

COURSERA CAPSTONE

IBM DATA SCIENCE PROFESSIONAL CAPSTONE

Opening a New Shopping Mall in India

By

Praveen kumar

Introduction

Recently, Machine Learning (ML) algorithms are widely used by the people . ML algorithms bring advantages because they offer solutions to problems related to the top questions. The ML algorithms are mainly used for the purpose of predicting the future and recommending the best solution to user. Here I am using the clustering

For many People , visiting shopping malls is a great way to relax and enjoy themselves during holidays. They can do shopping and they can eat in restaurants, shop at the various fashion outlets, watch movies and perform many more activities. Shopping malls are like a best destination for all types of shoppers. For retailers, the main location and the large crowd at the shopping malls provides a great distribution channel to market their products and services. Property developers are also taking advantage of this trend to build more shopping malls to cater to the demand. As a result, there are many shopping malls in india.

Business Problem

The objective of this capstone project is to analyse and select the best locations india to open a new shopping mall. Using data science methodology and machine learning techniques like clustering, if a property developer is looking to open a new shopping mall, where would you recommend that they open it? This is the business problem. This project is particularly useful to property developers and investors looking to open or invest in new shopping malls in india. This project may help the investors to choose the best city to open a new shopping mall.

Data

For this project the Foursquare API will be used. A list of shopping mall in india is downloaded and their respective location in longitude and latitude coordinates is obtained. The Latitude and longitude coordinates of those neighbourhoods. This is required in order to plot the map and also to get the venue data. The sources are the following

Shopping mall list : https://en.wikipedia.org/wiki/List_of_shopping_malls_in_India

The above link the used as a dataset to find which city is best place to open a new shopping mall in india. The data set will be crawled with help of python beautifulsoup and created csv in pandas dataframe

Methodology Section –

In this project india data is taken as a base which contains info of all shopping malls and their corresponding city. Some exploratory data analysis applied –

- 1) From the dataframe we will subset the dataset to get only the records for Shopping mall,
- 2) Data cleaning with removal of all NaN values plus any blank fields.
- 3) Visualize all neighborhood of Shopping mall by using FOLIUM library.
- 4) Fetching Latitude and Longitude of neighborhoods using Geopy Library based on the location
- 5) Normalizing the data fetched from Foursquare API.

Machine learning technique –

We will use Clustering technique of Unsupervised Learning to segment the neighborhoods and cluster them using K-MEANS clustering to get the clusters with similarities in their venues.

This cluster will help the property developer to open a new shopping mall in india

Results Section –

When we run this clustering on india dataset we get the 3 clusters having city which are similar based on venues they have nearby. And cluster 2 is best to open a new shopping mall. Below are top venues in each city.

Cluster 1 :

	City	Shopping Mall	Cluster Labels	Latitude	Longitude
2	Aurangabad	0.083333	0	22.351115	78.667743
25	Lucknow	0.074074	0	22.351115	78.667743
16	Howrah	0.062500	0	22.351115	78.667743
5	Bhubaneswar	0.086957	0	22.351115	78.667743
13	Greater Noida	0.100000	0	22.351115	78.667743
11	Faridabad	0.120000	0	22.351115	78.667743

Cluster 2 :

	City	Shopping Mall	Cluster Labels	Latitude	Longitude
29	Nashik	0.500000	1	22.351115	78.667743
12	Ghaziabad	0.333333	1	22.351115	78.667743

Cluster 3 :

