Using Machine Learning to decide on suitable locations to start up an Italian Resturant

by
Ezeoke Chidera
(www.github.com/willdera)

1.Introduction

1.1 Background

For this Capstone project, I am creating a scenario under assumption for an idea of an Italian Chef, who has made up his mind on starting up an Italian Restaurant in Toronto but is indecisive on where to pick and wants to know which areas are his best option. The idea behind this project is to find out if there are areas in Toronto with a low number of Italian Restaurants which could present as a great opportunity for this chef, who is based in Canada to make the best impact and not get overshadowed by existing italian restaurants. This chef is thinking of opening this restaurant in locations where Italian food is a fun try and is enthusiastic of the experience. With the purpose in mind, finding the location to open such a restaurant is one of the most important decisions for this chef and I am designing this project to help him find the most suitable location for his mission.

1.2 Business Problem

The objective of this capstone project is to find the most suitable location for the chef to start his new Italian Restaurant in Toronto, Canada. By using data science methods and machine learning methods such as data mining, data wrangling and clustering, this project aims to provide solutions to answer the business question:

In Toronto, If a chef wants to start up an Italian Restaurant to introduce people to italian food, where should he launch it.

1.3 Target Audience

The chef who wishes to find the most suitable location to launch his Italian Restaurant.

2. Data To Be Used

To solve this problem, I will need below data :

- List of neighborhoods in Toronto, Canada.
- Latitude and Longitude coordinates of these neighborhoods.
- Venue data related to Italian restaurants. This will help us find the neighborhoods that are most suitable to open an Italian restaurant.

3. Extracting Data

- Scrapping of Toronto neighborhoods via Postal Codes on Wikipedia.
- Getting Latitude and Longitude coordinates of these neighborhoods via Geocoder package.
- Using Foursquare API to get venue data related to these neighborhoods.