**Predicting the Best Location for Restaurants in North York Toronto**

1. **Introduction**
   1. **Background**

Toronto is one of the most popular cities in the world for immigrates. The diversity of culture brings a diverse food culture as well. North York is the outer suburbs of Toronto, growing into the secondary business districts outside Downtown Toronto. High-rise development in the area has given the former municipalities distinguishable skylines of its own, with high-density transit corridors serving.

* 1. **Problem**

Restaurants are always popular business across the world. Location is the key for profit. However, finding the best location to start a restaurant in the city can be difficult. This project aims to find the most popular neighbourhoods in North York to establish a new restaurant.

* 1. **Interest**

This project may potentially benefit entities who wants to start a new restaurant in North York Toronto. As well as investor who are looking for the area in North York where restaurants are the most popular venue choice.

1. **Data acquisition and cleaning**
   1. **Data sources**

The Toronto post code, borough and neighbourhood data are scrapped from [here](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M), and the venue details are provided by Foursquare.

* 1. **Data cleaning**

Geospatial co-ordinate is one of the key elements for API requests from Foursquare. Therefore, after scraping the data frame of the neighbourhoods in Toronto, geospatial co-ordinates are imported from csv file and joined with existing data frame, sharing the primary key as neighbourhood.

When the geospatial co-ordinate of North York is acquired, API requests were made for the detailed venues of North York.

Missing data were dropped, venues are returned with name, location and categories, categorised by neighbourhoods in North York.

* 1. **Feature selection**

Only the popularity feature of venues is selected during the feature engineering process, given the purpose of the project is to find the popularity of different venues. 99 unique venue categories were filtered out and one hot encoded for model fitting. The target is the name of neighbourhood, where Chinese restaurant is the first most common venue.

1. **Exploratory Data Analysis**
   1. **Grouping neighbourhood**

As mentioned above, the target label of this project is neighbourhood in North York Toronto. Therefore, data collected from Foursquare related to venues are grouped by neighbourhood as shown in Figure 1 below. Venue types are not evenly distributed in different areas. The variety of venues in Fairview, Henry Farm, Oriole is significantly higher than other areas, which can indicate this is a more centre district in North York, having the potential of attracting more customer traffic.

*Figure 1*

* 1. **Categorising venues**

Known the variety of venues in each neighbourhood in North York is insufficient for the project, based on the outcome data frame, 97 unique venue categories are identified and shown as below.



*Figure 2*

As the target of this project is restaurant, the popularity of this venue category will be used as main criteria.

1. **Predictive Modelling**
   1. **Model selection**

Classification models will be used as the target label is categorical. This project adopted a K-mean clustering model for evaluation. K=5 is set as hyperparameter for this model. The outcome of this model is shown in the map below.

Map

Description automatically generated

*Figure 3*

Each colour represents a cluster of neighbourhoods based on the 10 most popular venue categories. The red dots represent a cluster marked as 0, which most of the popular venues are restaurants.

* 1. **Model evaluation**

Model is evaluated by examine clusters. The results are show in Figure 4.

Table

Description automatically generated

*Figure 4*

Most top 10 popular venues among these neighbourhoods are restaurants and have a variety of different venues to attract more customers. Therefore, when building new restaurants, neighbourhood within cluster 0 can be recommended to investors.

1. **Conclusions**

Victoria Village, Lawrence Manor, Lawrence Heights, Don Mills North, Glencairn, Don Mills South, Hillcrest Village, Bathurst Manor, Wilson Heights, Downsview North, Fairview, Henry Farm, Oriole, Northwood Park, York University, Bayview Village, Downsview East, Downsview West, North Park, Maple Leaf Park, Upwood Park, Downsview Central, Bedford Park, Lawrence Manor East, Willowdale South, Downsview Northwest and Willowdale West are the popular neighbourhoods in North York Toronto for restaurants.