop item-based and rating-based recommendations for customers. The dataset comprised of two files: ”food.csv” and ”ratings.csv”. These datasets were employed to enhance the performance of the recommendation systems. The ”food.csv” dataset included 401 samples, containing attributes such as Food_id, Name, C_type, veg_non, and describe. All of these attributes were selected and used in the analysis to improve the recommendation systems. The dataset provided valuable information about different food items and their characteristics. Additionally, the ”ratings.csv” dataset consisted of 512 samples with features including User_ID, Food_ID, and Rating. Prior to constructing the model and developing the recommendation system, relevant attributes were carefully selected from both the dataset. This selection process ensured that the most relevant and informative attributes were utilized in order to enhance the accuracy and effectiveness of the recommendation system. By combining the attributes from both the datasets and selectively choosing relevant attributes, the research aimed to optimize the recommendation systems and provide more accurate and personalized recommendations to users. To provide a visual representation of the ratings distribution in the ”ratings.csv” dataset, we have included Figure 2.

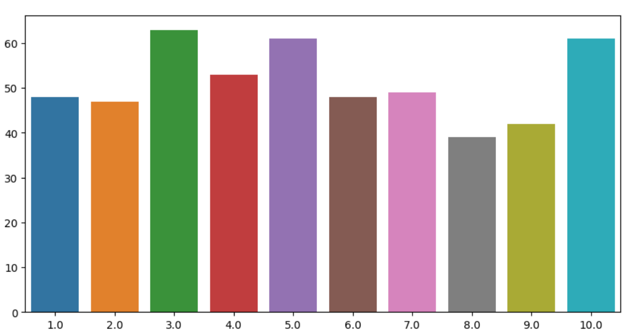

Fig. 2. Rating distribution

B. Classifier

A classifier is an algorithm that automatically assigns data into specific categories or classes. It utilizes predefined rules or patterns to make these categorizations. The result of the machine learning process applied to the classifier is a classification model. This model serves as the foundation for performing data classification. In the proposed recommendation system, the K-nearest neighbors classifier is employed [11]. By utilizing this technique, we provide the system with data and obtain a precise model that enables accurate classification. The nearest neighbors algorithm is a simple yet effective approach that relies on local patterns in the data to make predictions or recommendations. It is a non-parametric and lazy learning algorithm, means it does not explicitly build a model during the training phase but directly uses the training instances for prediction. In the nearest neighbors algorithm, the similarity between data points is determined by measuring the distance between them using a chosen distance metric, such as Euclidean distance or cosine similarity as shown in figure 3.

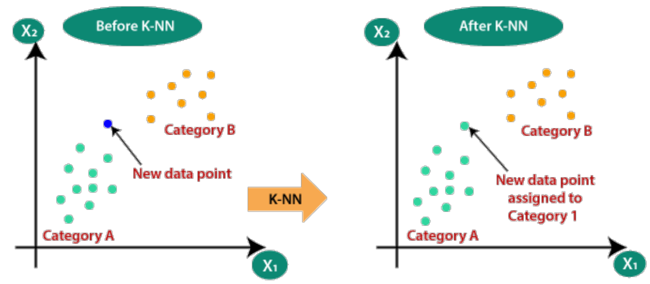

Fig. 3. K-nearest-neighbor-algorithm

In recommendation systems, the nearest neighbors algorithm is used to identify users or items that are similar to a target user or item. By analyzing the preferences or characteristics of these nearest neighbors, personalized recommendations can be generated.

C. Cosine similarity

Cosine similarity is a measure used to determine the similarity between two vectors in a high-dimensional space. In the context of food recommendation, cosine similarity is commonly applied to calculate the similarity between the preferences or characteristics of different food items or user profiles. The cosine similarity between two vectors, A and B, is calculated using the following formula as reflected in equation 1.

$$cosine\_similarity(A, B) = \frac{A \cdot B}{||A||||B||} \tag{1}$$

A.B represents the dot product of vectors A and B, which is the sum of the element-wise multiplication of their corresponding components.
|| A|| and || B|| denote the Euclidean norms (also known as the magnitude or length) of vectors A and B, respectively. To apply cosine similarity in food recommendation, the vectors A and B can represent the preferences or characteristics of two food items or user profiles.
The elements of the vectors could correspond to attributes like taste, ingredients, nutritional values, or any other relevant features. By calculating the cosine similarity between food items or user profiles, it is possible to determine how similar their preferences or characteristics are. This similarity information can then be used to generate recommendations, such as suggesting food items that are similar to the ones a user has enjoyed in the past or identifying users with similar preferences for collaborative filtering-based recommendations.

