Recommendation System for Food Dishes in Specific Restaurants Based on Sentiment Analysis

MSc Research Project
Data Analytics

Sunanda Pal
Student ID: x21195820

School of Computing
National College of Ireland

Supervisor: Vitor Horta

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Sunanda Pal
Student ID:	x21195820
Programme:	MSc Data Analytics **Year:** 2023
Module:	MSc Research Project
Supervisor:	Vitor Horta
Submission Due Date:	18th August, 2023
Project Title:	Recommendation System for Food Dishes in Specific Restaurants Based on Sentiment Analysis
Word Count:	9364 **Page Count** 23

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Sunanda Pal

Date: 17th August, 2023

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	□
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	□
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	□

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Recommendation System for Food Dishes in Specific Restaurants Based on Sentiment Analysis

Sunanda Pal
x21195820

Abstract

In the food and restaurant industry, recommendation systems are immensely used to enhance the customer experience. Customers often leave comments to share their dining experiences, either on dedicated food review platforms or on social media. Researchers collect massive amounts of such relevant data to build advanced recommendation systems and improve upon the existing ones. Current research, however, primarily focuses on recommending restaurants based on client preferences or dishes based on higher restaurant reviews. Thus, in this study, a new kind of recommendation system is proposed that suggests certain dishes in a specific restaurant to customers based on their culinary preferences and similarity with other users. To accomplish this, first, NLP methods are used to extract dishes from reviews. Then, sentiment analysis is applied in order to detect which users like which dish and restaurant. Finally, collaborative filtering is used to recommend food and restaurants to the users. Due to the large volume of datasets and available resource restrictions, the model implementation is carried out in two ways. In the first approach, KNN-based algorithms (KNNBasic and KNNwithMeans) are trained on the sampled data, where hyperparameter-tuned KNNBasic with cross-validation has shown better accuracy. In the second approach, matrix factorization techniques are implemented on the entire data, where optimized SVD has performed better than NMF in terms of RMSE and MAE. The research also identifies the limitations of the proposed methodology and provides directions towards the scope of improvements.

Keywords: *Natural language processing, text-mining, food recommendation, aspect-based sentiment analysis, restaurant-feedbacks, collaborative filtering.*

1 Introduction

In recent years, food delivery applications and restaurant review websites have gotten more and more popular, and their use has become more widespread, especially with the advent of hybrid work arrangements wherein people do stay at home more often and enjoy good food, ordering them online (Teck-Chai Lau and David Ng, 2019) and providing invaluable feedback to such services. Also, to share their dining experiences, customers opt to rate and or comment on food they've tried at either restaurants or online. Researchers accumulate such voluminous data to develop recommendation engines that aid web applications, relying heavily on user recommendations, to operate. These are influenced by a variety of variables, such as previous purchases, reviews, and searches made by users. More often than not, people like exploring new cuisines (Ester Giacomoni, 2017). Based on their previous orders, ratings, or other data, food and restaurant suggestions can be provided precisely. It becomes crucial to provide pertinent and individualized information about food and restaurants to elevate their dining experience. Such efforts contribute to a happier consumer base, which increases patronage for restaurants and engagement for their web apps. Furthermore, decision-making is streamlined by a recommender system, saving consumers' time and efforts in addition to allowing diners to discover new culinary experiences and dining options. Overall, recommendation engines for restaurants expand businesses, boost sales, attract customers, and improve customer satisfaction.

There have been numerous studies on recommendation systems in the dining and food industries. Much of the initial research focused on collaborative filtering and user-rating-based recommendations. However, with increasing growth in available data, research focus has shifted towards incorporating sentiment analysis of opinions to develop enhanced recommendation engines. One such study conducted by Mohana et al. (2022) proposes a recommender system that suggests restaurants based on people's food preferences extracted from their remarks and then recommends nearby eateries based on how closely they match the user's preferences. Saha and Santra (2017) developed a recommendation system that suggests the top restaurants in a specific city by analyzing user comments and computing sentiment scores for food items and restaurants, followed by collaborative filtering.

The majority of the experiments focus on recommending restaurants to consumers by analyzing the sentiments of comments, followed by collaborative filtering. But there aren't many studies that recommend food dishes based on consumers' preferences. Ratnaparkhi (2018), however, in his research, proposes a recommendation system for food dishes based on sentiment analysis of user comments, where the dishes are retrieved from the reviews themselves. The system recommends restaurant-specific but not user-centric, highest-rated dishes at a restaurant.

In our study, we attempt to recommend specific dishes at specific restaurants based on sentiment analysis of customer feedback on the dishes and then determining the similarity between their preferences. Previous studies have demonstrated the possibility of extracting the names of food items and sentiments directly from user reviews (Ratnaparkhi, 2018). For each customer, the names of the dishes are extracted from the reviews in order to determine which types of dishes and restaurants they prefer. Sentiment scores are then computed by performing sentiment analysis on their comments. A final score for the food and restaurant pair is calculated using the sentiment score and user rating. After gathering customer preferences, collaborative filtering is performed on a user-by-user basis to identify which customers have similar preferences. Dishes and restaurants are suggested based on this data. This study is crucial since it considers client preferences and recommends not only a restaurant but also a specific dish from within that restaurant. The goal of our study is to enhance the consumer experience, which will ultimately benefit all the stakeholders involved.

The aim of the study is to provide an answer to the following question:

"How can a recommendation system be developed to suggest food items in particular restaurants for consumers using NLP and aspect-based sentiment analysis and applying collaborative filtering to utilize similarities between user preferences?"

The entire research paper is divided into a few sections elaborating on different aspects of the study. The first section is all about the background and objectives of the study, which highlights the research question. Section 2 provides a detailed analysis of the previous research works in this domain, which is further divided into a few subsections to shed light on different areas of research. Section 3 elaborates on the methodology followed in this paper, describing detailed steps from data collection, pre-processing, and feature generation to collaborative filtering, model application, and the process of evaluation. Section 4 provides insights about the underlying architecture of the research process and describes the algorithms used. Furthermore, the fifth section details the implementation process, and the sixth section contains the evaluation methods and discussion of the research. The paper ends with Section 7, which provides the conclusion, limitations, and direction for future study of the project work.

2 Related Work

Numerous studies have been done on sentiment analysis-based restaurant and food recommendation systems. The study's objective is to use feedback analysis to suggest specific food items at a particular restaurant to a customer. To improve the customer experience, prior studies in Natural Language Processing (NLP), aspect-based sentiment analysis, Machine Learning (ML) algorithms, Collaborative Filtering (CF), and recommendation systems are studied.

To gain a thorough understanding of the domain and ultimately provide an answer to the research question, the literature review is divided into a few subsections. Section 2.1 contains the analysis of studies related to restaurant recommendations, whereas Section 2.2 has a literature survey related to food recommendations. Subsection 2.3 has the analysis of the papers in the domain of aspect-based sentiment analysis, and subsection 2.4 contains the discussion of the studies related to collaborative filtering algorithms used for recommendation, especially in the restaurant and food domains. Finally, an overview of the findings is described, which provides insights into the importance of answering the research question.

