

Data

Based on definition of our problem, factors that will influence our decision are:

- Number of existing restaurants in the neighborhood (any type of restaurant), found using Foursquare API
- Number of and distance to Italian restaurants in the neighborhood, if any
- Distance of neighborhood from city center

Following data sources will be needed to extract/generate the required information:

- Centers of candidate areas will be generated algorithmically and approximate addresses of centers of those areas will be obtained using
- Number of restaurants and their type and location in every neighborhood will be obtained using Foursquare API
- We get the neighbourhoods in Mumbai along with their latitude and longitudes from table present on the Wikipedia page : https://en.wikipedia.org/wiki/List_of_neighbourhoods_in_Mumbai

Based on the data sets obtained above we can extract more features from the data set like:

The distance can be extracted from the latitude and longitudes we have.

The nearby venues in particular neighborhoods can be obtained using Foursquare API and the latitude, longitudes for the neighborhoods.

The density of restaurants in a particular neighbourhood can also be found by getting venues filtered by restaurants in a particular radius and then looking at their counts in a particular area.