D. Block Diagram

The general block diagram of proposed system is mentioned in figure 4, with the help of this block diagram, we try to explain how our proposed system works on content based filtering and collaborative filtering. The steps are listed below.

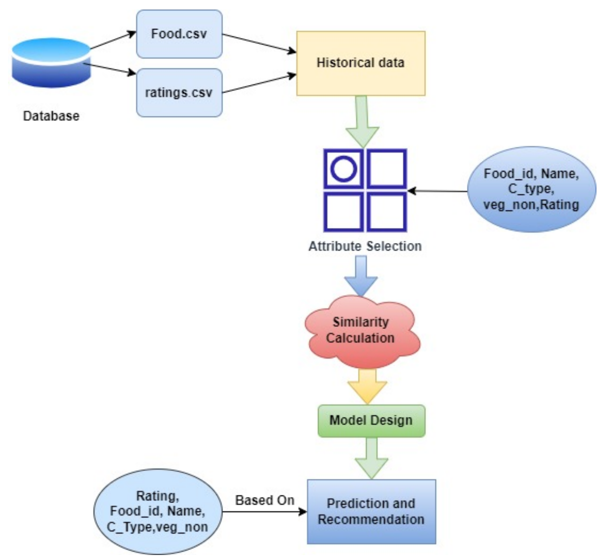

Fig. 4. Block diagram of proposed system

*1) **Food recommendation by using content-based filtering**:* We implemented the following steps, combining user-specific attributes and string matching, to create a personalized recommendation system for food items:

Step 1: We compiled a list of user-specific attributes for each food item, including the food name, ID, C_type, and veg _non. These attributes provided essential information about the food items.

Step 2: We incorporated string matching techniques to preprocess and analyze the textual data associated with food items. This involved removing punctuation, converting to lowercase, and eliminating stop words.

Step 3: We also applied string matching algorithm cosine similarity distance to compare the textual attributes between food items.

Step 4: Based on the preprocessed textual data, we constructed a feature vector for each food item, incorporating both numerical attributes and the results of the string matching algorithms.

Step 5: This feature vector represented the characteristics of the food item, capturing both the user-specific attributes and the similarity scores obtained through string matching.

Step 6: To determine the similarities between different food items, we utilized the cosine similarity measure on the feature vectors. This measure quantified the similarity between two feature vectors, considering both the user-specific attributes and the string matching results.

Step 7: We sorted the food items based on the calculated similarity scores, taking into account both the user-specific attributes and the textual similarities obtained through string matching. This sorting process allowed us to prioritize the most similar food items and rank them accordingly.

Step 8: Finally, based on the sorted list, we generated a list of suggested food items for the user.

By integrating string matching concepts into the content-based filtering algorithm, we enhanced the recommendation system's ability to capture textual similarities and provide more accurate and personalized food recommendations based on user preferences.

*2) **Food recommendation by using collaborative filtering**:* We implemented the following steps to develop a recommendation system based on user ratings and to incorporate food names for prediction:

Step 1: We gathered relevant information such as user_id, food_id, and ratings from the dataset. Additionally, we collected food names associated with each food _id to enrich the recommendation process.

Step 2: Using the collected data, we created a matrix that included user_id, food_id, and corresponding ratings. This matrix served as the foundation for capturing user-item interactions and preferences.

Step 3: To determine the similarity between user_id and food_id, we employed cosine similarity. In addition to considering the ratings, we also incorporated the textual information of food names by converting them into feature vectors.

