Predicting Restaurant Review Rating
GA Data Science - Final Project Outline
Craig Sakuma
February 13, 2014

1. Problem to be solved

Recommend a restaurant a user would like based on their previous reviews, information about the restaurant and information about other users that have similar preferences.

2. Description of dataset

Yelp public data set from Phoenix, AZ. Data set includes reviews, user profiles, business profiles, check-ins and tips.

3. Hypothesis

A Yelp user's review history can provide insight into how they will review future restaurants. By comparing a user with other members of the Yelp community (e.g., elite members, friends of the user, and average Yelp users), you can select a group of members that are similar to the user. Filtering the popular restaurants from this group of similar members will provide appropriate restaurant recommendations.

4. Statistical methods I plan to use and why

The primary method will be a recommendation engine. There also may be opportunities to use NLP to gain insight from review text and PCA to reduce dimensionality.

5. Applications the finding may have

This model will help Yelp accurately recommend restaurants and optimize ads by selecting ads that match the user's profile.

REFERENCES: TBD