Capstone Project - The Battle of Neighbourhoods

Introduction / Business Problem

Opening of a new Indian food restaurant in Mumbai city

1. Introduction

“People who love to eat are always the best people”. [1] Restaurants are where we always go to eat. These are the places where we eat peacefully. Try new dishes, cuisines. In India, people love to eat Indian food. Why? Because India is a diversified country. Each state has different dishes, different tastes. Hence you will find Indian restaurant all over the places. As expected, Mumbai city has so many Indian food restaurants. Hence to open a new restaurant, an analysis must be done. The objective of this project is to analyse and get the best suitable location to open an Indian food restaurant in Mumbai city. For this analysis machine learning technique like clustering will be used.

Problem Statement- “Best suitable location to open Indian food restaurant in Mumbai city.”

So, the target audience for this project is a businessman who is willing to invest in any location regardless of property cost. Or for someone who is trying to extend his Indian food restaurant chain in Mumbai city. For small investors this may not be a good idea.

2. Data and Methodology

Data from following sources will be used:

- List of neighbourhoods from Wikipedia [2]
- Coordinates (Geocoder Python) using OpenStreetMaps
- Restaurant venues using Foursquare API [3]

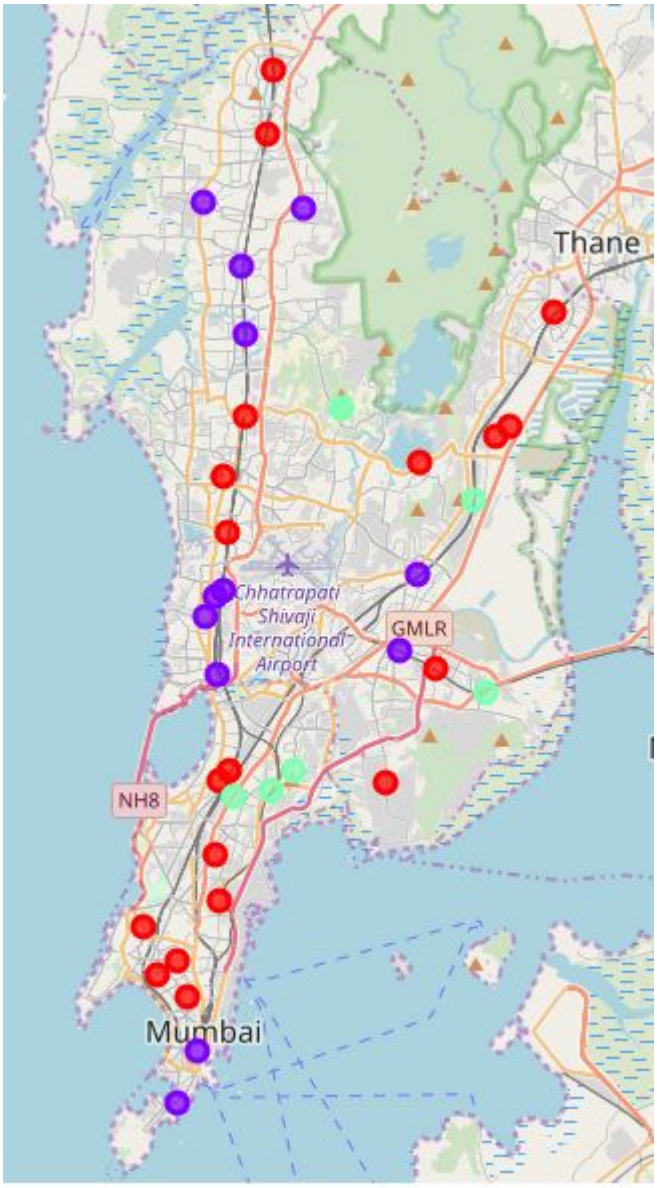
Now, usage of this data will be explained:

- From Wikipedia page, we will get the list of neighbourhoods. For this process, BeautifulSoup Python library will be used. Using these neighbourhoods will be added in the list and later into a data frame.
- Geocoder in Python will be used to get coordinates of the neighbourhoods. Coordinates will be joined with neighbourhoods into the data frame.
- Now using Foursquare API we will get a list of venues and their category.
- Maximum 300 venues in 4000 meters radius will be fetched.
- Data cleaning will be done at all stages whenever required.
- Also, data visualisation will be done using Folium maps, a Python library.
- Machine Learning technique clustering will be used to get results.
- Depending on the frequency occurrence of Indian Restaurants, neighbourhoods will be clustered.
- K-means will help us to understand the concentation of Indian Restaurants in neighbourhoods.

3. Results:

As you can see in the image below, the result is categorised into 3 clusters. These clusters are based on the frequency of occurrence for ‘Indian Restaurants’

- Cluster 0: Neighbourhoods with a moderate number of shopping malls in red colour
- Cluster 1: Neighbourhoods with a low number to no existence of shopping malls in purple colour
- Cluster 2: Neighbourhoods with a high concentration of shopping malls in light green colour



4. Discussion

- As seen in results, Mumbai city has so many Indian food restaurants. Though the prediction pattern is very different, restaurants are located all over the places.
- A high number of restaurants can be found all over the Mumbai city i.e. Cluster 0. A very high number of restaurants are in little outskirts. Probably, there was pretty much free or open space previously i.e. Cluster 2.
- A moderate number of restaurants can be seen in Cluster 1, however most of the locations in cluster 1 are places too far from the centre of the city and only a few are within or around the

centre. So if anyone wants to open a new restaurant Cluster 1 is the best suitable location.
- Also, those who are trying to extend their food chain restaurant, they can also invest in Cluster 0. But investor/businessman is highly suggested to avoid Cluster 2 since they will have intense computation already.

5. Conclusion

- In this project, I have gone through the process of identifying the business problem, specifying the data required, extracting and preparing the data, performing machine learning by clustering the data into 3 clusters based on their similarities.
- Limitation of this project is that I have not constraint the money need to invest. If we consider that then scenario could be different.
- Ultimately suggestion is given to open a new Indian food restaurant in Mumbai city. I conclude that the following areas as shown in the image are the best areas to open a new Indian food restaurant. Hence the neighbourhoods in cluster 1 are the most preferred locations to open a new Indian food restaurant and particularly the 4 regions near the mumbai airport along with Fort,Colaba,Chembur,Ghatkopar.

34	4.5 Fort	0.100000	1	18.932260	72.832880
32	4.3 Colaba	0.080000	1	18.915270	72.826140
24	3.1 Chembur	0.100000	1	19.062200	72.902420
9	1.3 Bandra	0.070000	1	19.054220	72.840190
17	2.2 Ghatkopar	0.100000	1	19.086477	72.908956
5	1.14 Vasai	0.070000	1	19.079340	72.839160
4	1.13 Santacruz	0.070000	1	19.081770	72.842050
2	1.11 Khar	0.040000	1	19.073447	72.835949

6. References:

- The All-Time Greatest Quotes About Food And Eating: https://www.delish.com/food/g25438962/food-quotes/ (https://www.delish.com/food/g25438962/food-quotes/)
- List of neighborhoods in Mumbai: https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Mumbai (https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Mumbai)
- Foursquare Developers: https://developer.foursquare.com/ (https://developer.foursquare.com/)