Coursera Applied Data Science Capstone Project

The Battle of Neighborhood

Author: Anoop Kohale, IBM

Introduction/Business Problem

Opening any business in any neighborhood requires carefully analyzing number of aspects of the business for it to become successful. There are number of factors influencing the decision, viz. will there be enough customers to buy my offerings, is there any competition around that can take away by revenue, etc. To success the businessman must carefully study these factors and come up with the strategy or plan of operating the business in the neighborhood.

Fortunately, Advanced data analysis and machine learning will help taking this decision with the information available in abundance around the internet. Foursquare is such an information provider. Foursquare provides data about the interesting venues around any neighborhood. We can utilize the machine learning algorithms and find out the clustering of specific business in the neighborhood. This will empower us with the understanding of demographics and we can then take better decision that will result in making the business a success.

Data to be used

Any location you open a business has number of other similar businesses operating. One needs to analyse the data available at your hands to come up with a decision. During this project we will gather data from various data sources. Below is the list of the data sources used for this project.

1. Foursquare Venues data

   a. Type: API Call to Foursquare. JSON data about the venue.

   b. Description: The data has various venues around a location within specific radius. Venues are categorized and reviewed by users of Foursquare

   c. Source: https://www.foursquare.com

2. Geocoder data

   a. Type: Latitude and Longitude data for given location

   b. Description: The latitude and longitude data of a given location can be extracted using GEOCODER library

   c. Source: Geocoder library

3. Neighbourhood data

   a. Type: Neighbourhoods around New York city

   b. Description: Neighbourhoods of New York city

   c. Format: GeoJSON data

   d. Source: https://geo.nyu.edu/catalog/nyu-2451-34572