**Peek-Pick: a real-time system for displaying and recommending services.**

In recent years, numerous online food platforms have been developed with a number of features for a customer to rate and review the food and services. The ratings and reviews become critical factors when new consumers select their taste. Though the ratings have been quantified from 1-5, the closeness of ratings (e.g., 4.5 *vs* 4.6) does not well distinguish which restaurant is more popular and preferable. However, reviews are more informative and often evaluated by customers before making a decision, especially for some pricy restaurants or services. Due to the limited space for each restaurant to show reviews on a webpage, only one or two reviews for each restaurant are displayed, which does not represent the overall impression by the reviewers. Users need to click the reviews webpage to view more content. Therefore, it is critical to provide a concise and informative summary of the reviews for customers. Recently, many text summarization techniques in Natural Language Processing have been developed to extract key information across a single document and multiple documents. Since most of the text summarization research focuses on model optimization and application large corpus of data in news, books or articles, the application to real-world recommendation systems in still very rare.

In this project, I am going to implement text summarization models, primarily extractive text summarization, to all reviews for each service in Yelp to extract the key information such as favorite food recommended by customers. Based on the summarization of reviews, I will combine other information such as ratings, prices to develop a recommendation system for restaurants in Yelp. The recommendation system (Peek-Pick) aims to provide a quick yet informative overview of one restaurant for consumers.

The data set in proposed using Yelp API to fetch real-time data on the Yelp website. However, it’s found that only three reviews are allowed to retrieve for pubic users. Therefore, the public review data set published by Yelp and the API retrieved information are evaluated in this project. The data set and API links are <https://www.kaggle.com/yelp-dataset/yelp-dataset> and <https://www.yelp.com/developers/documentation/v3>. The whole recommendation system will be hosted using the Heroku web application. I performed exploratory data analysis is access the feasibility of the project, including access Yelp data by Yelp API, generating insights based on the distribution of restaurant data and baseline model of text summarization of reviews.