New shop analysis

Comperison between MCD shop or SBX shop

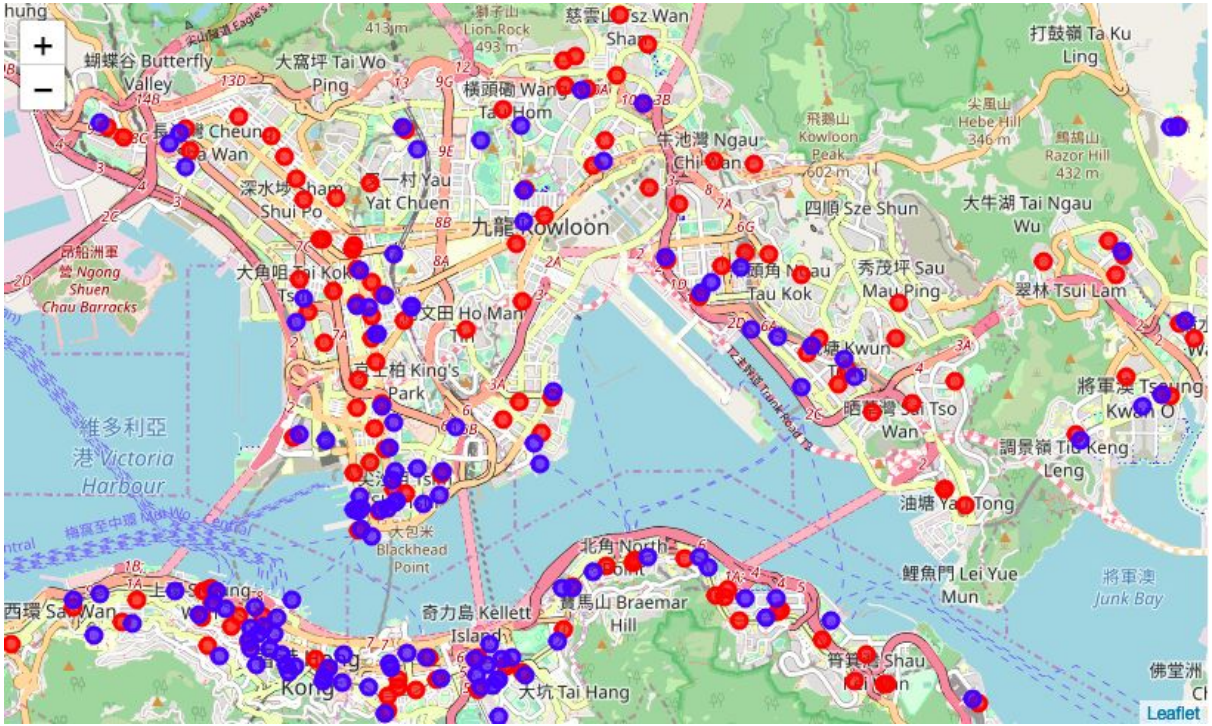
Introduction

XYZ Co. is a shop properties owner in Hong Kong. Recently they have a tenant moved out, so they have a shop for lease. MCD and SBX are the potential talent. We will analyze the nearby venues of the two chains in Hong Kong. And analyze the nearby venues of XYZ Co.’s place, predict the locals like which one more.

Data

1. MTR station location data generated from the website
2. MCD and SBX location data generated from Foursquare API and MTR locations
3. MCD and SBX nearby venues generated by Foursquare API

