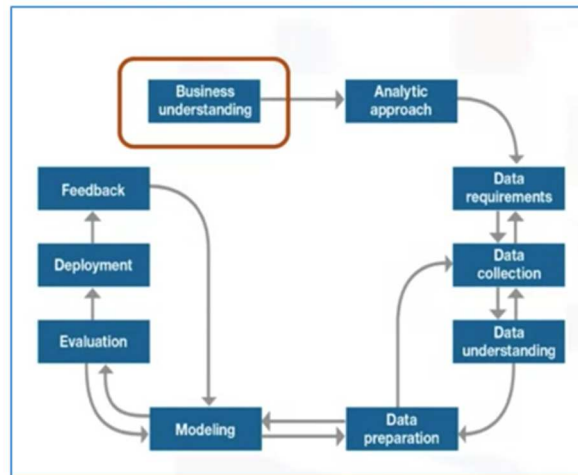


## Capstone Project - Battle of Neighbourhoods (Week 4) – Part2

### 2. Data

We will take advantage of the addresses used in the third week assignment to create this algorithm. We will put into practice techniques of scrapping, filtering and data transforming information.



Here are the main sources used:

- Wikipedia ( [https://en.wikipedia.org/wiki/List\\_of\\_postal\\_codes\\_of\\_Canada:\\_M](https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M)) to get Postal Code, Borough and Neighbourhood in Toronto

Toronto - 103 FSAs [ + edit ]

Note: There are no rural FSAs in Toronto, hence no postal codes should start with suggesting that Canada Post may have reserved the M0 FSA for high volume add

Postal Code	Borough	
M1A	Not assigned	Not assigned
M2A	Not assigned	Not assigned
M3A	North York	Parkwoods
M4A	North York	Victoria Village
M5A	Downtown Toronto	Regent Park, Harbourfront
M6A	North York	Lawrence Manor, Lawrence Heights
M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government
M8A	Not assigned	Not assigned
M9A	Eglinton	Islington Avenue, Humber Valley Village
M1B	Scarborough	Malvern, Rouge
M2B	Not assigned	Not assigned
M3B	North York	Don Mills
M4B	East York	Parkview Hill, Woodbine Gardens

- Geospatial data for Toronto ([http://cocl.us/Geospatial\\_data](http://cocl.us/Geospatial_data)) to get the geographical coordinates of each postal code

```
Postal Code, Latitude, Longitude
M1B, 43.8066863, -79.1943534
M1C, 43.7845351, -79.1604971
M1E, 43.7635726, -79.1887115
M1G, 43.7709921, -79.2169174
M1H, 43.773136, -79.2394761
M1J, 43.7447342, -79.2394761
M1K, 43.7279292, -79.2620294
M1L, 43.7111117, -79.2845772
M1M, 43.716316, -79.2394761
M1N, 43.692657, -79.2648481
M1P, 43.7574096, -79.273304
M1R, 43.7500715, -79.2958491
```

- Foursquare API to obtain more information about venues



- Random user data that will be used to train the model and later show the results.

```
['Italian Restaurant',  
'Asian Restaurant',
```