

Part 2.txt

Capstone - Battle of Neighborhoods

Part 2 :

Data Description

Since the Contractor wants to build his warehouse in Bangalore, India, I will need geo-locational information of Bangalore and its neighborhood. For this part of the problem, I will use postal code information provided by the India Spatial Data Repository. By using this information, I will analyze the target area and then with help of Foursquare I will find which restaurants are visited most. I will do this by getting data of location (latitude, longitude), distance to center, category and popularity that is provided by a typical request from Foursquare.

Following Libraries are used while executing Capstone project

- Pandas - Library for Data Analysis
- NumPy - Library to handle data in a vectorized manner
- Requests - Library to handle http requests
- Matplotlib - Python Plotting Module
- Sklearn - Python machine learning Library
- JSON - Library to handle JSON files
- Geopy - To retrieve Location Data
- Folium - Map rendering Library

Following data sources are used while executing the Capstone Project

- India Spatial Data Repository - Postal codes of India - File type JSON
- Foursquare data - typical requests data package - File type JSON

Foursquare API :

Foursquare is a social location service that allows users to explore the world around them. The Foursquare API allows application developers to interact with the Foursquare platform. The API itself is a RESTful set of addresses to which you can send requests, so there's really nothing downloaded into your server. You can currently request output in XML or JSON format, making requests to URLs that look like this :
<http://api.foursquare.com/v1/user.json>

HTTP requests parameters :
Number of neighborhoods : 200
Range : 1000

Folium :

Folium builds on the data wrangling strengths of the Python ecosystem and the mapping strengths of the leaflet.js library. Folium makes it easy to visualize data that's been manipulated in Python on an interactive leaflet map. I will use Folium to visualize the best location for our contractor.

K-mean :

Part 2.txt

It is an unsupervised machine learning calculation and i will use it to making clusters and top five restaurants in each neighborhood.