Immigrant location recommendation using machine learning

Rashmith Reddy Boppidi

Introduction

Business problem:

People who always wanted to move to toronto often fail to choose the place to move in toronto surrounded with their similar interests of living, this issue can be solved if they can find out the neighbourhood which are similar to their interest. This problem can be solved by clustering.

Objective:

To find similar neighborhood for an immigrant if he decides to move to toronto.

Target Audience:

Any immigrant who wants to move into toronto.

Stakeholders:

1)Any organisation who can act as a middle party to recommend their clients.

2)And individual Client by himself

Data acquisition

1. List of postal codes of Canada: M from Wikipedia. (https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)
2. Coordinates of the Toronto Neighbourhoods from csv file. (http://cocl.us/Geospatial_data)
3. Foursquare Location data. (http://foursquare.com/)
4. Cleaned data had 3 features.
5. Rows which were not available were removed.

K-means clustering for clustering the neighbourhoods

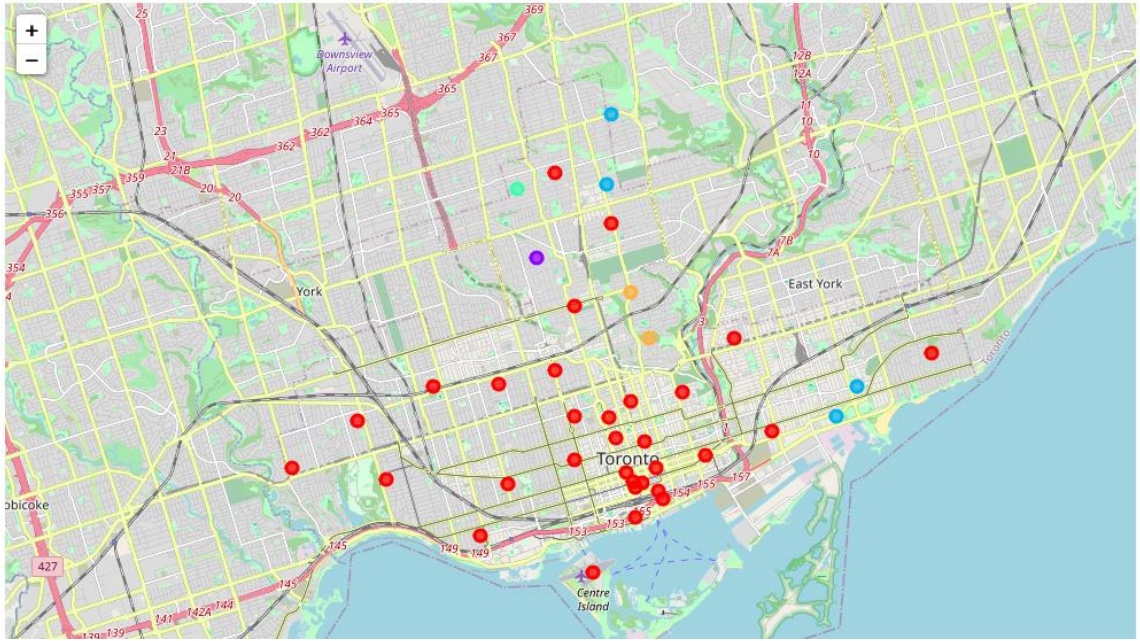
Based on the basis of common venues k-means clustering Machine Learning is used, which is unsupervised, in order to cluster the neighbourhoods together.