Step 4: For each user_id, we identified the nearest neighbors based on the calculated cosine similarity. These nearest neighbors consisted of other users who exhibited similar rating patterns and had expressed preferences for similar food items.

Step 5: To generate recommendations for each user_id, we leveraged the ratings and food names of the nearest neighbors. We aggregated the ratings of the nearest neighbors and considered the food names associated with those ratings. This approach allowed us to incorporate both the numerical ratings and the textual information to make personalized recommendations.

Step 6: In the final step, we sorted the recommendations for each user_id based on the average ratings, taking into account the food names as well. This sorting process ensured that the recommendations were presented in a descending order of preference, with the most highly rated and relevant food items appearing at the top.

By relating food IDs to their corresponding food names and incorporating them into the collaborative filtering algorithm, we enhanced the recommendation system's ability to provide personalized and context-aware food recommendations to users based on their ratings and the textual information associated with food items.

IV. Experiments and Results

The result of experiments performed on both the filtering techniques are shown below-

A. *Content-based Filtering*

Content-based filtering for food recommendation using string matching concept involves utilizing the textual information associated with food items to identify similarities and make recommendations based on these similarities.

The results obtained from the content-based filtering method are mentioned in figure 5.

```
# Content-based Filtering food recommendation
get_recommendations('chicken biryani', cosine_sim2)
```

```
95       methi chicken masala
99        spicy chicken curry
259       quick chicken curry
108    chicken quinoa biryani
261            chicken masala
```

Fig. 5. Recommended food for Chicken biryani

B. Collaborative filtering

Collaborative filtering is a recommendation technique that leverages user-item interactions to generate personalized recommendations. In the context of food recommendation, mapping the food ID to food names enhances the collaborative filtering approach by incorporating the textual information of food items. The results obtained from the collaborative filtering method are mentioned in figure 6.

```
# Get recommendations by collaborative filtering
Get_Recommendations('chicken biryani')
```

	Food_ID	index	Name	C_Type	Veg_Non	Describe
0	182.0	181	avial with red rice	Indian	veg	Red rice water potatoes carrots raw banana dru...
1	60.0	59	caramelized sesame smoked almonds	Snack	veg	red lentils or masoor dal halfboiled potato gr...
2	180.0	179	lotus leaf wrapped fried rice	Chinese	veg	Jasmine Rice Baked Edamame Beans Mock Meat Shi...
3	244.0	243	jalebi with fennel yogurt pudding	Dessert	veg	all purpose flour yogurt oil sugar water saffr...
4	184.0	183	vegetable bruschetta	Italian	veg	baguette grilled slices black olive tapenade a...

Fig. 6. Recommended food for Chicken biryani

The above result screenshot description column is unclear, so we have taken another screenshot to clearly show all the columns in the resulting output mentioned in figure 7.

```
# Get recommendations by collaborative filtering
Get_Recommendations('chicken biryani')
```

level_0	Food_ID	index	Name	C_Type	Veg_Non	Describe
0	182.0	181	avial with red rice	Indian	veg	Red rice water potatoes carrots raw banana drumstick small raw mango sour curd bean stick onion salt turmeric water coconut oil green chillies mustard seeds crushed
1	60.0	59	caramelized sesame smoked almonds	Snack	veg	red lentils or masoor dal halfboiled potato grated carrot french beans bread slices ground chickpea flour sattu green chillies ginger onions garlic minced salt sugar chaat masala red chilli powder garam masala corn flour besan bread crumbs for crunchyness coriander refined oil for fryingp
2	180.0	179	lotus leaf wrapped fried rice	Chinese	veg	Jasmine Rice Baked Edamame Beans Mock Meat Shitake Mushroom Spring Onion Dark Soy Sauce Sunflower Oil Salt
3	244.0	243	jalebi with fennel yogurt pudding	Dessert	veg	all purpose flour yogurt oil sugar water saffron green cardamom yogurt strained milk warm sugar nutmeg cardamom powder
4	184.0	183	vegetable bruschetta	Italian	veg	baguette grilled slices black olive tapenade artichoke hearts lettuce arugula trummed tomato confit fresh basil leaves mint leaves zucchini goat cheese parmesan cheese shavings mozzarella buffalo cheese

Fig. 7. Recommended food for Chicken biryani

V. Conclusion and Future Scope

The food recommender system, based on a content-based and collaborative filtering technique, has been effectively implemented in this study. By utilizing the datasets obtained from Kaggle, we developed a model that successfully recommends food items to users based on their preferences and ratings.