2.1 Recommendation System for Restaurants

Owing to the popularity of social media platforms and food applications, it is highly likely for customers to post about their dining experiences, opinions on food and services, and other topics. Using text analytics, researchers can utilize this data to forecast restaurant or food reviews or to build recommendation systems. By analyzing user attitudes toward food and restaurants from comments, the authors (Santra et al., 2017) attempt to recommend top eateries in Kolkata, India. They have used collaborative filtering to rate and rank eateries after classifying user sentiment as either good, negative, or neutral to determine the overall rating. As the sentiment analysis is done on a review level, despite being a good suggestion strategy, it lacks a finer analysis of opinion. In a different study by Limboi et al. (2019), the researchers sought to forecast ratings for unvisited restaurants and suggest customers eat at the top-n eateries. The reviews were divided into two categories, namely, positive and negative. The use of sentiment analysis has been made to generate restaurant ratings. The sentiment score and restaurant ratings were then utilized as input to collaborative filtering, which produces a list of the top-n restaurants that the public should visit. This study demonstrates how to forecast a customer's possible rating for a restaurant they haven't yet visited using previous ratings and sentiment analysis from reviews, but it also lacks fine-grained sentiment analysis where different aspects can be examined. On the other hand, a few research studies by E. Asani et al. (2021) and Mohana et al. (2022) focus on building context-aware recommendation systems for restaurants, which emphasize the evaluation of user feedback and identification of food preferences to result in the delivery of customized recommendations. The preferences are derived from the reviews through clustering, feature selection from a previously established list of words, and estimating the sentiment scores for them. User comments are first pre-processed through tokenization, POS (Part Of Speech) tagging, stop-word removal, stemming, and noun word filtering. Then, to extract the food names, the WordNet[1] ontology is used, and clusters are formed for the food items. Next, a sentiment score is generated for each feature and cluster, and clusters with the highest scores are selected to determine the preference. By comparing the similarities between the customer’s food preferences and the menu being served at the restaurant, recommendations are made for open dining places while considering factors like opening hours and duration. Even though both papers follow a similar approach, in the study conducted by E. Asani et al. (2021), only the Wu-Palmer similarity module is used, whereas in the study by Mohana et al. (2022), authors have compared cosine and Jaccard similarity along with Wu-Palmer to obtain the best accuracy. Moreover, the suggestion method utilized in these two papers is different from that in the previous papers, and they also offer insight into how to determine customer meal preferences. Luo et al. (2020), in their study, explored the important aspects of user comments in improving the restaurant recommendation system. Five restaurant attributes, such as location, value, dining ambiance, food and drinks, and customer service, are analyzed through sentiment analysis. The significance of restaurant features and overall customer satisfaction through a generic number rating are analyzed to determine which aspect of the restaurant holds more importance in the recommendation system. The authors have proposed a refined way of selecting the useful aspects of comments to build a recommendation engine. Unlike the previous papers, this study emphasizes selecting the important aspects from text feedback. The literature also gives ideas on how to go about answering the proposed research question and figuring out the methodology that needs to be followed.

2.2 Rating and Recommendation for Food Dishes

In his study, Kedar Ratnaparkhi (Ratnaparkhi, 2018) has used user comments from the Yelp[2] dataset and extracted meal items for each restaurant in a particular city. Aspect-based sentiment analysis is carried out on the review data after processes like tokenization, stop-

---

[1] https://wordnet.princeton.edu/

[2] https://www.yelp.com/dataset

word removal, N-gram preparation, POS tagging, etc. The food-level sentiment score is generated by analyzing the associated adjectives mentioned in the comments. After doing aspect-based sentiment analysis, the finest meals at a restaurant are categorized using popular machine learning classifiers such as KNN, Random Forest, and SVM. In the food-restaurant recommendation domain, this paper aims to suggest the most popular meals to customers in a restaurant. However, as it suggests top-rated items in an eatery, it will recommend the same items to all the customers at a specific restaurant. So, the chances of recommending dishes to customers despite their disliking them are quite high.

M. Gupta et al. (2021) have worked on a food recommendation system depending on the mood of the user. This mood-based recommender model utilizes Zomato's dataset to locate restaurants based on user location, and attributes related to food items like cuisine, location, etc. are obtained from the dataset and used to construct modes like rating and mood. KNN clustering is used for the recommendation system's development. The system provides suggestions through a website where the user has to input their current mood. Although the system does not involve past data from users and only focuses on meal items and their characteristics, the idea of associating characteristics with food can be utilized to answer our research question. In the study conducted by Mathayomchan and Taecharungroj (2020), authors have analyzed the impact of different aspects of restaurants, i.e., attributes like service, value, atmosphere, etc., on customer meal experiences, which provided the inspiration for analyzing aspects in this project. Ha C. H. (2021) builds a collaborative filtering-based recommendation system for Yelp data focusing on cuisine types like Italian, American, and so on. This study aims to predict latent ratings for personalized restaurant ranking while considering the user's culinary preferences. However, our study focuses on specific dishes rather than cuisine.

Trending research in the food recommendation domain focuses on predicting healthy food based on diet and health issues. Authors like Rostami et al. (2022), Gao et al. (2022), and Tran et al. (2021) have conducted research on recommending nutritional foods based on their nutrients, ingredients, and recipes to encourage a healthier diet for individuals. Our research differs in its fundamental idea in that it aims to recommend dishes available at restaurants based on customer preferences. It does not take into account attributes like nutritional value or the ingredients present. However, all these studies provide elaborated techniques used in food recommendation, where aspect-based sentiment analysis is used while analyzing the opinions related to food. Ideas related to the granular-level study, i.e., the food-level study, are considered while designing the methodology of the proposed research. This in turn invokes the need for analyzing the studies related to aspect-based sentiments.

2.3 Aspect Based Sentiment Analysis

Opinion mining is critical to learning whether a customer likes or dislikes a product, and aspect-based sentiment analysis (ABSA) helps with that. Khodra et al. (2017) emphasize ABSA's significance in restaurant reviews. The authors have used Indonesian Tripadvisor data, employing preprocessing steps like sentence splitting and normalization. Detailed sentiment analysis steps include aspect extraction, categorization, sentiment classification, opinion structuring, and rating calculation. Various methods, like bags of N-grams, CBOW clusters, and MaxEnt classifiers, were used that yielded good F1 measures of 0.793, 0.823, and 0.642 for aspect extraction, categorization, and sentiment classification, respectively. This study presents a comprehensive approach to extracting aspects and assessing sentiments, enhancing the understanding of consumer viewpoints. Zhang et al. (2021) aimed to enhance aspect-based sentiment analysis (ABSA) accuracy while delving into customer behavior, genuine sentiments, and product inclination. They have utilized a restaurant review dataset and performed pre-processing like segmenting, eliminating redundant tags, punctuation, and garbled text, as well as stop-word removal, normalization, and feature extraction using techniques like TF-IDF and Word2vec. The analysis extracts five primary aspects, namely, location, service, price, environment, and dish, along with 18 detailed aspects such as traffic convenience, wait-time, cost-performance, discounts, portion, and taste. They have also made use of a BERT-based weakly supervised classifier and multi-aspect stacking model using LR, SVM, GBDT, and XgBoost for the purpose of analysis. While both studies present innovative sentiment analysis methodologies, Zhang et al. (2021) focused on extracting finer aspects and detailed sentiments, utilizing deep learning to optimize language models. However, our study concentrates on food dish aspects and sentiment scoring.

In another study, Md Shad Akhtar et al. (2017) explored diverse natural language processing techniques like POS tagging, headword detection, chunk analysis, lemmatization, stop-word removal, and WordNet utilization for feature extraction from text reviews. The researchers have created an ensemble method for sentiment classification, combining Maximum Entropy (ME), Support Vector Machine (SVM), and Conditional Random Field (CRF)-based learners. While effective for manual aspect identification, such comprehensive preprocessing requires substantial processing time and resources. Gomathi et al. (2019), in their research, also highlighted the NLP techniques such as tokenization, suffix-prefix analysis, POS tagging, and so on, used for extracting sentiment and aspect from feedback. After the analysis, user comments are tagged as either positive, negative, or neutral. This data is then utilized to build recommendation engines for restaurants. Also, in other studies (Hossain et al., 2020; Zahoor et al., 2020; R Gupta et al., 2021), authors have emphasized the importance of sentiment analysis of user feedback for recommendation purposes in the hospitality domain. In those studies, the pre-processing and feature engineering included steps like tokenization, stop-word removal, etc., which will be utilized in our study as well.

