Capstone Project (The Battle of Neighborhoods)

IBM Data Science

Rupali Panda

City of Mumbai (Photo credit: https://en.wikipedia.org/wiki/Bangalore)

Introduction

Shopping malls have become a very important part of our daily lives in metro cities. With increase in working hours and busy lifestyles, many households and professionals find it difficult to shop for merchandise and daily needs in places which are geographically apart. In addition to the basic household needs, shopping malls provide other amenities such as entertainment, health-related services and restaurants among many others, under one roof. These centres have always attracted people, especially the younger generation, who can shop for the latest fashion trends, do shopping (including grocery and home essentials), dine in restaurants, go to movies and let kids play in well-maintained facilities meant for them. Shopping malls provide a wider variety of good and services, which was not available in local stores.

Bangalore is one of the biggest cities of India. This city with a population of 10 million, is also known as the silicon valley of India due to a large number of software professionals working

for many national and international companies. Busy schedules and modern life choices leave little time for these professionals, who have set the trend of "quality shopping under one roof" and spend considerable amount of money at these malls. Therefore, opening a shopping mall is not only a lucrative business for the property developers but proves very useful for the consumers. However, opening a shopping mall requires a lot of strategic planning, especially the success of such a venture depends heavily on the choice of a suitable location.

Business Problem:

The property developers and investors need to review and then strategically target a suitable location for the shopping mall to maximize the return on their investments. The objective of this project is to analyze and select the best locations in the neighbourhoods of Bangalore to open a new shopping mall. For this analysis, Data science methodology and machine learning techniques such as clustering were used.

Target Audience:

This project will particularly be very useful for the property developers who are looking for investment in a new shopping mall in the Bangalore metropolitan area. It will also be very helpful for the investors/businesses who can open the stores inside the mall.

Data:

The following data are required for this project:

- List of all the neighbourhoods in Bangalore, which is needed to determine the location of existing shopping malls.
- Latitude and Longitude of the neighbourhoods - required to plot and map venue data.
- Venue data, which will be used to perform clustering of the neighbourhoods based on the existing malls in the city.

Data Sources and Data Science Methodology

- The data were collected by using web scarping techniques from the Wikipedia page: https://en.wikipedia.org/wiki/Category:Neighbourhoods_in_Bangalore.
- Then the Python Geocoder package was used to get the geographical co-ordinates of the neighbourhoods of Bangalore city.
- Foursquare API was used to get venue data for the neighbourhoods.
- Finally to analyze each neighbourhood and clustering the k-means clustering is used.