Predicting Location for Opening a New Hotel

Final Report

By-Shyam Bedre August 3, 2020

1.Introduction:

1.1 Description of the problem

The business problem we are currently posing is: We want to find a suitable location for opening a new hotel/restaurant attracted to all local people and foreign tourists in the center of all famous venues.

1.2 Discussion of the background

With nearly 5 million population and a thriving vibrant metropolis, Pune is the center of academia and business that epitomes 'New India' with baffling mixture of capitalism and spiritualism. Geographically Pune might not be the biggest city but it has more to offer than any other town. From being a cultural capital to educational hub, Pune has lot to offer.

Pune is also well known for its hotels and restaurants, with places to choose from traditional street food to 7 courses dine at a five-star restaurant, you will get ever option to try.

I believe it's difficult for to open a new hotel/restaurant, to make a choice from among many options since there are a lot of places available

So, how could we leverage Foursquare location data and machine learning to help us make decision and find appropriate location? This is the problem I would like to address in this capstone project.

2. Data Requirements:

For this project we need following data:

- **Venues in 1km radius neighborhood of Pune:**
- **Data source: Foursquare APIs**
- **Description: By using this API we will get all the venues in neighborhood.**

We will be completely working on Foursquare data to explore and try to locate our new hotel where more venues like church, temples, beach, museums, memorials that are present nearby.

Problem

Considering all the constraints pertaining to the data and weights for each category of the venue, what is the best location in USA to set up a gaming arcade?

Tools we need:

- **Foursquare API**
- **IBM Watson Account**
- **List of United States countries by per capita income**
- **List of United States cities by population**
- **Jupiter Notebooks**

3. Methodology section:

In order to do the analysis and suggest the best location, following are the steps we have to follow:

1. **Removing Outlier - Hotel four seasons restaurant and bar**
2. **How Far are hotels from the core location**
3. **Explore for other venues around Pune**
4. **Extract Venues using Search Queries**
5. **Location of all venues**
6. **How far are venues from the core location?**
7. **Venue Categories**
8. **Rating of all Venues**
9. **Number of Tips for all Venues**
10. **Extracting Rated and Tips Venues**
11. **Final list of Venues**
12. **Clustering based on venues**
13. **Center of all clusters & Midpoint of all venues**

4. Result section:

We got a glimpse of the venues in Pune and were able to find out some interesting insights which might be useful to travelers as well as people with business interests. Let's summarize our findings:

4.a My hotel location

- **Final location is pointed at 18.5194°, 73.8519°.**
- **This location is at Narayan peth near to Bajirao Road and Omkareshwar temple.**
- **Located at exact junction of two cross roads which can give more attention to people who pass by.**

4.b Top Rated Venues

- **Nana Nani Park**
- **Shaniwar Wada**
- **Lal Mahal**
- **Lokmanya Tilak Museum**

All these venues are rated well than other and also, they have more tips and located within 320 meters to core location of Pune. So, tourists may like to visit these places.

4.c Spot my hotel against others

- **Green - My hotel location**
- **Red - Pune's core location.**
- **Black - Venues.**
- **Blue - Other hotels.**

My predicted location and core location are very close to each other which is expected. As this has central attraction, the predicted one almost matched with the core.

Plot to show the final location we got for opening a new hotel/restaurant:

5. Discussion section:

From above reports, we could get an idea why the predicted one is pointed/clustered on the given spot. First most thing could be the center of attraction for the place.

KMeans have figured out the most common place for all the venues. This output was very adjacent to the core location. This proves the accurate spotting of our predicted algorithm.

In the Four-Square API, we have queried the Venues of a locality by specifying the LIMIT and Radius of our choice. We have chosen less LIMIT as the number of API calls that can be done using a free account in Four Square are less. We can increase the limit for more accurate results.

Despite of the findings, there were some lack in data. Tips and ratings were missing for most of the venues. Also, when I compared foursquare data with google map, I could see there were many hotels and venues found missing in foursquare.

6. Conclusion section:

In a fast-moving world, there are many real-life problems or scenarios where data can be used to find solutions to those problems like seen in the example above.

As a business person, one would be able to set up a hotel/restaurant on given spot. This will bring revenue automatically as we have located in very near to core one. We proved this with Kmeans.

Similarly, data can also be used to solve other problems, which most people face in metropolitan cities. Potential for this kind of analysis in a real-life problem is discussed in great detail. Also, some of the drawbacks and chance for improvements to represent even more realistic pictures are mentioned.