Sentiment analysis is performed to determine polarity scores for specific comments or different aspects retrieved from the comments. In Chitalia's (2023) research, the author has utilized the Yelp dataset to determine restaurant popularity scores with the help of sentiment analysis and natural language processing techniques. It utilizes the idea that, while user reviews hold substantial influence on business popularity, the star rating serves as a quick decision-making metric. This study recognizes the potential of additional aspects across reviews, users, and businesses in refining recommendations. It combines the sentiment score with the restaurant star rating given by the user to determine the popularity score. A similar approach is followed in our research while estimating the final rating for the food and restaurant pair. Together, these studies discuss the concepts of aspect-based sentiment analysis and pre-processing. Our study adapts these methods for achieving text pre-processing, concentrating on extracting food aspects and assigning sentiment scores for each customer, facilitating collaborative filtering for producing suggestion lists.

2.4 Collaborative Filtering for Recommendation or Rating Prediction

To develop recommendation systems based on similar user choices, collaborative filtering (CF) algorithms are widely used. Numerous studies have been conducted on CF in restaurant and food industry contexts to enhance the user experience through customized recommendations. A novel prediction score formula and an enhanced collaborative filtering algorithm were proposed by the authors (Ling et al., 2017). For the purpose of making suggestions, the researchers have taken into account both user reviews of a restaurant and user traits. They have given the average scoring component a greater weight when grading similarity and included the confidence level correction factor to increase the mean score for the similarity measure. A new CF algorithm is developed using the new formula, and RMSE and MAE are considered as evaluation metrics. The study, however, incorporates user ratings and other features but not user sentiments. The authors, Tripathi and Sharma (2020), took an alternative approach when it came to collaborative-filtering-based recommendations, attempting to forecast restaurant and user ratings solely based on previous customer reviews of a restaurant. A multiclass SVM classifier and KNN algorithms are used to perform collaborative filtering, and the MSE value is utilized to obtain the best model, which in this case is the SVM classifier. While the first study concentrated on raising the mean score for similarity to create a novel algorithm, the second study evaluated machine learning-based techniques to determine the one with better performance. The second study examines and takes into account user sentiments.

Authors Mara Deac-Petrusel and Sergiu Limboi conducted a study (Limboi et al., 2020) to improve the recommendation process by incorporating user sentiment into collaborative filtering. The traditional approach of considering user ratings for recommendations does not reflect accurately, a user's feelings about a particular item, service, or restaurant as a whole. However, the authors of the study developed a sentiment rating approach to obtain sentiment scores for each item. Three factors determined the similarity measure, namely, the attractiveness, relevance, and popularity (ARP) of comments and users. Then, K-nearest neighbour was used for collaborative filtering, and the new ARP measure outperformed certain standard similarity measures in terms of MAE and RMSE, such as the Pearson correlation coefficient, Jaccard index, cosine similarity, Spearman coefficient, etc. Review data from Yelp and Datafiniti Hotel has been used for this purpose. This study demonstrated how sentiment scores may be effectively included in collaborative filtering to enhance the recommendation system. Siddik et al. (2023), in another study, explored matrix factorization algorithms such as Singular Value Decomposition (SVD), SVD with Implicit Ratings (SVD++), and Non-Negative Matrix Factorization (NMF) in the context of collaborative filtering. While evaluating food product ratings in an Amazon Fine Food Reviews dataset, this study compares these algorithms with varying latent factors. NMF and SVD++ emerge as superior, with NMF achieving an average prediction error of 0.7311 (MAE) and SVD++ recording a prediction error of 1.0607 (RMSE). However, SVD++ is used here as the dataset contains implicit information about user feedback.

The use of collaborative filtering is predominant in recommendation systems in other fields as well, such as while recommending a movie, a song, etc. M. Gupta et al. (2020), in their research study, have employed KNN algorithms and collaborative filtering to increase accuracy in comparison to content-based filtering while making a movie recommendation. Utilizing cosine similarity with k-nearest neighbor and collaborative filtering, the approach overcomes the shortcomings of content-based filtering. The study sought to optimize movie recommendations for an enhanced user experience. In this section, papers related to collaborative filtering are analyzed to come up with techniques suitable for our data and research question.

All of the previous literature offers comprehensive insights into the various research methodologies for recommendation systems in the restaurant and food domains. These studies have explored diverse approaches and methods such as customer feedback analysis, natural language processing (NLP), aspect-based sentiment analysis, and collaborative filtering, to name a few. Our research study draws inspiration from such ideas and leverages them to define the purpose of our research and formulate our methodology. Previous research studies mostly focused on recommending highly rated food items and making restaurant suggestions based on ratings or sentiment scores from feedback. In our study, we provide a unique contribution by recommending specific food items at restaurants based on the individual's personal preferences, derived from reviews and user similarities. In our study, we extracted food-related sentiments from reviews and employed collaborative filtering to generate personalized recommendations.

3 Research Methodology

Drawing on existing literature, the methodology of our study is outlined. The objective of this study is to suggest restaurant food pairings to customers based on their food preferences. This research employs the KDD (Knowledge Discovery in Databases) approach, comprising several systematic stages to derive insights from data. These stages encompass defining the problem, selecting relevant raw data, preprocessing and transforming the data, applying collaborative filtering algorithms for data mining to reveal valuable patterns, and ultimately assessing the results to interpret and present knowledge. The diagram below illustrates the proposed KDD-based methodology.

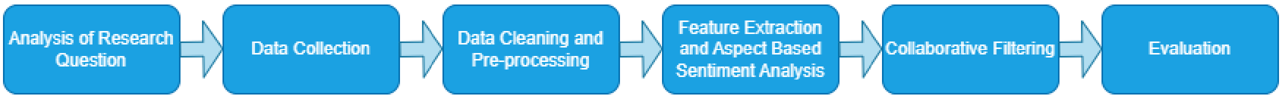


Figure 1: Methodology for the proposed research

3.1 Data Collection

The initial step for the study is to search for an appropriate dataset. Numerous well-known datasets from the restaurant domain have been utilized in prior studies, predominantly emphasizing sentiment analysis. Research conducted by Gomathi et al. (2019), Hossain et al. (2020), and many others has used the Tripadvisor[3] dataset, which contains user reviews on

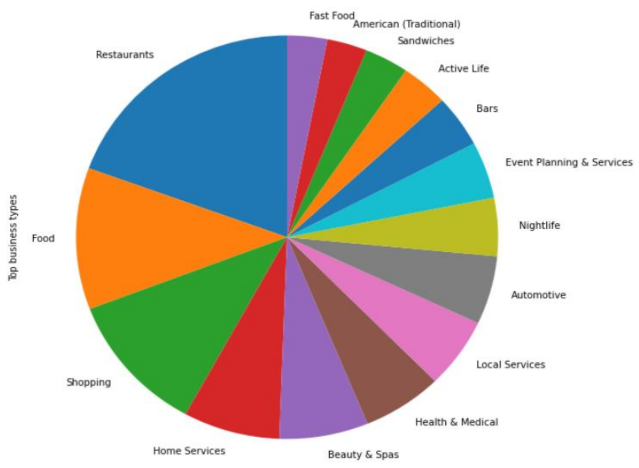
[3] https://tripadvisor.com/

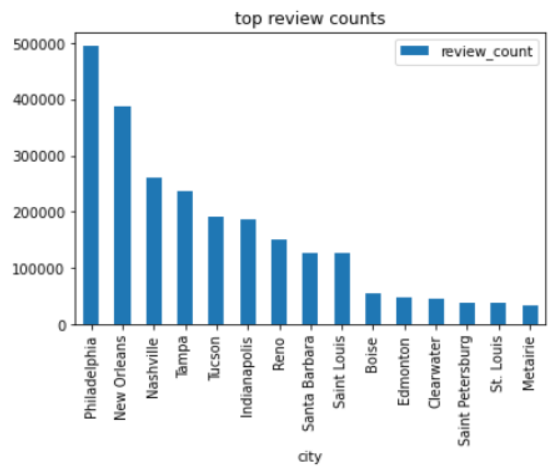
their website about hotels and restaurants. However, the volume of data from Tripadvisor is huge, and gathering data from there requires web scraping; therefore, it is not considered for our study due to the limited time frame. Another popular dataset used in this domain is the Zomato dataset, which is utilized by researchers like R. Gupta et al. (2021), M Gupta et al. (2021), etc. in their studies of sentiment analysis and food recommendation. The dataset provides information about the businesses and reviews but not about the users, so it is discarded in our project as it is mainly user-focused. Ratnaparkhi (2018), Limboi (2020), and many others have used another popular dataset, publicly available at the Yelp website, which comprises of approximately 6,990,280 reviews, related to 150,346 businesses, distributed across 11 metropolitan cities, contained within six JSON files: business, reviews, user, checkin, tip, and photo. The dataset was selected for our project as it is most suitable to work with users and perform sentiment analysis to build a recommendation system.

