

The Battle of Neighborhood

Finding a Better Place in Scarborough, Toronto

Introduction

- The purpose of this Project is to help people explore better facilities around their neighborhood
- It will help people making smart and efficient decision on selecting great neighborhood out of numbers of other neighborhoods in Scarborough, Toronto
- This project aims to create an analysis of features for people migrating to Scarborough to search for the best neighborhood as a comparative analysis between neighborhoods

Data

- We will need data about different venues in different neighborhoods of that specific borough
- In order to gain that information, we will use "Foursquare" locational information
- Foursquare is a location data provider with information about all manner of venues and events within an area of interest
- Such information includes venue names, locations, menus and even photos
- For each neighborhood, we have chosen the radius to be 100 meters

Methodology

- To compare the similarities of two cities, we decided to explore neighborhoods, segment them, and group them into clusters to find similar neighborhoods in a big city like New York and Toronto
- To be able to do that, we need to cluster data which is a form of unsupervised machine learning: k-means clustering algorithm
- Using credentials of Foursquare API features of near-by places of the neighborhoods would be mined
- Due to http request limitations the number of places per neighborhood parameter would reasonably be set to 100 and the radius parameter would be set to 500

Results

- Scarborough is a popular destination for new immigrants in Canada to reside
- As a result, it is one of the most diverse and multicultural areas in the Greater Toronto Area, being home to various religious groups and places of worship
- This project has used Four-square API as its prime data gathering source as it has a database of millions of places, especially their places API which provides the ability to perform location search, location sharing and details about a business
- The best neighborhood is Adelaide, King, Richmond – based on most common venues

Conclusion

- In this project, using k-means cluster algorithm I separated the neighborhood into 10(Ten) different clusters and for 103 different latitude and longitude from dataset, which have very-similar neighborhoods around them
- Using the charts in the final report, results presented to a neighborhood based on average house prices and school rating have been made
- I feel very learned now on Data Science after successfully completing the specialization on IBM Data Science