

Recipe Recommendation System Using TF-IDF

Shubham Chhipa¹, Vishal Berwal², Tushar Hirapure³ and Soumi Banerjee⁴

¹Department of Information Technology, Ramrao Adik Institute of Technology, Nerul, Navi Mumbai.

² Department of Information Technology, Ramrao Adik Institute of Technology, Nerul, Navi Mumbai.

³ Department of Information Technology, Ramrao Adik Institute of Technology, Nerul, Navi Mumbai.

⁴ Department of Information Technology, Ramrao Adik Institute of Technology, D.Y. Patil Deemed to be University, Navi Mumbai.

Abstract - A Recipe Recommendation System is being proposed in this following paper. Food recommendation is a new area, with few systems that are focus on analysing and user preferences and constraints such as ingredients available at their side being deployed in real settings in the form of web application or mobile application [4]. The proposed model is a mobile application which allows users to search recipes using ingredients available at them including vegetables. For this work we have find a dataset which is a collection of Indian cuisines recipes and apply the content-based recommendation using Term Frequency – Inverse Document Frequency (TF-IDF) and Cosine Similarity [1]. This application gives the recommendation of Indian recipes based on ingredients available at them and allows users to filter out the recipes on course type, diet type, etc.

Keywords— Machine Learning, Recommendation System, TF-IDF, Cosine Similarity, bag-of-words, NLP, Flask.

1. Introduction

People make decisions related to food every day. They all think about what to eat, what to cook, how much nutritional value this food has, can this food make me healthy i.e., is it gluten free or sugar free, is it diabetic friendly or not and other questions. Recommendation systems help the user to make fast decisions in these complex information spaces.

Diet management systems, which have mostly replaced conventional paper-and-pen approaches, have gotten a lot of attention. These systems leverage educational information and services to persuade people to change their habits. Because of the popularity of these diet monitoring capabilities, these systems have a wealth of user preference data, which might be used to personalise interactive elements and boost user engagement with the system and diet programme. A meal recommender is one such tailored service that is great for educating diet. The nutritional properties of the food could be used by this recommender to inform its recommendations.

Recommender systems are one of the most widely used data science applications today. They are used to anticipate a user's "rating" or "preference" for a certain

item. Almost every big IT firm has used them in some capacity. It's used by Amazon to recommend products to customers, YouTube to determine which video to play next on auto-play, and Facebook to suggest pages to like and people to follow [5].

For as long as people have existed, the challenge of what to cook/how to cook delicious food and make it nutritious has existed. For thousands of years, the only way to learn about amazing dishes and recipes was through trial and error or word of mouth, and neither approach took into account the nutritional content of a dish. [3]

Thinking about “**What to cook today?**” is a problem that every housewife and working women face. It seems to be easy decision but it requires some strategic and smart decisions when it comes to decide heavy or light food such that their family members digestion and health doesn't get disturbed and face acidity and gas issues and they also have to think about their budget.

The proposed application helps users to cook recipes from available ingredients at their end and filter the recipes based on course type, diet type, etc.

2. Literature Survey

2.1. Recommendation of Indian Cuisine Recipes based on Ingredients-Nilesh, Vishal, Pritam, Dr. Madhu. [1]

They have created datasets by using web scraping and extract features from ingredients column using bag of words. Bag of Words contain the specific keywords for each recipe and for recommendation they have used cosine similarity between these keywords.

2.2. Recipe Recommender System Using Image Recognition of Food Ingredients – Mona Mishra, Yifan Gong. [2]

They have picked up twenty-two fruits and vegetables categories from ImageNet. They have created 5layer CNN model for image recognition. They have used Microsoft Azure for cleaning the dataset and then searching the ingredients and sends the response containing the recipes.

2.3. Healthy Recipe Recommendation using Nutrition and Ratings Models - Yew Siang Tang, Anita Hanzhi Zheng, Nicholas Lai. [3]

They have used dataset from Kaggle which contains 50,000 recipes which is scraped by allrecipes.com. They have created two models Nutritional Model and Ratings Model. They have used GNN for prediction of the recipes.