3.2 Data Cleaning and Pre-processing

For our study, only business, review, and tip files will be used, as others have minimal relevance. The business.json file holds business details like ID, name, location, ratings, and categories. The review.json file contains user and long textual reviews, while the tip.json file contains short user comments about food and services. First, due to their large size, all the JSON files are converted into CSV, and EDA and data cleaning are performed to select only the necessary data.

Data Cleaning: The Yelp dataset contains numerous records beyond restaurants, including home services, grocery shops, pharmacies, etc. For our study, only data related to restaurants was selected. Moreover, as a consumer will only be recommended food and restaurants in his or her city, reviews from only one city are selected. From EDA, it is observed that although Philadelphia has the highest number of reviews, New Orleans has the highest median number of reviews per restaurant, i.e., 88, and also has the second highest total review count, i.e., 386986. Hence, data related to only New Orleans is selected, which includes information about 1391 restaurants. It is observed that review data has long feedback, while tip data contains short feedback. Hence, both are considered as reviews, and business, review, and tip data are merged for the specific city to create one data frame. Both the reviews and tips are concatenated into a single column and finally the data is checked to remove duplicates and null values.





```
Avg reviews per restaurant:
             city  review_count
474   New Orleans          88.0
739        Tucson          53.0
467     Nashville          52.0
702         Tampa          46.0
530  Philadelphia          46.0
```

Figure 2: EDA- Types of businesses (upper left), Top review counts (upper right), Average review count per restaurants (bottom).

Pre-processing: In the case of pre-processing, steps shown in studies like Ratnaparkhi (2018) and Md Shad Akhtar et al. (2017), are followed to prepare data for feature extraction. Once the data is cleaned, text reviews are converted to lower-case. With the help of NLTK library, tokenization is performed, and stop-words that do not have much contribution towards sentiments, such as ‘a’, ‘the’, ‘is’, etc., are removed to enhance processing efficiency by reducing the data. Next, from the tokens, N-grams are formed with value 3, as in natural language sentences, descriptive words generally appear near nouns, and 3-grams have shown better efficiency than others (Don A. et al., 2007).

3.3 Feature Engineering

Feature engineering for this study involves four major steps: food word extraction, food description creation, estimating sentiment score, and computing the overall score.

Food Dish Extraction: To identify food dishes within the text reviews, two approaches exist: gathering individual restaurant-specific items and searching feedback, or utilizing a predefined, efficient dictionary related to food items. The latter approach is chosen due to its time efficiency and genericness, which can be applied to any restaurant. Research that focuses on extracting food-related terms as aspects has commonly utilized food dictionaries or ontologies (E Asani et al., 2020; Md Shad Akhtar et al., 2017). In this project, the WordNet ontology is utilized to create a food dictionary. The food dictionary created in such a way contains few irrelevant words and ingredient names such as sugar, wheat, etc., and are filtered out by understanding the content of the dictionary. Next, all the food names are extracted from the corresponding reviews for all the users using the refined dictionary, and eventually, rows that do not have any food items are removed due to their lack of significance. Furthermore, irrelevant N-grams lacking dish names are filtered out too.

Food-Opinion Extraction: Once the dishes are extracted, the next vital step is to generate opinions regarding the food from user comments. For this purpose, first POS tagging is performed on the N-grams using the python NLTK package to generate more information at the granular level, which is crucial for the analysis. As descriptive words, mostly adjectives are used, along with some verbs and nouns. Although considering verbs and nouns as descriptive words increases the presence of irrelevant words as descriptions, excluding those might result in losing important information. For example, if a sentence contains “loved burrito”, ‘loved’ will be tagged as ‘VBD’, i.e., verb, and it will give a positive sentiment, and excluding that means losing information. Also, as these will be utilized in sentiment analysis and those irrelevant words yield neutral sentiments, all the nouns (‘NN’), verbs (‘VB’, ‘VBD’, 'VBG', 'VBN', 'VBP', 'VBZ'), and adjectives (‘JJ’) present in N-grams are considered to create a food-description pair. In his study, Ratnaparkhi

(2018), follows a similar approach. To generate the food-description pair, a {key: value} pair is generated for each food item attached to each of the corresponding descriptive words. Now for every user and review, there are multiple such pairs, which will be further utilized to analyze sentiment.

Sentiment Analysis: After {food: description} creation, sentiment analysis is done on the words used as descriptions to compute positive and negative sentiment scores for each word. In previous literature (Ratnaparkhi, 2018; K. Zahoor et al., 2020), authors have made use of the open-source lexicon resource named SentiWordNet (Esuli and Sebastiani, 2006) to estimate sentiment scores for opinionated words that are present in the English dictionary. Similar to this, each description's positive and negative scores are determined and assigned to the relevant food dish. Next, the average positive and negative score for each dish is estimated along with the total sentiment score by subtracting the negative from the positive score.

Determining Score for Food-Restaurant Pair: As the objective of the research is to recommend dishes and restaurants, an overall score needs to be estimated from a food and restaurant perspective. The sentiment score is already computed at the food token level. In the case of restaurants, we have a rating of customers present for every business. Hence, that is utilized to determine a total rating score. The sentiment score is in the range of 0-1. However, the restaurant ratings are in the range of 0-5, and therefore a min-max normalization is performed to obtain a normalized value for ratings. The idea is taken from the study conducted by Chitalia (2023), where the author combined the sentiment score and normalized ratings to determine a total score for Yelp restaurant popularity. Next, (restaurant, food) pairs are created, taking the business_id and dishes for each customer. Finally, the average of sentiment value and normalized ratings is calculated to obtain an overall score for every restaurant-food pair.

Only the records containing a total positive score are considered, and hence, the dataset is prepared with all the necessary information to proceed with collaborative filtering.

3.4 Model Implementation: Collaborative Filtering

Once the data is transformed, it contains scores that indicate how much each user has liked each food-restaurant pair. Hence, the next step is to apply collaborative filtering to generate recommendations of food and restaurants depending on the similar scores of users, which signifies their likings. Researchers like Limboi et al. (2020) and M Gupta et al. (2020) have utilized KNN collaborative filtering algorithms to build recommendation engines. On the other hand, while dealing with huge datasets, authors (Siddik et al., 2023) have implemented matrix factorization-based collaborative filtering algorithms such as SVD, SVD++, and NMF to get good results. Since the dataset used is quite large, the experiment is divided into two sections as the implementation of KNN algorithms requires huge memory space and is not possible due to physical resource restrictions.

Experiment 1: In the first case, the original dataset is sampled into a small set and then split into the train and test data in an 80:20 ratio. Then, KNN algorithms such as KNNBasic and KNNWithMeans are applied to develop the recommendation system.

Experiment 2: In the second case, the full dataset is utilized to apply models based on matrix factorization suitable for large datasets. The dataset is also split into 80:20 ratios to create the train and test datasets. Two algorithms, SVD and NMF, are applied and compared to obtain the best model, depending on the performance. As our dataset contains only ratings and no implicit feedback, SVD++ is not applied.

In both of the experimental approaches, cross-validation and hyperparameter tunning are carried out to enhance the efficiency and performance of the models.

3.5 Interpretation and Evaluation

Once the models are trained, the evaluation is done on the basis of metrics like RMSE and MAE, mainly as done in previous literature (Tripathi and Sharma, 2020; Limboi et al., 2020). A lower RMSE and MAE value is indicative of a better model. In both experiments, the respective models are evaluated to determine the best collaborative filtering algorithm suitable for our data.

4 Design Specification

To build a recommendation system, a proper design needs to be followed throughout the project.

