CS 410 Text Information System - Final Project Proposal

Team Name: Pacific Search

Project Name: Restaurant Finder

Team Members:

Daniel Felker (Captain)	dfelker2@illinois.edu
Sanjay Babbar	babbar2@illinois.edu
Meng-Yun Tsay	mtsay2@illinois.edu

There are times when you are browsing on the web for a restaurant to go to but cannot figure out because the list of restaurants returned by the search engine does not consider various aspects of restaurant that you are looking for like it should have good ambiance, tasty food items for kids, should be an Italian restaurant, etc. We are trying to solve this problem for our users by creating a search engine meant for searching restaurants only.

We are going to design and implement a chrome extension to recommend relevant restaurant results near the user's location based on user's query. Users can enter their query, such as "cheap", or "good Italian food" in text box on chrome extension and the relevant restaurants will appear in the extension or the browser. By using this extension, users can describe their needs for restaurants as a query in text box and our system will tailor users' needs accordingly and return a ranked list of relevant restaurants so that users are likely to explore hidden gem restaurants around their corners without extra efforts.

Our planned approach is to use a popular text retrieval model that best suits our needs. We are first going to identify if the user's query mentions restaurant type, such as "Italian", "Mexican", "Indian". This will be done by comparing each word in the query with the list of restaurant types we have created. If there is a match, that suggests the user is looking for a specific type of restaurant. So, we will filter the dataset for that restaurant type. We are then going to use the user's query to search using various aspects of filtered restaurants like reviews, ratings of the reviews etc. We will then rank the search results based on relevance, overall restaurant rating and location and then return relevant restaurants to the user.

We plan to use MeTA toolkit to achieve the above-mentioned objectives. For the dataset, we are going to use Yelp's open dataset (https://www.yelp.com/dataset). Our main programming language is Python.

For evaluation, we are going to create a subset of the actual dataset which will serve the purpose of test collection. Based on that dataset, we will create an ideal ranked list of restaurants for various test queries. We will then measure how well our program performed by matching its results with the ideal ranked list.

The main tasks incurred in this project are main restaurant searcher/ranker system algorithm (35-hour work), chrome extension (8-hour work), progress report due on November 13th (3-hour work), final demonstration and presentation (8-hour work), self-evaluation report (8-hour work), source code documentation and other administrative tasks (6-hour work)