2.4. Personalized Food Recommendations - Jorge Imeida, Prof. Doutor Anibal Cavaco Silva, Technical University of Lisbon.[4]

They have used Yolp Content-based component and Yolp Collaborative Component. They have used Rochios Algorithm using FF-IRF for extracting weights.

ingredients. In such cases, users find themselves shopping for ingredients or they decide to substitute the missing ingredient with something else.

In the proposed system the user can explore recipes based on different course types such as lunch, dinner, breakfast, snacks, etc. as well as different diet types such as gluten free, sugar free, diabetic friendly, etc. This system takes care of user's biggest constraints such as limited number of ingredients available at them.

The user have to enter or dictate the ingredients available at their end and then they have to click on see recipes button. Then app will make a request to the recommendation system passing the ingredients as a query parameters and then recommendation system sends the recipes that can be cooked using that ingredients to the app. Then app display all that recipes in a list format with the filter so that user can filter recipes based on their requirements.

Users can also create their shopping list for groceries in our app while cooking so that they don't procrastinate the task of making a list. This system also takes care of user's budget by creating multiple combinations of the ingredients and trying to give best suggestions / recipes so that their resources get used perfectly.

Fig. 1. displays the working of our system.

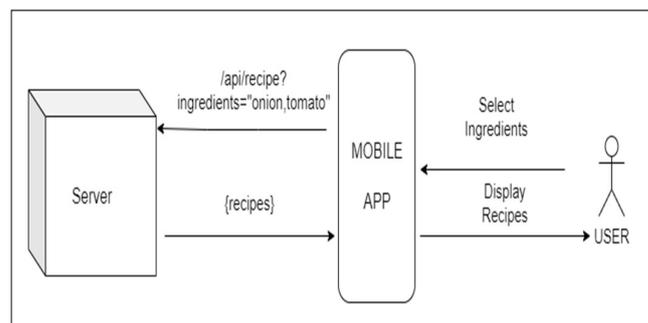


Fig. 1. Block Diagram of System.

3. Proposed System

This paper proposes a system which has greater accessibility as it is a mobile application and it is easy to use for users.

The existing system consists of approaches such as thinking what to cook today manually leading to stress, time consuming and worries about their budget for groceries. There are several cooking apps or websites available today, that are used to find recipes based on some keyword, like name of the food ingredient or type of cuisine, etc. These apps are mindful of the needs and interests of their users, but they fail at identifying their user's constraints, i.e., limited number of food

4. Proposed Methodology

4.1 Dataset:

This paper have used the dataset available on the Kaggle. The name of dataset is **6000+ Indian Food Recipes Dataset** by Kanishk. This dataset was selected because it contains all the recipes which are of Indian Palette and it is created by scraping the archanaskitchen.com website.

Since this dataset does not contain images of a recipes then we have created a script that scrape all the images of the recipe from the same website and adds them to a new column Images in the dataset and created a new dataset.

The dataset can be used to answer a lot of questions related to Food Recipes such as serving sizes, time required to prepare a dish, most common ingredients, different cuisines, diets, courses and what not.

4.2 Pre-processing and Parsing of Ingredients:

There was a lot of redundant information in this dataset for example weights and measures such as tbsp, grams, etc are not going to add value to the vector encodings of the recipes. There are also other words in the ingredients that are useless to us. These are the words that are very common among all the recipes such as oil, fresh, leaf, or any other words.

To remove all this information from the ingredients column in the dataset we have created **ingredient_parser** function. At first we are removing all the punctuations from the ingredients using **maketrans** function from **str** library. Then we have lemmatize the ingredients using **WordNetLemmatizer()** from **nlk** to get root of the different words. After creating **ingredient_parser** function we have used lambda functions which makes it easy to parse all the ingredients.

Fig. 2. shows the flowchart of the preprocessing method.