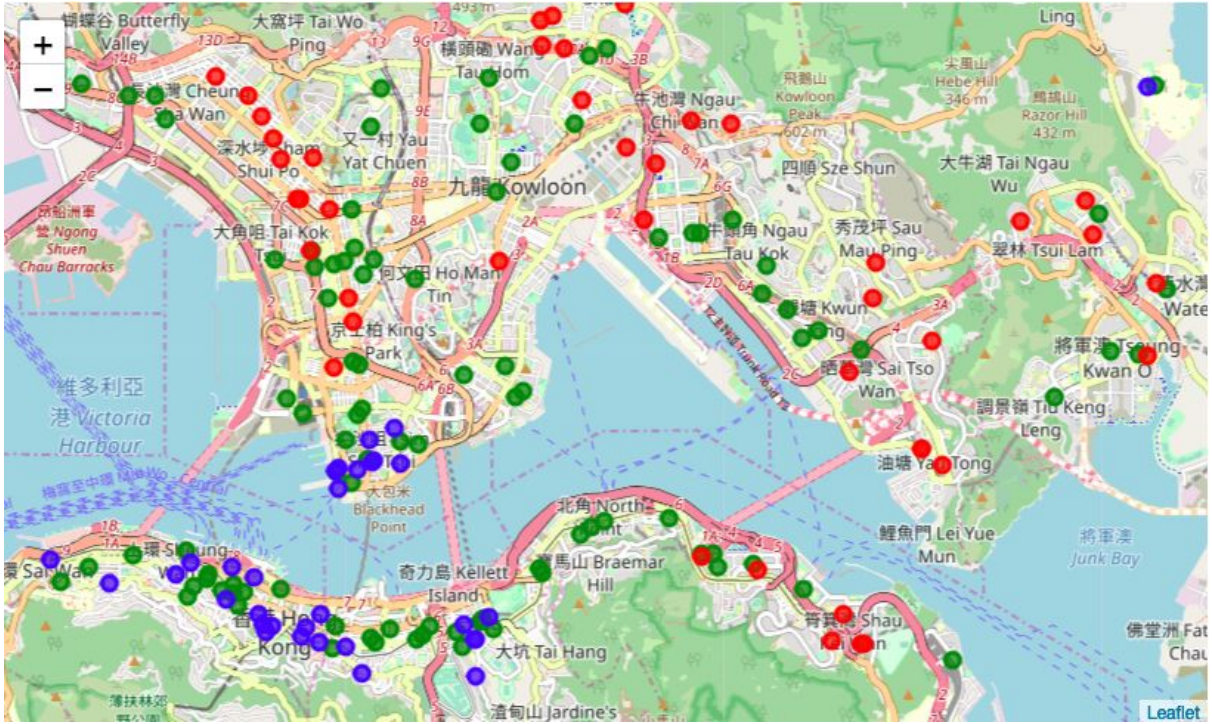
MCD and SBX locations



Blue: SBX shops

Red: MCD shops

MCD and SBX locations with combined points



Blue: SBX shops

Red: MCD shops

Green: 2 shops

Classification - Evaluation

	K	Accuracy Score
0	1	0.555849
1	2	0.545188
2	3	0.524354
3	4	0.559126
4	5	0.524208
5	6	0.534918
6	7	0.548856
7	8	0.573552
8	9	0.524795

KNN

	Solver	C	Accuracy Score
0	newton-cg	0.01	0.464984
1	newton-cg	0.10	0.472124
2	newton-cg	1.00	0.569885
3	lbfgs	0.01	0.464935
4	lbfgs	0.10	0.465033
5	lbfgs	1.00	0.563380
6	sag	0.01	0.464887
7	sag	0.10	0.472222
8	sag	1.00	0.556289
9	saga	0.01	0.465180
10	saga	0.10	0.475352
11	saga	1.00	0.580350

Logistic regression

	Kernel	Accuracy Score
0	linear	0.510808
1	poly	0.487945
2	rbf	0.480373
3	sigmoid	0.476538

