The Battle of Neighborhoods

1. Introduction

People search for rental listing of any city would be highly interested in the localities with restaurants, cafes, bars, Grocery stores,etc. People might want to know what kind of venuses are nearby a particular rental listing, so they can make informed decisions in selecting the rental place based on the venuses.The information of nearby venuses of various rental listing in the city and their localities in form of map would help people decide which rental place to choose among the many in the city and also which locality to prefer. Also combining the frequency of the venus nearby a rental listing would help people make easy decisions about selecting the rental place. A map of all the rental listing with specific color attributes will be plotted to highlight the rental listing based on nearby venuses. We will classify the various rental lists into different clusters using a Machine Learning Algorithm, the K-means clustering Algorithm. This enables any user to take a quick glance and decide which place to select for renting out.

2. Data

For this project the Foursquare API will be used. A list of available rental place will be obtained with their respective location in longitude and latitude coordinates.

The source of the rental data is Craigslist for rental ad. We will be importing CraigslistHousing package(package designed to get rental ads directly from the Craigslist site) from craigslist library to obtain the relevant rental data. For now we will only be obtaining first one hundred rental posts in the area of 'San Francisco' using the get_result method of CraigslistHousing object.

A Foursquare API GET request is sent in order to acquire the venues that are within a radius of 500m of each rental post. The data is formatted using one hot encoding with the categories of each venue. Then, the venues are grouped by neighborhoods computing the mean of each feature.

The clustering of rental place will be done based on the frequency of venus near a rental place. This would help use to make a final decision of selecting a particular rental place over the other.