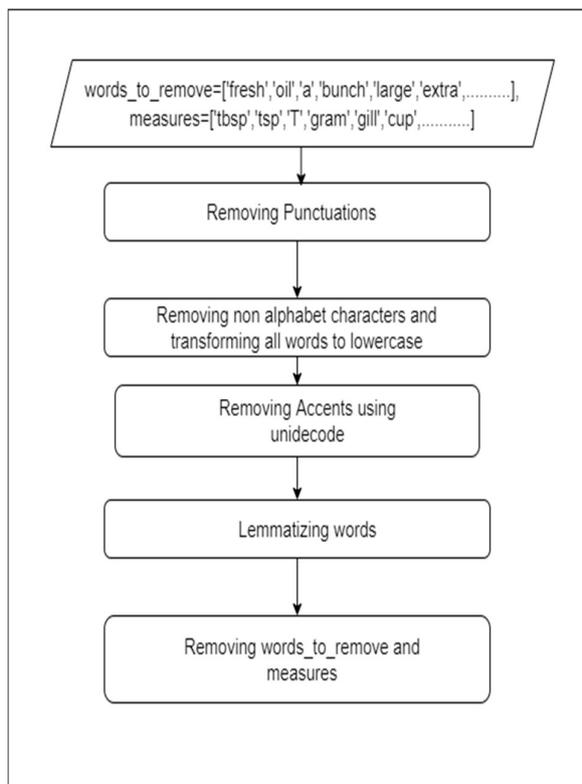


Fig. 2. Flowchart of Preprocessing Method.

4.3 Extracting Features

To encode each document (recipe ingredients) we have two choices NLP's bag of words and Term Frequency – Inverse Document Frequency (TF-IDF). Bag of Words performed ok but TF-IDF marginally out-performed it, so we opted for this instead.

The TF-IDF determines the relevance of a phrase by considering its importance across all documents. [6].

$$tfidf(t,d,D) = tf(t,d) \times idf(t,D) \tag{1}$$

- t = term
- d = document
- D = document

$$tf(t,d) = f_{t,d} / \sum_{t \in d} f_{t,d} \tag{2}$$

$$idf(t,D) = \log(|D| / | \{ d \in D : t \in d \} |) \tag{3}$$

The paper have used Scikit-learn **TfidfVectorizer** method. We then saved the model and encodings using pickle in two different files.

4.4 Recommendation System

This paper have used content-based filtering which enabled us to recommend recipes to people based on the attributes i.e. ingredients.

To measure the similarity between the ingredients provided by user and ingredients of a recipe we have used Cosine Similarity.

Cosine Similarity is a metric that is used to measure how similar two items are and generate a scores based on the similarity [7]. By using the cosine similarity function we get the the scores of the matching recipes and then that scores are passed to **getRecommendation** function to rank these scores and output a pandas containing all the details of top N recommendations.

We have created an API for our model using Python web framework Flask [8] and hosted it on pythonanywhere.

Fig. 3. show the flowchart of the recommendation system.

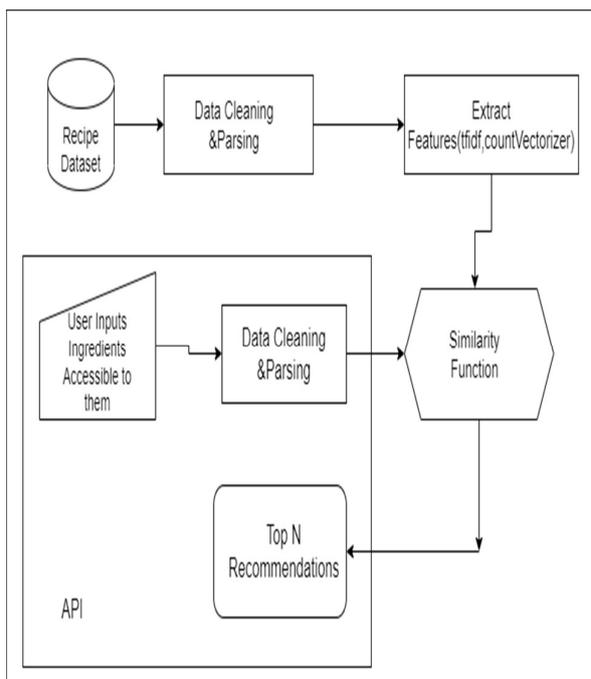


Fig. 3. Flowchart of recommendation system.

4.5 Developing An APP:

For developing an app for the users we have used React Native [9], which is open source cross platform UI Software framework developed by facebook.

This paper have also used React Native Paper [10] which is a UI library for react native and Redux for state management.

5. Results

In our evaluation we have tested the performance of our recommendation system by comparing the ingredients given by the users as input and the ingredients of the recipes that our recommendation system displays.

We have also taken the budget of the users into an account hence we create the powerset of all the ingredients given by the users to generate multiple combinations of different ingredients and then run our recommendation on each element or combination of a powerset.