4.1 Architecture of the Entire System

The architecture and methods used to create a collaborative filtering recommendation system are described in the design specification. The diagram below depicts the underlying architecture of the methods followed.

Figure 3: Design of the Implementation Flow

4.2 Techniques and Algorithms Used

Mainly two types of algorithms are used in this study: KNN-based and Matrix-factorization-based.

1. **KNN-based Collaborative Filtering:** K-Nearest Neighbors (KNN) based collaborative filtering is a neighbourhood-based approach used to develop recommendation systems. Based on similarities between users or items, recommendations are made. The basic idea is to identify the most similar users or items to a target user or item and use their preferences to generate recommendations.

   KNNBasic: KNNBasic is a collaborative filtering method that predicts a user's preference for an item based on the preferences of its K most similar users or items. Similarity scores are computed using measures such as cosine similarity or Pearson correlation, and preferences are combined to generate recommendations. This method is simple and effective for recording local user-item interactions.

   KNNWithMeans: KNNWithMeans extends KNNBasic by incorporating user or item mean preferences into predictions. In this method, recommendations are modified by correcting the biases in the data. In certain scenarios, this method gives better performance by taking into consideration both local similarities and global user or item preferences while maintaining simplicity.

2. **Matrix-Factorization based Collaborative Filtering:** Matrix factorization is a mathematical approach to decomposing a matrix into the product of two lower-dimensional matrices. It is used in recommendation systems to detect hidden characteristics in user-item interaction data, allowing for customized item recommendations.

   SVD: Singular Value Decomposition (SVD) is a matrix factorization technique used in collaborative filtering. SVD decomposes the user-item interaction matrix into lower-dimensional user and item matrices, capturing latent features that represent user preferences and item characteristics. SVD extracts significant patterns and enables customized recommendations based on latent factors by approximating the original matrix using a selection of singular values and their corresponding vectors.

   NMF: Non-Negative Matrix Factorization (NMF) is a matrix factorization approach employed in collaborative filtering. NMF factorizes the user-item interaction matrix into non-negative user and item matrices, ensuring that latent features remain non-negative and interpretable. This intrinsic limitation yields meaningful patterns that can more intuitively reflect user preferences and item attributes.

4.3 Requirements

The architecture of the system provides insights into various requirements. First of all, in terms of software resources, Python and its relevant packages and or libraries such as Pandas, NumPy, Scikit, Surprise, etc. need to be installed in the system to enable smooth processing. Hardware should also be supportive of the project’s implementation, considering the large dataset. In terms of the data requirement, it should be fetched from the identified source, which is Yelp.com, by following the proper guidelines to maintain data ethics throughout the project.

5 Implementation

In this section, implementation steps and corresponding outputs are described. The research methodology is followed rigorously during the entire implementation process.

5.1 Data Cleaning

The original dataset downloaded from the Yelp website contains six JSON files, out of which only business, review, and tip are selected. All the data is loaded into a Python data frame in tabular format. Yelp has data related to restaurants along with other businesses such as beauty, home services, medicinal stores, etc. Only restaurant-related data is selected first, and then the city containing the maximum number of median reviews per business is chosen, which also contains the second highest number of total reviews. Therefore, for the project, only the data related to restaurants in New Orleans is chosen. Necessary fields from business, tip, and review are merged, and the resulting data frame consists of 9 attributes (review_id, user_id, business_id, stars, text, city, categories, name, and address) and 387003 records about 180812 users and 1391 restaurants after duplicate and null value removal. The figure below shows a portion of the data frame. Here, the 'text' field contains the concatenated review and tip comment data, which is most useful for our study. This step ensures that the data used for analysis and modelling is accurate, consistent, and free from noise.

	review_id	user_id	business_id	stars	text	city	categories	name	address
0	z0osLHDvXvzfm57D4DmD2Q	xVKE_HJ2pwUtTdLbL3pnCg	S2Ho8yLxhKAa26pBAm6rxA	3.0	Service was crappy, and food was mediocre. I ...	New Orleans	Cajun/Creole, Seafood, Restaurants, Breakfast ...	Creole House Restaurant & Oyster Bar	509 Canal St
1	tXHWJWnTdrraHGUqaPWj3g	zKAHSNzqvwsyoFCw3QpafA	S2Ho8yLxhKAa26pBAm6rxA	4.0	Enjoyed my fish out at a sidewalk table. A bi...	New Orleans	Cajun/Creole, Seafood, Restaurants, Breakfast ...	Creole House Restaurant & Oyster Bar	509 Canal St

Figure 4: Merged and cleaned dataset

5.2 Data Pre-processing

Data preprocessing involves steps like lower-case conversion, tokenization, stop-word removal, and preparing N-grams, which in this case are 3-grams. All these steps are executed with the help of the Python NLTK package, and results are stored in different columns, resulting in a total of 11 attributes after pre-processing, which is shown in the diagram below.

business_id	user_id	review_id	stars	text	city	categories	name	address	text_tokens	ngrams
0__F9fnKt8uioCKztF5Ww	002sVJCpSdFDqb6mCx9okg	i51UYC-axeOZAp8eyR3O-Q	1.0	located in the back of the catahoula hotel i t...	New Orleans	Cafes, Nightlife, Cocktail Bars, Peruvian, Res...	Piscobar	914 Union St	[located, back, catahoula, hotel, thought, fou...	[(located, back, catahoula), (back, catahoula,...
0__F9fnKt8uioCKztF5Ww	0G-QF457q_0Z_jKqh6xWiA	pF1BBNKDrQgfxLEOiZsyCg	5.0	i absolutely love this bar! even though i live...	New Orleans	Cafes, Nightlife, Cocktail Bars, Peruvian, Res...	Piscobar	914 Union St	[absolutely, love, bar, !, even, though, live,...	[(absolutely, love, bar), (love, bar, !), (bar...

Figure 5: Data-frame after pre-processing

5.3 Feature Engineering

This is the most crucial step of the research, where features like dish names and corresponding sentiments are extracted and further used in the analysis to build a recommendation engine.

Food Dish Extraction: The first step in feature engineering is to extract the most important feature of our research, i.e., the dish names. As discussed in the methodology, the WordNet dictionary is used to generate a food dictionary, which is further refined by filtering out irrelevant words such as sugar, wheat, etc., which are not dish names. It is done by understanding the elements present in the dictionary through manual intervention, thereby reducing the size from 3583 words to 2907 words. Now the refined dictionary is used to extract the food items from the comments by comparing them with text tokens produced in the pre-processing step. Finally, records from which no food names are extracted are removed, and the index is reset. Moreover, N-grams that do not contain relevant dish names are also removed due to their lack of significance for analysis. The output is shown below.

user_id	review_id	stars	text	city	categories	name	address	text_tokens	ngrams	food_names	filtered_ngrams
QF457q_0Z_jKqh6xWiA	pF1BBNKDrQgfxLEOiZsyCg	5.0	i absolutely love this bar! even though i live...	New Orleans	Cafes, Nightlife, Cocktail Bars, Peruvian, Res...	Piscobar	914 Union St	[absolutely, love, bar, !, even, though, live,...	[(absolutely, love, bar), (love, bar, !), (bar...	[coffee, espresso]	[(., delicious, coffee), (delicious, coffee, d...
PbctTZ4SV_MS3TaeTQ	MKLDHCphgJ2SCTsLwF7IWg	5.0	what a beautiful way to make use of gorgeous o...	New Orleans	Cafes, Nightlife, Cocktail Bars, Peruvian, Res...	Piscobar	914 Union St	[beautiful, way, make, use, gorgeous, old, str...	[(beautiful, way, make), (way, make, use), (ma...	[popcorn, jambalaya]	[(rooftop, (, popcorn), ((, popcorn, jambalaya...

