A Recommender System for Chinese Restaurants in Toronto

Paul Yonghao Li

August 9th , 2019

I. Introduction

1.1 Background

Toronto has the best Chinese restaurants in North America, and it is a highly diverse city. Many new immigrants choose Toronto as their first place to live in Canada, and there is also a huge internation student community, which provides lots of business opportunities. Many new immigrants and students like Chinese cuisine, while there is not a good platform that give recommendations of the Chinese restaurants based on data.

1.2 Business Problem

I want to build a recommendation system that suggests your next place to eat based on the places your have been to, the information about the venues including users likes and tips, the locations, etc. so the customers will get a short list of recommendations that fit their taste.

This recommendation system can be used in services and applications like food delivery (UberEats) or online-seat-booking platform. With the understanding of customers' preference, the business or application could also send promotion messages to their users to increase the revenue and create better customers experience.

II. Data Acquisition and Cleaning

2.1 Data Acquisition

I will only use Foursquare APIs for data acquisition.

First, I will use "Search for venues" API to find the restaurants.

GET https://api.foursquare.com/v2/venues/search

I will choose Finch Subway Station as centre point, and find venues under "Chinese Restaurant" category, with distance <8km to the centre point. The API gives me 50 results, in which we only need the venue_id, location (latitude and longitude), and name of the venues.

Secondly, I will use "Get Details of a Venue" to find the detail information of the restaurants.

GET https://api.foursquare.com/v2/venues/VENUE_ID/

The Foursquare free API also provide other information such as how many users gave “like” to a venue and the tips from users. However, it only gave me three users and 2 tips

per venue, even there are hundreds of likes in some of the venues. So, in this project, I will not use the user likes and tips information. In real business environment, we should include this information since it is all very helpful of building our recommendation engine.

2.2 Data cleaning and feature selection

The results from Foursquare API are in json format, from which we extract the data we need.

Although in the venue details, we have all the basic features of the restaurants, we found that the subcategories are the most important feature for a Chinese restaurant customer. Because there are a lot of subcategories under the "Chinese Restaurant". A customer wants to have dim sum will not be satisfied in a Szechuan restaurant.

We found about 20 subcategories and use them as labels. To build a content-based recommendation system, we use the One Hot Encoding technique to convert the list to a vector where each column corresponds to one possible value of the label.

III. Create the user profile

Now assume that we have a user who have already been to 5 of the restaurants and gave his ratings. We can use the dot product to create the user's profile, including the preference of each subcategories (labels).

IV. Give the recommendations

Multiply the labels of restaurants by the weights of subcategories in users' profile, we will get the recommendation table. We will recommend the restaurants with highest scores.

V. Further thoughts

This recommendations system is purely content based. In reality we should combine the content features with collaborative filtering to create better customer experience.