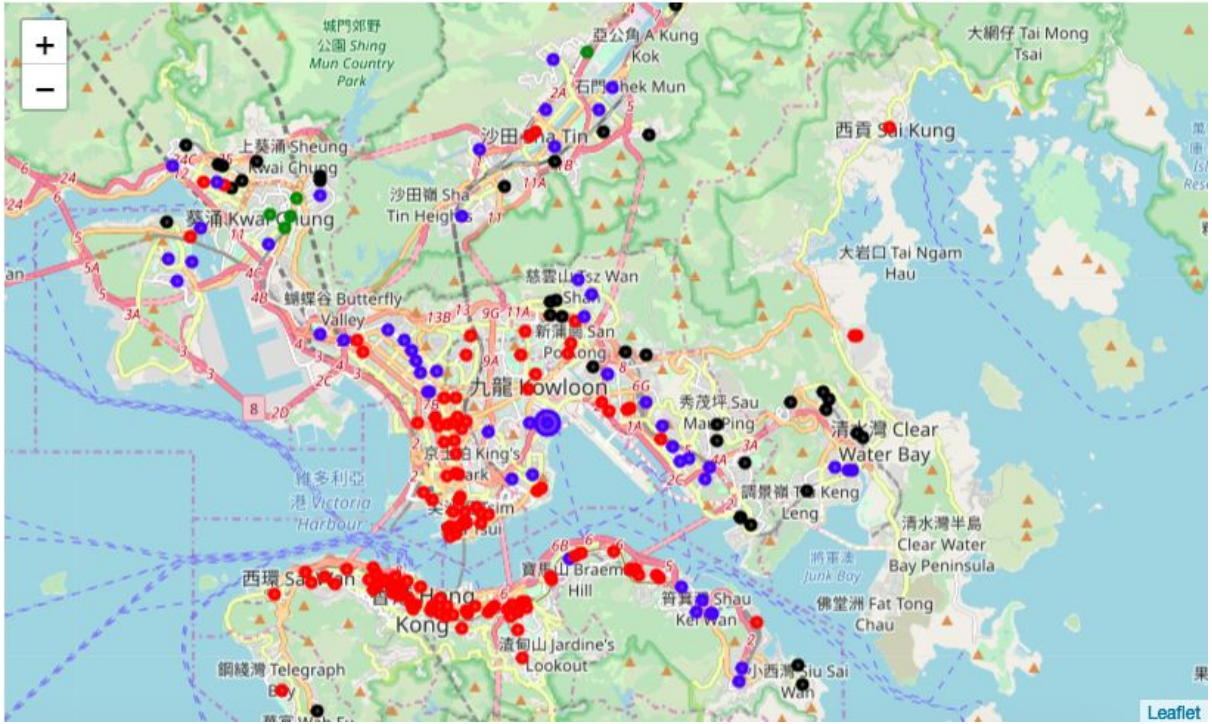
SVM

Classification - Result

After we examine the above 3 algorithms with different parameters, we will choose to use the “K-nearest neighbor” algorithm with K = 8 to build the model. And the prediction is Type 0.

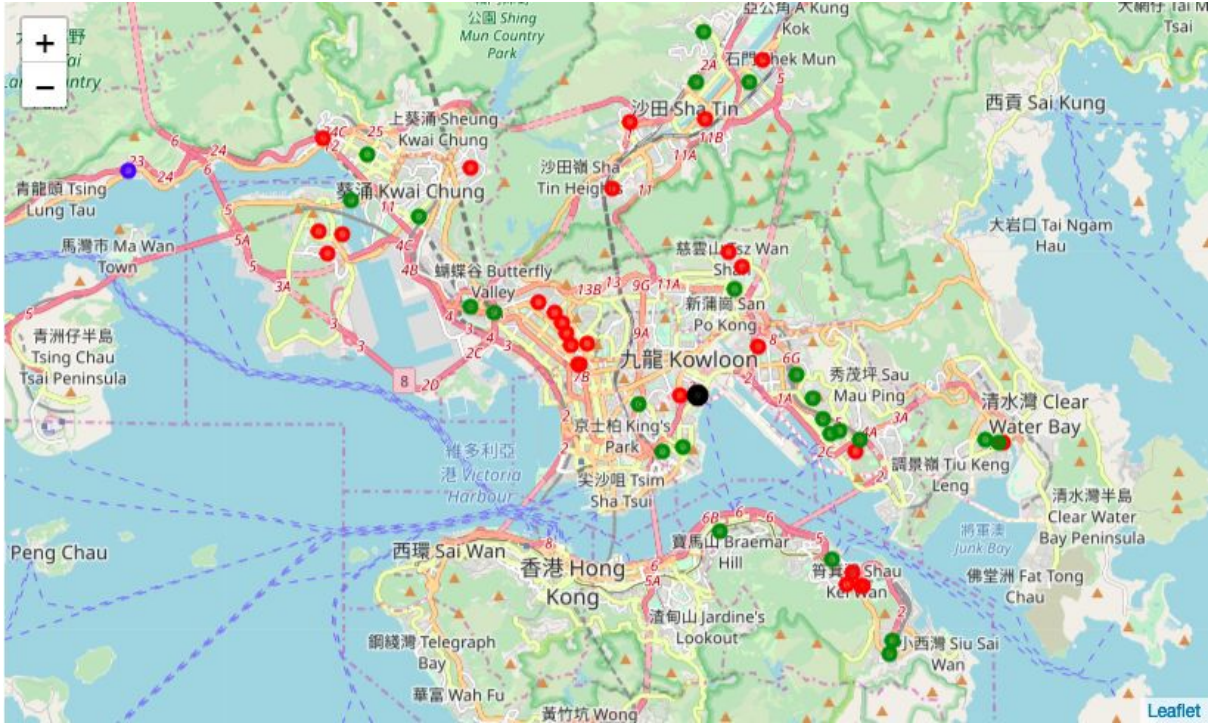
It means the model suggest the MCD shop.

Clustering



We cluster all the 2 chains' shops and the target shops into 5 categories. And the biggest circle are our target shop. It shows the target shop belongs to category 1 (in blue).

Clustering (analyzing Category 1)



This shows that more point in Blue Category are MCD shops (in red) and the place sof 2 shops (in green). Only a few points are SBX shops (in blue).

Conclusion

After we analyze the data by Classification and Clustering, MCD is more likely to open a shop where nearby venues are similar to XYZ shop. The classification result shows the XYZ shop's nearby environment is more similar to MCD shops. And the clustering result shows the XYZ shop is clustered to a category that there are far more MCD shops than SBX shops.