How Machine Learning Can Help a Food Truck Manager to Find the Best Neighbourhood to Sell in Manhattan

Israel Pinheiro de Siqueira*

E-mail: israel.pinheiro@aluno.unb.br

Introduction

Find a profitable neighborhood for a food truck is not an easy task. Today, finding a profitable area is done by experimentation and observation and requires some kind of expertise, and obviously, luck.

With the advance of computational resources, machine learning algorithms became more popular and can be used to solve a vast class of problems that can't be solved a few years ago.

Using machine learning algorithms to find and group neighborhoods based on the residents preferred restaurants and based on the idea that we, as human beings, tend to group with others that are similar in some sense (tastes, lifestyle, thoughts, ideas, etc),[1 2] we will show in this work how machine learning can be used to make good indications of a profitable neighborhood based on the type of food (or foods) he offer.

Data

The principal idea is to cluster and segment neighborhoods based on locals' favorite restaurants.

To do this we need:

- Names and geographic coordinates(latitude, longitude) of all neighborhoods of Manhattan;
- Trending restaurants around each neighborhood;

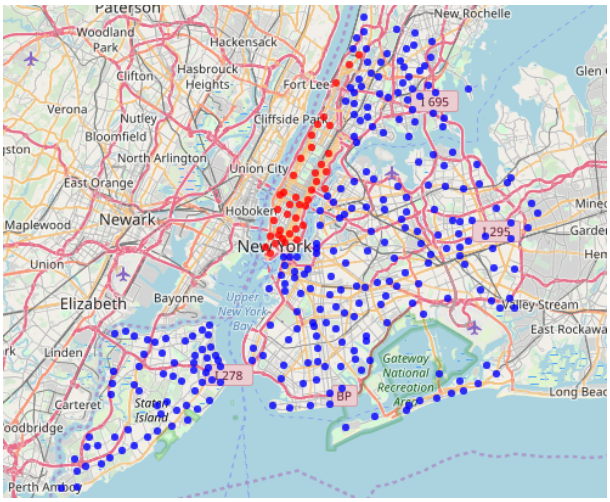

Figure 1: New York neighborhoods in red. In blue we have the neighborhoods of Manhattan that will be used in this work.

To get the names and geographic coordinates of all neighborhoods of Manhattan we resorted to the dataset: 2014 New York City Neighborhood Names available for free in the New York University Spatial Data Repository (https://geo.nyu.edu/) and filter the Borough to match Manhattan. In figure [1] we have all neighborhoods in New York and in red we have only the ones that belong to Manhattan and will be used in this work.

To get the Trending restaurants we resort to the Foursquare API available in (https://developer.foursquare.com/). First, we need to create a free developer account. After that we create a python script to get the top 40 venues for each neighborhood. An example of the dataset is shown in table [01].

Table 1: Example of Marble Hill trending venues

Name	Category	Lat.	Lon.
Cold Cut City	Sandwich Place	40.88	-73.91
Kam Wah	Asian Restaurant	40.88	-73.91
Gold Mine	Café	40.88	-73.90
Q'Kachapa	Spanish Restaurant	40.88	-73.90
Sam's Pizza	Pizza Place	40.88	-73.91

Given that we are now able to procced with the metodology.

Methodology

Results

Discussion

Conclusions

Acknowledgment

A special thank's to all instructors from IBM Data Science Specialization for the great work providing us with unvalueabe content through all specialization courses making this project (and all other) possible.

References

(1) Castano, E.; Yzerbyt, V.; Paladino, M.-P.; Sacchi, S. I belong, therefore, I exist: Ingroup identification, ingroup entitativity, and ingroup bias. *Personality and Social Psychology Bulletin* **2002**, *28*, 135–143.

(2) Ellemers, N.; Spears, R.; Doosje, B. Sticking together or falling apart: In-group identification as a psychological determinant of group commitment versus individual mobility. *Journal of personality and social psychology* **1997**, *72*, 617.