K = 5 was selected to go with.

```
        color=rainbow[cluster-1],
        fill=True,
        fill_color=rainbow[cluster-1],
        fill_opacity=0.7).add_to(map_clusters)

map_clusters
```

Out[78]:



Visualization of different clusters

Cluster 1

```
In [87]: toronto_merged.loc[toronto_merged['Cluster Labels'] == 0, toronto_merged.columns[[1] + list(range(5, toronto_merged.sha
```

Out[87]:

	borough	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	East Toronto	0	Trail	Neighborhood	Health Food Store	Pub	Doner Restaurant	Diner	Discount Store	Distribution Center	Dog Run	Yoga Studio
20	Downtown Toronto	0	Coffee Shop	Café	Hotel	Restaurant	Salad Place	Seafood Restaurant	American Restaurant	Japanese Restaurant	Italian Restaurant	Beer Bar
21	Downtown Toronto	0	Coffee Shop	Restaurant	Café	Hotel	Gym	American Restaurant	Japanese Restaurant	Seafood Restaurant	Italian Restaurant	Beer Bar
24	Central Toronto	0	Café	Sandwich Place	Coffee Shop	Pizza Place	Pub	Donut Shop	BBQ Joint	Indian Restaurant	History Museum	Pharmacy
25	Downtown Toronto	0	Café	Restaurant	Bar	Bookstore	Sandwich Place	Japanese Restaurant	Bakery	Chinese Restaurant	Dessert Shop	Pub

Cluster 2

```
n [80]: toronto_merged.loc[toronto_merged['Cluster Labels'] == 1, toronto_merged.columns[[1] + list(range(5, toronto_merged.shap
```

ut[80]:

	borough	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
23	Central Toronto	1	Park	Jewelry Store	Trail	Sushi Restaurant	Dessert Shop	Ethiopian Restaurant	Electronics Store	Eastern European Restaurant	Dumpling Restaurant	Donut Shop

Cluster 3

```
[81]: erged.loc[toronto_merged['Cluster Labels'] == 2, toronto_merged.columns[[1] + list(range(5, toronto_merged.shape[1]))]]
```

[81]:

	borough	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
5	Central Toronto	2	Gym	Hotel	Department Store	Sandwich Place	Breakfast Spot	Food & Drink Shop	Park	General Travel	Gluten-free Restaurant	Eastern European Restaurant
4	Central Toronto	2	Dim Sum Restaurant	Park	Bus Line	Swim School	Yoga Studio	Ethiopian Restaurant	Electronics Store	Eastern European Restaurant	Dumpling Restaurant	Donut Shop
2	East Toronto	2	Park	Brewery	Movie Theater	Ice Cream Shop	Sushi Restaurant	Fast Food Restaurant	Restaurant	Pub	Italian Restaurant	Board Shop
38	East Toronto	2	Light Rail Station	Yoga Studio	Garden Center	Skate Park	Restaurant	Recording Studio	Pizza Place	Park	Garden	Spa

Cluster 4

output; double click to hide

```
nto_merged['Cluster Labels'] == 3, toronto_merged.columns[[1] + list(range(5, toronto_merged.shape[1]))]
```

:

	borough	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
22	Central Toronto	3	Home Service	Music Venue	Garden	Department Store	Ethiopian Restaurant	Electronics Store	Eastern European Restaurant	Dumpling Restaurant	Donut Shop	Doner Restaurant

Cluster 5

```
3]: erged.loc[toronto_merged['Cluster Labels'] == 4, toronto_merged.columns[[1] + list(range(5, toronto_merged.shape[1]))]]
```

3]:

	borough	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
10	Downtown Toronto	4	Park	Playground	Trail	Yoga Studio	Department Store	Electronics Store	Eastern European Restaurant	Dumpling Restaurant	Donut Shop	Doner Restaurant
8	Central Toronto	4	Restaurant	Park	Trail	Deli / Bodega	Electronics Store	Eastern European Restaurant	Dumpling Restaurant	Donut Shop	Doner Restaurant	Dog Run

Result and Observations

We have successfully divided our data into 5 clusters. Now the person can select the locations based on his similarity. most of the neighbourhood fall under cluster 0, they have mostly coffee shops and cluster 1 has park, cluster 2 has gym and restaurants with swimming pool has most common , cluster 3 home services, cluster 4 parks and other activity places