**Coursera Capstone Project**

**Utilizing Foursquare API to analyze Chennai’s Neighborhoods**

Introduction:

Chennai is the capital of Tamilnadu and it is one of the metropolitan cities of India. Chennai is the 6th Largest city in India with the population of 9,714,000 as of 2015 with Land area of 971 km2.Diversity of cuisine available is reflective of social and economic diversity of Chennai.

The aim of project is to find neighborhoods with a high frequency of restaurants/food stalls/cafés in Chennai, India. Chennai is one of the Metropolitan City in India. To find popular restaurants or food stalls or cafes, respective coordinates need to be retrieved, so that Foursquare can be used to find nearby venues.

Problem Description:

Chennai has many Ancient temples like Triplicane parthasarathy temple, Mylapore Sivan temple and Hertiage sites like Mahabalipuram, St.Geroge Fort etc. Due to these, many foreign tourists across globe and people from various parts of India are visiting Chennai daily. The main problem they are facing is good lodging and food.

In such situation, they mainly relied on Internet and searching various restaurants nearby their lodging and there are few questions to be addressed:

1. How many types of foods available in the restaurant?
2. Which is the most nearest me having good rating?
3. How many “similar” restaurants are available nearby?
4. Do the “similar” restaurants cost more? If so what is the specialty they will have?

Expectation from proposed recommender system is to address these questions and also to recommend new places based on their rankings.

Target Audience:

Target Audience is not particular to a touristor who are new to Chennai city but also to everyone who is living in Chennai city also since no one aware of popular restaurants in all the places. People who rarely visits restaurants would prefer to have the most rated restaurants nearby them and all this could be handled by our recommender system.

Data

Chennai data is not readily available and i have gathered data from wikipedia page "https://en.wikipedia.org/wiki/Areas\_of\_Chennai" . Data is gathered using web scrapping techniques using bs4, and then grabbed all the hyperlinks. Using urllib, these links are visited individually and the coordinates and pincodes are scraped and put into a pandas dataframe.

The next step is to get all the venues in each neighborhood within a specified radius, in this case, 500 metres. To use Foursquare, we need to create an account as they offer a limited number of API calls per day for a free user (more if you give your credit card details). After signing up, we will get a CLIENT\_ID and a CLIENT\_SECRET which are then appended to a URL along with other parameters, which is then used to send a GET request to that URL. The resulting JSON file is then parsed and stored for further use. After doing this for each neighborhood, the resulting venues are again put into a pandas dataframe.

Methodology

After this, the venue categories are one-hot encoded and then the 10 most frequent venues (only five shown here) in each neighborhood are found out. After this, k-means clustering is used to group these neighborhoods. I forgot to choose the best k for this algorithm. I didn’t notice until after I published this. Oh well. After they are clustered, we can use Folium (which is a map visualization library for Python), to see all the neighborhoods and the cluster they belong to on a nice map:

Success Rate:

With Chennai developing very fast and becomes IT hub, new restaurants are emerging quickly. We need a system that could help us access vast number of food varieties. It’s impossible for a person to ask each and every one about their visit to a particular place and also not every one remembers everything. On the other hand computers remembers everything and machine learning evolves to its peak, its high time technology will be our personal guidance and help us personally based on our likes and dislikes. So people would care about the project as their personal assistance and success rate could increase with time.