Capstone Project - The Battle of Neighborhoods

Coursera

William Ho Swee Chiong

July 30 2021

Introduction:

Los Angeles, one of the greatest cities in the United States of America and even in the world. LA is known for Hollywood and also the sunny weather. LA is also a diverse city where the resident in Los Angeles is made up of every races and ethnics, which also make this city shine. However, LA is a huge cities which has a 500 mi^2 of total land area. As comparison, Manhattan Island only has 33 mi^2 of land area. Therefore, every different part of LA will show different characteristics.

Our client MR. Wang is a new immigrant to the United States. Here's is some background of our client: He's a member of middle class income group in China. He owned a Chinese canteen in China before he and his family migrate to the US. He wish to start a business in Los Angeles city. However, he doesn't know much about this cities. Therefore, he seek for business advice to find the best location for him to start the business (open a restaurant) and also find a neighborhood that has an easier accessibility to Asian community to ease his life as a new immigrant.

Our goal

- Using the dataset and webpage available online to build the model to predict the area
- Using foursquare API to cluster the neighborhood that consist of more Asian restaurant and Asian grocery store.

Dataset Chosen

- https://data.lacounty.gov/GIS-Data/ZIP-Codes-and-Postal-Cities/c3xr-3jw2/data – This website consist of a list of LA neighborhood ,it's postal code and also coordinate
- FourSquare API – Using this API also us to access the venue around the location

Using this LA neighborhood dataset to help us to find the solution for our client MR. Wang to found a neighborhood that is most suitable for him to settle down and start his business to ensure him to have a better life. We will perform use some data science skill to process the data such as Data cleaning, data wrangling and also map visualization by using Folium. Next, we will use K-means clustering to perform the machine learning computation.