Below is the table that shows the recommendations provided by our system on a certain input.

Users Input – Potato or Onion

Sr No	Recommendations
1	Aloo Ke Gutte Recipe

2	Konkani Style Batata Recipe
3	Parsi Style Chas Payelo Sakarkand Recipe
4	Niramish Aloo Recie
5	Cajun Style Potatoes Recipe
6	Spring Onion Recipe
7	Vaghrela Chawal Recipe
8	Odisha Style Aloo Recipe

Using Lunch Filter

Sr No	Recommendations
1	Aloo Ke Gutte Recipe
2	Niramish Aloo Recipe

Using Side Dish Filter

Sr No	Recommendations
1	Konkani Style Batata Recipe
2	Parsi Style Chas Payelo Sakarkand Recipe
3	Odisha Style Aloo Recipe

Users can also explore recipes based on different types of cuisines such as South Indian, North Indian, Italian, Chinese, Fusion.

On exploring Chinese cuisine type our recommendation system recommends recipes that are of Indo-Chinese taste. Our system recommends 48 recipes based on Chinese cuisines. Some of the recipes are shown below in the table.

Chinese Cuisine Recipes

Sr No	Recommendations
1	Idli Manchurian Recipe With Oriental Twist
2	Chilli Baby Corn Manchurian Recipe
3	Chilli Paneer Momo Recipe
4	Spicy Schezwaan Noodles

Users can also explore recipes based on their diet type such as Veg, Non Veg, Gluten Free, Diabetic Friendly and sugar free.

On exploring Diabetic Friendly Diet type our recommendation system recommends recipes that are diabetic friendly. Our system recommends 260 recipes based on Diabetic Friendly diet. Some of the recipes are shown below in the table.

Diabetic Friendly Recipes

Sr No	Recommendations
1	Masala Karela Recipe

2	Thalipeeth Recipe
3	Ragi Masala Roti
4	Kovakkai Podi Curry Recipe

All responses from our server is in JSON format and in app we have parsed and transformed these JSON. Users will see list of recommended recipe in the flatlist format in app.

6. Conclusion

The Indian cuisines are broadly refreshing for its remarkable utilization of herbs and spices. Mostly Indian Cuisines are result of an innovation/jugaad done by various Indian housewives and working class women cause they were having the constraints of limited ingredients and keeping in mind of the health of their family members.

This paper presented a method for recipe recommendation using ingredients matching of recipes. In this paper, we tried to solve the burning problem of Indian housewives or working women that is **What to cook today?** Through this paper we have learned to build a recommendation system using TF-IDF method and understands how similarity scores gets calculated using cosine similarity.

7. Future Scope

The current system allows users to directly explore all the recipes without login, in future we are planning to add registration and login functionalities. If there will be login functionality then we will be able to personalise all the user experience by recommending recipes of their flavours preferences and we will be able to store the bookmarks and likes of the user.

The current system allows two approaches for taking input from users which is by selecting ingredients from list and other is by dictating the ingredients one by one. In future we are planning to add image recognition of vegetables so that users can directly scan the ingredients and search for recipes. [2]

References

[1]. Recommendation of Indian Cuisine Recipes based on Ingredients - Nilesh, V. Raman, P. Hazarika, Dr. M. Kumari.
 [2]. Recipe recommendation system using Image recognition of Food Ingredients – M. Mishra and Y. Gong.
 [3]. Healthy Recipe Recommendation using Nutrition

and Ratings Models – Y. Tang, A. Zheng, N. Lai.

[4]. Personalised Food Recommendations – J. Almedia, D. Anibal.

[5]. <https://www.datacamp.com/community/tutorials/recommender-systems-python>

[6]. <https://towardsdatascience.com/tf-idf-explained-and-python-sklearn-implementation-b020c5e83275>

[7]. <https://towardsdatascience.com/using-cosine-similarity-to-build-a-movie-recommendation-system-ae7f20842599#:~:text=Cosine%20similarity%20is%20a%20metric,the%20items%20are%20100%25%20similar>

[8]. <https://flask.palletsprojects.com/en/2.1.x/>

[9]. <https://reactnative.dev/>

[10]. <https://callstack.github.io/react-native-paper/>