

A Location-based Recommender System for Restaurant Business Expansion (Proposal Draft)

IBM Data Science Professional Capstone Project

Junkang Gu – 2020

jgu8@u.rochester.edu

University of Rochester, NY

Problem & Background

With the rise of social media, insta-famous restaurants, such as Shake Shack, are fast growing among younger customers. In contrast to mature franchises like McDonald's and Burger King, which already have established networks of existing stores, emerging franchises are growing and expanding their business from cities where they are originally located, to a wider range of places by opening new stores. However, choosing the location for opening a new store may turn out complicated, as cities and neighborhoods vary from each other, sometimes significantly.

This project will focus on Shake Shack as an example, while the recommender system can be generalized and applied to any similar business scenarios. Shake Shack is an insta-famous burger restaurant originated in Madison Square Park, New York City. Shake Shack now has 33% stores in NYC and 7 in the busiest region, Manhattan. Here we suppose that Shake Shack is looking for business expansion to other cities, Rochester for example. Rochester is a city a lot less populated than the NYC, and differs significantly in almost every aspect. However, it might still be worth consideration, due to multiple universities and colleges which bring a huge number of young college students who are potential customers for Instagrammable restaurants. "Rochesterfoodies", an Instagram account, has 11.6K followers, signaling great potential for these restaurants. Therefore, with data analysis, we would like to figure out the best neighborhood in Rochester, if any, for opening a new Shake Shack store.

Data and application

To solve the problem, we will need the following data:

List of neighborhoods in Rochester.

Geographic data of those neighborhoods for map plotting

Venue data of Rochester for clustering the neighborhoods

Venue data for the NYC Shake Shack store

Firstly, I will acquire a list of neighborhoods in Rochester, as well as corresponding data such as coordinates. After that, I will acquire the venue data of these neighborhoods, run K-means clustering to cluster them into groups. Finally, I will calculate the vector distance of venue data of the Shake Shack store and the centroids of each group, to determine which group is the closest to the Shake Shack neighborhood.