Figure 6: Data-frame after food extraction

Food-Opinion Extraction: Next, every N-gram undergoes processing through the Part-Of-Speech tagger, which attaches a corresponding part of speech to each word based on the rules of the English language. This process is executed utilizing the Stanford POS tagger, available in the Python NLTK package. Subsequently, a {food: description} pair is created for every food item in each record by attaching the food names with the words tagged as nouns (‘NN’), verbs (‘VB’, ‘VBD’, 'VBG', 'VBN', 'VBP', 'VBZ'), or adjectives (‘JJ’) in their corresponding N-grams. If in any N-gram there are multiple descriptive words present, then those many pairs are generated.

```
filtered_df['filtered_pos_tags'][0]

[[('.', '.'), ('delicious', 'JJ'), ('coffee', 'NN')],
 [('delicious', 'JJ'), ('coffee', 'NN'), ('drinks', 'NNS')],
 [('coffee', 'NN'), ('drinks', 'NNS'), ('morning', 'NN')],
 [('meant', 'NN'), ('sipping', 'VBG'), ('espresso', 'NN')],
 [('sipping', 'VBG'), ('espresso', 'JJ'), ('drinks', 'NNS')],
 [('espresso', 'JJ'), ('drinks', 'NNS'), ('surrounded', 'VBD')]]
```

```
filtered_df['food_descriptions'][0]

[{'coffee': 'delicious'},
 {'coffee': 'delicious'},
 {'coffee': 'morning'},
 {'espresso': 'meant'},
 {'espresso': 'sipping'},
 {'espresso': 'surrounded'}]
```

Figure 7: Example of POS tagging (left) and {Food: description} pair (right)

Sentiment Analysis: After extracting the opinions for food items, sentiment analysis is carried out with the help of SentiWordNet to obtain a positive and negative score for each descriptive word and store it in a separate column. If one food item has multiple descriptions, a score is calculated for each of them, and a final positive and negative sentiment score is estimated by taking the average for each item. Only the attributes like user_id, business_id, stars, and average_scores that contain the food and sentiment value are kept, as other fields do not hold importance anymore. Finally, a total sentiment score is calculated from the positive and negative values for each food item.

food_tokens	filtered_ngrams	filtered_pos_tags	food_descriptions	sentiment_scores	average_scores
['coffee', 'espresso']	[('.', 'delicious', 'coffee'), ('delicious', '...	[[('.', '.'), ('delicious', 'JJ'), ('coffee', ...	[{'coffee': 'delicious'}, {'coffee': 'deliciou...	[{'food': 'coffee', 'description': 'delicious'...	[{'food': 'coffee', 'pos': 0.0, 'neg': 0.0}, {...
['popcorn', 'jambalaya']	[('rooftop', '(', 'popcorn'), ('(', 'popcorn',...	[[('rooftop', 'NN'), ('(', '('), ('popcorn', '...	[{'popcorn': 'rooftop'}, {'popcorn': 'jambalay...	[{'food': 'popcorn', 'description': 'rooftop',...	[{'food': 'popcorn', 'pos': 0.0, 'neg': 0.0}, ...
['banana', 'smoothie']	[(':', 'order', 'banana'), ('order', 'banana',...	[[(':', ':'), ('order', 'NN'), ('banana', 'NN'...	[{'banana': 'order'}, {'banana': 'order'}, {'b...	[{'food': 'banana', 'description': 'order', 'p...	[{'food': 'banana', 'pos': 0.20833333333333334...

	user_id	business_id	stars	average_scores
0	0G-QF457q_0Z_jKqh6xWiA	-0__F9fnKt8uioCKztF5Ww	5.0	[{'food': 'coffee', 'pos': 0.0, 'neg': 0.0}, {...
1	1DjkPbctTZ4SV_MS3TaeTQ	-0__F9fnKt8uioCKztF5Ww	5.0	[{'food': 'popcorn', 'pos': 0.0, 'neg': 0.0}, ...
2	556ZKNJZRscqbPszuDnO-w	-0__F9fnKt8uioCKztF5Ww	5.0	[{'food': 'banana', 'pos': 0.20833333333333334...

	user_id	business_id	stars	food	pos	neg	total_score
0	0G-QF457q_0Z_jKqh6xWiA	-0__F9fnKt8uioCKztF5Ww	5.0	coffee	0.000000	0.0	0.000000
1	0G-QF457q_0Z_jKqh6xWiA	-0__F9fnKt8uioCKztF5Ww	5.0	espresso	0.000000	0.0	0.000000
2	1DjkPbctTZ4SV_MS3TaeTQ	-0__F9fnKt8uioCKztF5Ww	5.0	popcorn	0.000000	0.0	0.000000
3	1DjkPbctTZ4SV_MS3TaeTQ	-0__F9fnKt8uioCKztF5Ww	5.0	jambalaya	0.000000	0.0	0.000000
4	556ZKNJZRscqbPszuDnO-w	-0__F9fnKt8uioCKztF5Ww	5.0	banana	0.208333	0.0	0.208333

	user_id	business_id	stars	food	total_score
0	0G-QF457q_0Z_jKqh6xWiA	-0__F9fnKt8uioCKztF5Ww	5.0	coffee	0.000000
1	0G-QF457q_0Z_jKqh6xWiA	-0__F9fnKt8uioCKztF5Ww	5.0	espresso	0.000000
2	1DjkPbctTZ4SV_MS3TaeTQ	-0__F9fnKt8uioCKztF5Ww	5.0	popcorn	0.000000
3	1DjkPbctTZ4SV_MS3TaeTQ	-0__F9fnKt8uioCKztF5Ww	5.0	jambalaya	0.000000
4	556ZKNJZRscqbPszuDnO-w	-0__F9fnKt8uioCKztF5Ww	5.0	banana	0.208333

Figure 8: Flow of Sentiment-analysis

Overall-score Calculation for Restaurant-food pair: In the last phase of feature generation, business_id and food attribute (restaurant, food) pairs are created for every user. Min-max normalization is carried out for the star ratings of restaurants given by the user, and then a final rating is estimated for the pairs from the star rating and sentiment score. In the final dataset, only the records with a positive rating are retained by removing around 10755 rows. The final dataset contains 134150 unique user records, for a total of 819541 records.

	user_id	business_id	stars	food	final_rating	restaurant_food_pair
0	0G-QF457q_0Z_jKqh6xWiA	-0__F9fnKt8uioCKztF5Ww	5	coffee	0.500000	(-0__F9fnKt8uioCKztF5Ww, coffee)
1	0G-QF457q_0Z_jKqh6xWiA	-0__F9fnKt8uioCKztF5Ww	5	espresso	0.500000	(-0__F9fnKt8uioCKztF5Ww, espresso)
2	1DjkPbctTZ4SV_MS3TaeTQ	-0__F9fnKt8uioCKztF5Ww	5	popcorn	0.500000	(-0__F9fnKt8uioCKztF5Ww, popcorn)
3	1DjkPbctTZ4SV_MS3TaeTQ	-0__F9fnKt8uioCKztF5Ww	5	jambalaya	0.500000	(-0__F9fnKt8uioCKztF5Ww, jambalaya)
4	556ZKNJZRscqbPszuDnO-w	-0__F9fnKt8uioCKztF5Ww	5	banana	0.604167	(-0__F9fnKt8uioCKztF5Ww, banana)

Figure 9: Final Dataset after Feature Generation

5.4 Model Implementation

As mentioned in the methodology, two different experiments are carried out in this study. For this purpose, the Python Surprise package is used to apply collaborative filtering algorithms.

Experiment 1: In this case, only 10,000 rows are sampled from the original data due to memory restrictions, and data is loaded into the Surprise dataset using only the user_id, restaurant_food_pair, and final_rating. The dataset is then divided into train-test sets in an 80:20 ratio by keeping random_state at 42.

Case 1: The KNNBasic algorithm is applied first to the training dataset. Then K-fold cross-validation is applied with 5 folds, and finally GridSearchCV is applied to obtain the best combination of hyperparameters for similarity metric, number of neighbors, minimum number of neighbors to consider, and user or item-based.

Case 2: The KNNWithMeans algorithm is applied in the second case. 5-fold cross-validation and hyperparameter tuning using GridSearchCV are applied here too.

