# New York Restaurant Recommender System

**Contributors**

1. **Karen Amanya**
2. **Penina Wanyama**
3. **David Mwiti**
4. **Stephen Thuo**
5. **Nicholus Magak**
6. **Mercy Onduso**

**Overview**

The aim of this project is to develop a recommender system for different restaurants in yelp using data scraped from Yelp. The reason for choosing this project is to help stakeholders improve their business deliverables and know what the clients require from them. In this project, we are targeting the hospitality industry and our main stakeholders are the restaurants registered in yelp.

**Business Problem**

The restaurants registered in Yelp in New York City are struggling to attract and retain customers due to fierce competition and a crowded market. With so many dining options available in the city, customers often find it difficult to choose where to eat, and restaurants struggle to differentiate themselves from their competitors. As a result, restaurants are experiencing lower customer footfall, lower average check sizes, and lower customer loyalty.

To address this business problem, we can build a recommender system that provides personalized restaurant recommendations to its users based on their preferences, search history, and location. By leveraging Yelp's vast user and business data, the recommender system can make personalized recommendations that are highly relevant to the user, helping them discover new restaurants that they are likely to enjoy.

The recommender system can take into account several factors when making recommendations, such as the user's preferred cuisine type, location and rating. By providing personalized recommendations, the recommender system can help restaurants stand out in a crowded market and attract and retain customers who are more likely to enjoy their offerings.

By building a personalized recommender system for restaurants in New York City, Yelp can help its users discover new and exciting dining experiences, while also helping restaurants differentiate themselves in a crowded market and increase their customer base and loyalty.

In this social media age, individuals are keen on restaurant recommendations given by influencers and other users on platforms such as Instagram and TikTok as they come with a personal review of the food, service and aesthetic feel of the place. This is primarily because a 1-to-5-star rating is not as informative as a videos and personal reviews of a restaurant from a person that has been to place. As a result, many restaurants and hotels are also shifting more towards such platforms as a marketing tool and away from sites like yelp and as a result, there is a reduction in website visits which eventually contributes to a drop in revenue for the business. Restaurants in yelp can gain a competitive edge in the market by adding a more personalized feel to their website that will allow users to easily find restaurants.

**General Objective**

The objective of this project is to develop a recommender system that provides personalized recommendation to users based on their preferences and past historical data. The final algorithm will be integrated with the scraped data into a user-friendly interface that allows users to easily find and select the best restaurant for their needs

**Specific Objectives**

1. To build a recommender system that will accurately match a user to a restaurant that is most likely to provide the experience they are looking to have.
2. To identify the factors that greatly affect a person’s rating or review of a restaurant e.g., service, location, ambience or food.
3. To improve on the optimization of the yelp website for both the user and business.
4. Develop a recommendation algorithm that takes into account user preferences, historical data, and restaurant characteristics.
5. Create a web application that can be used to integrate the recommendation algorithm into a user-friendly interface that allows users to easily find and select the best restaurant for their needs.

**Research Questions**

1. What are the key features and attributes of Yelp's user data and business data that can be leveraged to build a robust recommender system?
2. What are the different types of recommendation algorithms that can be used in Yelp's recommender system and how do they compare in terms of accuracy and efficiency?
3. How can Yelp's recommender system take into account user preferences, such as location, cuisine, price range, and rating, to make personalized recommendations?
4. How can Yelp evaluate the performance of its recommender system and continuously improve it over time?

**Success Metrics**

1. Root Mean Squared Error(RMSE). This will help in evaluating the performance of the recommender system in predicting the test data.
2. Precision: This metric measures the proportion of recommended items that are relevant to the user. A high precision means that the recommender system is suggesting items that the user is likely to be interested in.
3. Recall: This metric measures the proportion of relevant items that are recommended to the user. A high recall means that the recommender system is suggesting a large proportion of the items that the user is likely to be interested in.
4. F1 Score: This metric combines precision and recall into a single score, which provides a balance between the two metrics. A high F1 score means that the recommender system is performing well in terms of both precision and recall.

**Data Understanding**

We will be scraping data from Yelp. The link to the site is provided here: <https://www.yelp.com/>

We split the data into 5 parts, where each group member was expected to scrape the raw data from their end, and later we combined the files into one csv file. The features we are targeting when scraping the data is the user’s unique identifier, restaurant’s unique identifier, the reviews, ratings and location. The following are the various meanings of the selected features:

1. User’s unique identifier
2. Restaurant’s url
3. Reviews
4. Date
5. Location

**Data Preparation**