Foundation Project - 1

AMPBA Batch 15

SUBMISSION BY GROUP 10:
Kshitij Sharma
Saumaya Ranjan
Nidhi Srivastava
Mohua Sinha
Deep Kamal Singh

Table of Contents

1   Background

Food outlets are everywhere, whether it is online or offline, serving mouth-watering food has always been a trending business. If you look at the analytics being used in this food business it's negligible. Menu designs with attractive food images to get the attention of the customers, getting reviews, chatbots, etc are some of the basic technology-based solutions that are available globally today.

2   Objective

Track the customer's tastes and preferences based on seasons and moods, by using historical data along with personalized survey data. People may be allergic to specific foods or smells which may result in a very serious problem if neglected. You have to gather as many details as possible from primary and secondary data sources available and develop a Personalized Recommendation Model for Foodies.

2.1   Data Sources

- Public forums (Menus)
- Survey data (for personalized preferences)

3   Solution Approach

3.1   Business Understanding: Defining the business problem

3.1.1   Background

In the current scenario, consumers have multiple options for food both in terms of Dining out restaurants as well as delivery options. In order to select a restaurant for dining out or delivery , consumers have options to browse through the restaurant's menu on their website or to log on to food aggregator platforms to compare the restaurant in terms of Ratings, safe and hygiene, offers, distance etc. However, very limited information is available in terms of relating consumers preferences and taste of food based on seasons moods

3.1.2   Business problem

- Present food recommendation engines are not effective
- Significant churn on the account of limited/sub-optimal choices to customer, impacting future sales and retention

3.1.3   Business Objective

- Maximize revenue by providing the most preferred and likeable food options

3.1.4   Business Constraints

- Minimize consumer fatigue by providing most optimal choices
- Minimize the searching time for the preferred choices

4   Data requirements

4.1   Attributes/Features required

To gather the historical information and determine people taste and preferences, we would need to collect the data around the following attributes:

Attribute/features	Rationale
Food outlet location	Geographic region along with review date will help to determine the seasonality
Cuisines/Menu information	Type of food provided by the food outlet
Popular dishes	Help to determine the preferences for the dishes for the outlet
People Say This Place Is Known For	Help to determine the reasons for visiting the outlet/possible mood
Events	Help to determine the mood on the day by linking with the review date
Ratings/Reviews	To determine the likeability of a dish and likeability of the outlet
Review date	+2 days can be considered as the possible visit date and can help to determine the seasonality impact as well as mood by relating to events around those dates
Reviewer info	To determine the demographics of the reviewer

4.2   Data sources

In order to capture the above attributes, the following sources can be leveraged:

Data Sources	Possible Challenge
Swiggy	Web Scraping
Zomato	Scraping
Yelp	
Dine out	Scraping
Social Media Platforms-	
Personalized Survey	Number of respondents will be limited
Kaggle	Limited information availability

4.3   Data collection process

The data collection process would be done through two modes:
- Unstructured data through ***Web scraping*** on websites, online aggregator platforms etc. This will be done for a geography (say North or South) and for a one-year data in order to capture the seasonality aspect

■   Structured data through ***Personalized survey***: This will be done through a pan-India survey and would be targeting 200+ respondents to capture the food preferences along with demographic information of the respondents

5   Data understanding with summary

In order to build a Recommendation system, we will follow MCDA – Multi criteria decision analysis matrix match, where in we will figure out top driving factors for a customer's food preference and cross match it with attribute set of food options (dishes), further learning in this direction led us to form below indices, which are set of data points we will feed as input to our model

5.1   Data Matrices

5.1.1   Customer food preference indices
-   Nutritional compatibility matrix
-   **Allergic**/medical compatibility matrix
-   **Flavor** preference matrix

5.1.2   Food Dish composition matrix
-   Ingredient's list – composition
-   Complimentary compatibility
-   **Seasonal**/weather compatibility score – humidity, temperature data
-   Customer ratings and order history – likelihood pairs

5.2   Challenges
•   sourcing the data which is not owned by Business is major challenge – even if the data is arranged from external sources, the reliability and accuracy of data remains a big challenge
•   sourcing data from trusted sources has its own cost and we need to define "discovery" cost and continuous data sourcing cost before we can envisage upon viability of our model

6   Implementation Model - future strategy
We can design the solution in 2 steps…Based on the demographics and season, using Naïve bay's theorem , we can suggest the food preference

In the second step, once we have identified the food preference (type of food) based on Naïve Bay's, using recommendation engine, we can suggest the restaurant (using the IMDB problem logic)