	City	Shopping Mall	Cluster Labels	Latitude	Longitude
0	Ahmedabad	0.044444	2	22.351115	78.667743
24	Kozhikode	0.016667	2	22.351115	78.667743
26	Mangalore	0.020000	2	22.351115	78.667743
27	Mangaluru	0.020000	2	22.351115	78.667743
28	Mumbai	0.000000	2	22.351115	78.667743
30	Navi Mumbai	0.017241	2	22.351115	78.667743
31	New Town, Kolkata	0.000000	2	22.351115	78.667743
32	Noida	0.000000	2	22.351115	78.667743
33	Pune	0.000000	2	22.351115	78.667743
34	Siliguri	0.000000	2	22.351115	78.667743
35	Sonapat	0.000000	2	22.351115	78.667743
23	Kolkata	0.030000	2	22.351115	78.667743
22	Kochi	0.000000	2	22.351115	78.667743
18	Jaipur	0.015873	2	22.351115	78.667743
20	Kalyan	0.000000	2	22.351115	78.667743
19	Jamshedpur	0.000000	2	22.351115	78.667743
36	Thane	0.021505	2	22.351115	78.667743
17	Hyderabad	0.000000	2	22.351115	78.667743
15	Hadapsar, Pune	0.027778	2	22.351115	78.667743
14	Gurgaon	0.000000	2	22.351115	78.667743
10	Dombivali	0.000000	2	22.351115	78.667743
9	Delhi	0.000000	2	22.351115	78.667743
8	Coimbatore	0.033708	2	22.351115	78.667743
7	Chennai	0.000000	2	22.351115	78.667743
6	Chandigarh	0.000000	2	22.351115	78.667743
4	Bhopal	0.000000	2	22.351115	78.667743
3	Bengaluru	0.010000	2	22.351115	78.667743
1	Amritsar	0.027778	2	22.351115	78.667743
21	Kanpur	0.000000	2	22.351115	78.667743
37	Thiruvananthapuram	0.023256	2	22.351115	78.667743

Map Outputs :



Some other outputs :

DataFrame

Out[146]:

	Mall Name	City	Year	size	Longitude	Latitude
0	None	None	None	None	7.540121	44.933143
1	DLF Mall of India	Noida	2016	2,000,000 sq ft (190,000 m2)	77.391073	28.535633
2	Sarath City Capital Mall	Hyderabad	2018	1,931,000 sq ft (179,400 m2)	78.474629	17.361608
3	Ambience Mall Gurgaon	Gurgaon	2007	1,800,000 sq ft (170,000 m2)	77.002700	28.428262
4	DB City Mall, Bhopal	Bhopal	2010	1,300,000 sq ft (120,000 m2)	77.401989	23.258486
5	World Trade Park, Jaipur	Jaipur	2012	1,300,000 sq ft (120,000 m2)	75.820349	26.916194
6	Ambience Mall, Vasant Kunj	Delhi	2008	1,200,000 sq ft (110,000 m2)	77.221939	28.651718
7	R City Mall	Mumbai	2009	1,200,000 sq ft (110,000 m2)	72.835335	18.938771
8	Phoenix Marketcity (Pune)	Pune	2011	1,190,000 sq ft (111,000 m2)	73.854454	18.521428
9	Elante Mall	Chandigarh	2013	1,150,000 sq ft (107,000 m2)	76.779714	30.733442
10	Phoenix Marketcity (Mumbai)	Mumbai	2011	1,140,000 sq ft (106,000 m2)	72.835335	18.938771
11	The Great India Place	Noida	2007	1,000,000 sq ft (93,000 m2)	77.391073	28.535633

Df Head

	Mall Name	City	Year	size
0	None	None	None	None
1	DLF Mall of India	Noida	2016	2,000,000 sq ft (190,000 m2)
2	Sarath City Capital Mall	Hyderabad	2018	1,931,000 sq ft (179,400 m2)
3	Ambience Mall Gurgaon	Gurgaon	2007	1,800,000 sq ft (170,000 m2)
4	DB City Mall, Bhopal	Bhopal	2010	1,300,000 sq ft (120,000 m2)

Df Tail

	Mall Name	City	Year	size
76	Gold Souk Grande, Kochi	Kochi	2011	500,000 sq ft (46,000 m2)
77	MSX Mall	Greater Noida	2013	500,000 sq ft (46,000 m2)
78	Unity One, Janakpuri	Delhi	2015	500,000 sq ft (46,000 m2)
79	Lodha Xperia Mall	Dombivali	2016	500,000 sq ft (46,000 m2)
80	Prozone Mall (Coimbatore)	Coimbatore	2017	500,000 sq ft (46,000 m2)

Conclusion :

In this Project we have used Foursquare API with Machine learning techniques to provide the best results in segmenting the neighborhood according to their venues.