Experiment 2: In the second approach, the entire dataset is considered for model building. Data loading and train-test splits are followed in a similar manner as in the first experiment. Here, two matrix-factorization-based algorithms are applied.

Case 1: The SVD model is applied first. Next, 5-fold cross-validation and GridSearchCV hyperparameter tuning are applied in a similar manner to get the best combination of epochs, learning rate, and L2 regularization term.

Case 2: Here, NMF is implemented, and 5-fold cross-validation and hyperparameter optimization are performed to obtain tuned results for latent factors, user factors, and item factors.

In all cases, RMSE and MAE values are obtained for evaluation purposes.

5.5 Tools and Languages Used

The entire implementation is carried out using the Python programming language. Python's data manipulation libraries, such as Pandas, Numpy, Seaborn, Matplotlib, Plotly, etc., are employed for data preprocessing and transformation. The Surprise library is utilized for collaborative filtering algorithm implementation, cross-validation, and model evaluation. Moreover, the implementation uses built-in functions from the Surprise library to calculate evaluation metrics like RMSE and MAE.

6 Evaluation

The final part of research is the evaluation of the process or models built. In our study, performance evaluation is also carried out in both of the experiments and is discussed below. Two main metrics, RMSE and MAE, are chosen for assessment purposes. RMSE (or Root Mean Squared Error) is a measure of the average magnitude of prediction errors, computed as the square root of the average of squared differences between the predicted values and the actual values. Alternatively, MAE (or Mean Absolute Error) is a metric that calculates the average of absolute differences between predicted and actual values, providing a measure of the average magnitude of errors without the direction.

6.1 Experiment 1: Sampled Data with KNN Algorithms

Case 1: KNNBasic is trained on the train dataset with MSD similarity, and predictions are generated on the test dataset. RMSE values of 0.1568 and MAE values of 0.1200 were obtained. Furthermore, the KNNBasic algorithm yielded an average RMSE of 0.1518 and an average MAE of 0.1165 across the 5-fold cross-validation. GridSearchCV further optimized the hyperparameters, and the best combination was found to be {'k': 10, 'min_k': 1, 'sim_options': {'name': 'cosine', 'user_based': True}}. After optimization, the RMSE and MAE values obtained are 0.1517 and 0.1164, respectively.

Case 2: Employing the KNNWithMeans algorithm resulted in an RMSE of 0.1581 and an MAE of 0.1210 while using MSD similarity. In the case of 5-fold cross validation, 0.1534 RMSE and 0.1174 MAE are achieved. Hyperparameter tuning provides {'k': 10, 'min_k': 1, 'sim_options': {'name': 'cosine', 'user_based': True}} as the optimal combination, which yields a 0.1532 RMSE and 0.1173 MAE value.

The results of both algorithms are compared below. From the table below, it can be said that both the KNNBasic and KNNWithMeans algorithms produced similar results with some minor differences. However, comparing all the values, KNNBasic algorithms, when hyperparameter tuned with 5-fold cross validation, perform better than others as they produce the lowest scores for RMSE and MAE.

Table 1: Comparison of KNN-based algorithms

Models->	KNNBasic			KNNWithMeans		
Evaluation Metrics	MSD similarity	5-fold cross validation	Hyperparameter tunning with 5-fold cross validation	MSD similarity	5-fold cross validation	Hyperparameter tunning with 5-fold cross validation
RMSE	0.1568	0.1518	0.1517	0.1581	0.1534	0.1532
MAE	0.12	0.1165	0.1164	0.121	0.1174	0.1173

6.2 Experiment 2: Full Dataset with Matrix-Factorization Algorithms

Case 1: The first case featured the application of the SVD algorithm, which initially attains an RMSE of 0.1462 and an MAE of 0.1125 with a basic implementation. With 5-fold cross validation, the RMSE and MAE slightly reduce to 0.1462 and 0.1123, respectively. Hyperparameter tuning via GridSearchCV fine-tuned the model, optimizing epochs, learning rate, and the L2 regularization term. The optimized hyperparameters are {'n_epochs': 15, 'lr_all': 0.01, 'reg_all': 0.6}. It produces the best result for the SVD model, with RMSE 0.1272 and MAE 0.0957.

Case 2: In Case 2, utilizing the NMF algorithm, we initially achieved an average RMSE of 0.1377 and an average MAE of 0.1058. After cross-validation with 5 folds, the RMSE dropped to 0.1376, but the RMSE remained the same. The optimized model after hyperparameter tuning achieves RMSE of 0.1304 and MAE of 0.0967, with the best combinations found as {'n_factors': 30, 'reg_pu': 0.02, 'reg_qi': 0.02}.

The table below compares the output of the two models for different scenarios. It is observed that the optimized SVD model performs better than the NMF, considering the RMSE and MAE values. While selecting the best model, papers like Tripathi and Sharma

(2020) and Limboi et al. (2020) have given more weightage to RMSE and MAE as evaluation metrics, and therefore, hyperparameter-tuned SVD is chosen as the best model in the case of the second experiment.

Table 2: Comparison of KNN-based algorithms

Models->	SVD			NMF		
Evaluation Metrics	SVD()	5-fold cross validation	Hyperparameter tunning with 5-fold cross validation	NMF()	5-fold cross validation	Hyperparameter tunning with 5-fold cross validation
RMSE	0.1464	0.1462	0.1272	0.1377	0.1376	0.1304
MAE	0.1125	0.1123	0.0957	0.1058	0.1057	0.0967

6.3 Discussion

The study showcases the application of both KNN-based algorithms and matrix-factorization techniques in the domain of recommendation systems for food and restaurants. The experiments provide key insights into the performance of different algorithms under varying conditions. With limited data, optimized KNNBasic algorithms perform better than KNNWithMeans, and with large amounts of data, optimized SVD has proven to give better RMSE and MAE values than NMF.

The research shows the feasibility of recommending food and restaurants simultaneously. Although a total of four models are compared here and the best in both cases is chosen, all the models are able to predict foods in a specific restaurant for the user. The figure below shows the prediction generated for a certain user, where ‘Restaurant_ID’ is the ID of the business and ‘Item’ is the food suggested.

```
 Prediction for User_ID :f6Ut7EVm0uW_UI2H2tw2Eg
1. Restaurant_ID: _DvAGDt7KnXPvlY9kxWbiQ, Item: brownie
2. Restaurant_ID: iSRTaT9WngzB8JJ2YKJUig, Item: mimosa
```

Figure 10: Example of a prediction generated for a certain user

Due to physical resource restrictions, the experiment is conducted in two ways. However, with proper setup and resource availability, all the algorithms could be tested on the entire dataset. In real-world scenarios, it may happen that one has to deal with an extremely huge dataset, and the availability of memory and computational resources is limited. In that case, sampling has proven to be effective. Moreover, KNN methods demonstrate reliability in scenarios with limited data, while matrix-factorization approaches excel in handling comprehensive datasets. These insights empower practitioners to choose the most suitable algorithm based on their specific requirements.

Although optimized SVD and KNNBasic models are proven to perform better than others, all the algorithms have showcased low RMSE and MAE scores, meaning good performance. Other collaborative filtering algorithms could also be tested and compared to figure out the best CF model.

7 Conclusion and Future Work

The research study proposes an innovative idea for extracting culinary dishes from user comments and recommending the dishes at specific restaurants using sentiment analysis and collaborative filtering algorithms. The project aims to use collaborative filtering, sentiment analysis, and NLP to generate and offer food and restaurants to customers based on similar tastes. Previous and current research frequently examines various restaurant and user-related variables, such as consumer feelings, ratings, etc., for suggesting restaurants or examining the restaurant dish with the highest ratings, but not many studies are done to recommend restaurant-based food. The unavailability of food item data is one factor in such limited studies. For the study, the Wordnet ontology is selected and used to extract food dishes; sentiment analysis is performed to determine the score for restaurants and food items; and finally, collaborative filtering algorithms are applied to generate recommendations.

