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

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

In this project, I am going to implement text summarization models to all reviews for each service in Yelp to extract the key information such as favorite food recommended by customers. With the summarization of reviews, I will combine other information such as ratings, prices to develop a recommendation system for restaurants in Yelp. To build recommendation system, Collaborative Filtering approach will be implemented. The data set in proposed using Yelp API to fetch real-time data in 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 link is <https://www.yelp.com/dataset> and <https://www.yelp.com/developers/documentation/v3>. The whole recommendation system will be hosted using Heroku web application. Initial exploratory data analysis is performed to access the feasibility of the project, including access Yelp data by Yelp API, generating the distribution of restaurant data and baseline model of text summarization of reviews.