In our approach, we incorporated string matching to enhance the recommendation process, ensuring that the food items suggested were closely related to the user's preferences. Furthermore, we considered multiple attributes, such as food name, ID, C_type, and veg_non, to understand each food item and provide more accurate recommendations comprehensively.

While the proposed model achieved the desired output of recommending foods based on user preferences and ratings, it is essential to acknowledge its limitations. First, the model heavily relies on the available datasets, and the coverage and representativeness of the data may limit the recommendations. Additionally, the model primarily focuses on user-specific attributes and collaborative filtering, potentially overlooking other relevant factors such as social influence or temporal dynamics.

As part of future work, we aim to expand the capabilities of the food recommender system by incorporating additional criteria and attributes for a recommendation. These criteria may include nutritional information, ingredient preferences, dietary restrictions, or user demographics. By combining a more diverse range of criteria, we can offer users more personalized and tailored recommendations.

Furthermore, we plan to evaluate the performance of the proposed model by calculating error values like root mean square error (RMSE) and mean absolute error (MAE) between the predicted and actual ratings. This evaluation will allow us to quantify the accuracy of the recommendations and make improvements based on the obtained error values.

References

[1] J. Bobadilla, F. Ortega, A. Hernando, and A. Gutiérrez, "Recommender systems survey," *Knowledge-based systems*, vol. 46, pp. 109–132, 2013.

[2] Y. Gao, W. Yu, P. Chao, R. Zhang, A. Zhou, and X. Yang, "A restaurant recommendation system by analyzing ratings and aspects in reviews," in *Database Systems for Advanced Applications: 20th International Conference, DASFAA 2015, Hanoi, Vietnam, April 20-23, 2015, Proceedings, Part II 20*. Springer, 2015, pp. 526–530.

[3] M.-R. Petrusel and S.-G. Limboi, "A restaurants recommendation system: Improving rating predictions using sentiment analysis," in *2019 21st International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)*. IEEE, 2019, pp. 190–197.

[4] S. Sawant and G. Pai, "Yelp food recommendation system," 2013.

[5] R. Y. Toledo, A. A. Alzahrani, and L. Martinez, "A food recommender system considering nutritional information and user preferences," *IEEE Access*, vol. 7, pp. 96 695–96 711, 2019.

[6] R. Maia and J. C. Ferreira, "Context-aware food recommendation system," *Context-aware food recommendation system*, pp. 349–356, 2018.

[7] L. Chen, G. Chen, and F. Wang, "Recommender systems based on user reviews: the state of the art," *User Modeling and User-Adapted Interaction*, vol. 25, pp. 99–154, 2015.

[8] K. Saito, T. Asada, Y. Yoshitomi, R. Kato, and M. Tabuse, "A recipe decision support system using knowledge information and agent," *Journal of Robotics, Networking and Artificial Life*, vol. 5, no. 3, pp. 204–207, 2018.

[9] M. Gupta, S. Mourila, S. Kotte, and K. B. Chandra, "Mood based food recommendation system," in *2021 Asian Conference on Innovation in Technology (ASIANCON)*. IEEE, 2021, pp. 1–6.

[10] K. Khan, B. Baharudin, A. Khan, and A. Ullah, "Mining opinion components from unstructured reviews: A review," *Journal of King Saud University-Computer and Information Sciences*, vol. 26, no. 3, pp. 258–275, 2014.

[11] R. A. Sowah, A. A. Bampoe-Addo, S. K. Armoo, F. K. Saalia, F. Gatsi, and B. Sarkodie-Mensah, "Design and development of diabetes management system using machine learning," *International journal of telemedicine and applications*, vol. 2020, 2020.