Due to the huge volume of data, two different approaches are taken where in one approach, a small volume of sampled data is taken, and in the other, the whole data is taken, and collaborative filtering algorithms are applied and compared. KNNBasic performs better with limited data, whereas SVD has shown better results on the entire dataset.

Limitations and Future Work: Though the study presents good results, there are some limitations. The food dishes extracted here are single words, so dish names containing more than one word, such as “Pepperoni Pizza”, are left out. Further work can be done to incorporate such dishes. Moreover, based on the previous literature studied, to find the most relevant descriptive words, 3-grams are chosen. It may leave important information about the sentiments regarding food. Hence, other values of N-grams can be tried and tested. Also, the food dictionary is created only with WordNet. Therefore, other ontologies and resources can be explored to expand the food dictionary and search for more items. In the case of generating the final score, food sentiments and user ratings of restaurants are considered. Other factors, like customers’ sentiments regarding service, ambience, location, etc., can be considered in future work to further enhance the efficiency of the system. Also, in our research, manual intervention is carried out to refine the food dictionary created from WordNet. In future work, it can be automated in two possible ways. One approach is to use GPT model to get only food dishes out of the dictionary and the other way is to obtain ingredient list from WordNet itself and then subtract it from food dictionary to get refined one. Limited computational resources are a prevalent issue in this project. In the future, scalable systems with high configurations can be used to work with large volumes of data. Lastly, as this research has only utilized data from Yelp, further work can be carried out on other data from the restaurant-food industry, like Tripadvisor and so on.

The system can be commercialized by integrating a real-time feedback system and modifying the algorithms to cater to diverse user preferences to improve customer engagement and satisfaction.

The study mainly focuses on the analysis of feedback related to restaurants. However, the idea of the recommendation system built in this research can be utilized in other domains as well. It provides valuable insights into personalized recommendations for food and eateries and future research endeavours.

References

Akhtar, M.S., Gupta, D., Ekbal, A. and Bhattacharyya, P., 2017. Feature selection and ensemble construction: A two-step method for aspect based sentiment analysis. Knowledge-Based Systems, 125, pp.116-135.

Asani, E., Vahdat-Nejad, H. and Sadri, J., 2021. Restaurant recommender system based on sentiment analysis. Machine Learning with Applications, 6, p.100114.

Asani, E., Vahdatnejad, H., Hosseinabadi, S. and Sadri, J., 2020, September. Extracting user's food preferences by sentiment analysis. In 2020 8th Iranian joint congress on fuzzy and intelligent systems (CFIS) (pp. 066-069). IEEE.

Chai, L.T. and Yat, D.N.C., 2019. Online food delivery services: Making food delivery the new normal. Journal of Marketing advances and Practices, 1(1), pp.62-77.

Chitalia, S.B., 2023. Yelp Restaurant Popularity Score Calculator.

Deac-Petruşel, M. and Limboi, S., 2020, September. A sentiment-based similarity model for recommendation systems. In 2020 22nd International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC) (pp. 224-230). IEEE.

Don, A., Zheleva, E., Gregory, M., Tarkan, S., Auvil, L., Clement, T., Shneiderman, B. and Plaisant, C., 2007, November. Discovering interesting usage patterns in text collections: integrating text mining with visualization. In Proceedings of the sixteenth ACM conference on Conference on information and knowledge management (pp. 213-222).

Ekawati, D. and Khodra, M.L., 2017, August. Aspect-based sentiment analysis for Indonesian restaurant reviews. In 2017 International Conference on Advanced Informatics, Concepts, Theory, and Applications (ICAICTA) (pp. 1-6). IEEE.

Esuli, A. and Sebastiani, F. (2006). SENTIWORDNET: A publicly available lexical resource for opinion mining, p. 417422.

Gao, X., Feng, F., Huang, H., Mao, X.L., Lan, T. and Chi, Z., 2022. Food recommendation with graph convolutional network. Information Sciences, 584, pp.170-183.

GIACOMONI, E. (2017) ‘Why Should You Try New Food Experiences?’, eatwith, 24 May. Available at: https://www.eatwith.com/blog/2017/05/24/new-food-experiences/ [Accessed 12 August, 2023]

Gomathi, R.M., Ajitha, P., Krishna, G.H.S. and Pranay, I.H., 2019, February. Restaurant recommendation system for user preference and services based on rating and amenities. In 2019 International Conference on Computational Intelligence in Data Science (ICCIDS) (pp. 1-6). IEEE.

Gupta, M., Mourila, S., Kotte, S. and Chandra, K.B., 2021, August. Mood Based Food Recommendation System. In 2021 Asian Conference on Innovation in Technology (ASIANCON) (pp. 1-6). IEEE.

Gupta, M., Thakkar, A., Gupta, V. and Rathore, D.P.S., 2020, July. Movie recommender system using collaborative filtering. In 2020 international conference on electronics and sustainable communication systems (ICESC) (pp. 415-420). IEEE.

Gupta, R., Sameer, S., Muppavarapu, H., Enduri, M.K. and Anamalamudi, S., 2021, September. Sentiment analysis on Zomato reviews. In 2021 13th International Conference on Computational Intelligence and Communication Networks (CICN) (pp. 34-38). IEEE.

Ha, C.H. and Univerisy, S., 2021. Yelp Recommendation System Using Advanced Collaborative Filtering.

Hossain, N., Bhuiyan, M.R., Tumpa, Z.N. and Hossain, S.A., 2020, July. Sentiment analysis of restaurant reviews using combined CNN-LSTM. In 2020 11th International conference on computing, communication and networking technologies (ICCCNT) (pp. 1-5). IEEE.

Li, L., Zhou, Y., Xiong, H., Hu, C. and Wei, X., 2017, March. Collaborative filtering based on user attributes and user ratings for restaurant recommendation. In 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC) (pp. 2592-2597). IEEE.

Luo, Y., Tang, L.R., Kim, E. and Wang, X., 2020. Finding the reviews on yelp that actually matter to me: Innovative approach of improving recommender systems. International Journal of Hospitality Management, 91, p.102697.

Mathayomchan, B. and Taecharungroj, V., 2020. "How was your meal?" Examining customer experience using Google maps reviews. International Journal of Hospitality Management, 90, p.102641.

Mohana, R.S., Kousalya, K., Krishnakumar, B. and Lohappriya, D., 2022, January. Restaurant based recommender system based on sentimental analysis. In 2022 International Conference on Computer Communication and Informatics (ICCCI) (pp. 1-6). IEEE.

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

Ratnaparkhi, K., 2018. Recommender system for food in a restaurant based on Natural Language Processing and Machine Learning (Doctoral dissertation, Dublin, National College of Ireland).

Rostami, M., Oussalah, M. and Farrahi, V., 2022. A novel time-aware food recommender-system based on deep learning and graph clustering. IEEE Access, 10, pp.52508-52524.

Saha, S. and Santra, A.K., 2017, August. Restaurant rating based on textual feedback. In 2017 International conference on Microelectronic Devices, Circuits and Systems (ICMDCS) (pp. 1-5). IEEE.

Siddik, M.B.S. and Wibowo, A.T., 2023. Collaborative Filtering Based Food Recommendation System Using Matrix Factorization. JURNAL MEDIA INFORMATIKA BUDIDARMA, 7(3), pp.1041-1049.

Tran, T.N.T., Felfernig, A., Trattner, C. and Holzinger, A., 2021. Recommender systems in the healthcare domain: state-of-the-art and research issues. Journal of Intelligent Information Systems, 57, pp.171-201.

Tripathi, A. and Sharma, A.K., 2020, June. Recommending restaurants: A collaborative filtering approach. In 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO) (pp. 1165-1169). IEEE.

Zahoor, K., Bawany, N.Z. and Hamid, S., 2020, November. Sentiment analysis and classification of restaurant reviews using machine learning. In 2020 21st International Arab Conference on Information Technology (ACIT) (pp. 1-6). IEEE.

Zhang, Y., Du, J., Ma, X., Wen, H. and Fortino, G., 2021. Aspect-based sentiment analysis for user reviews. Cognitive Computation, 13(5), pp.1114-1127.