



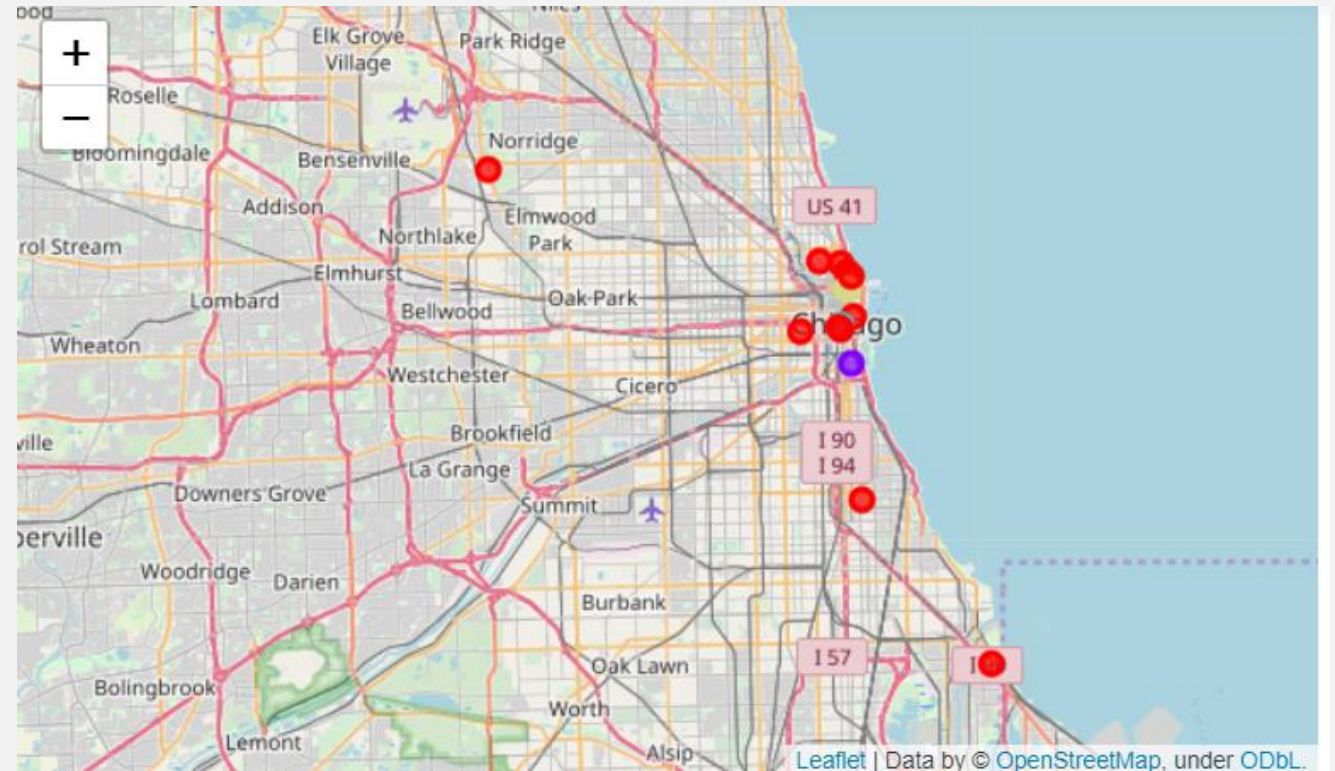
IBM APPLIED DATA SCIENCE CAPSTONE

- Handan Tan
- 26/03/2020

OUTLINE

- Introduction
- Data Preparation
- Methodologies
 - Exploratory Data Analysis
 - Machine Learning
- Results
- Conclusion

Venue clustering of 14 neighborhoods in Chicago



INTRODUCTION

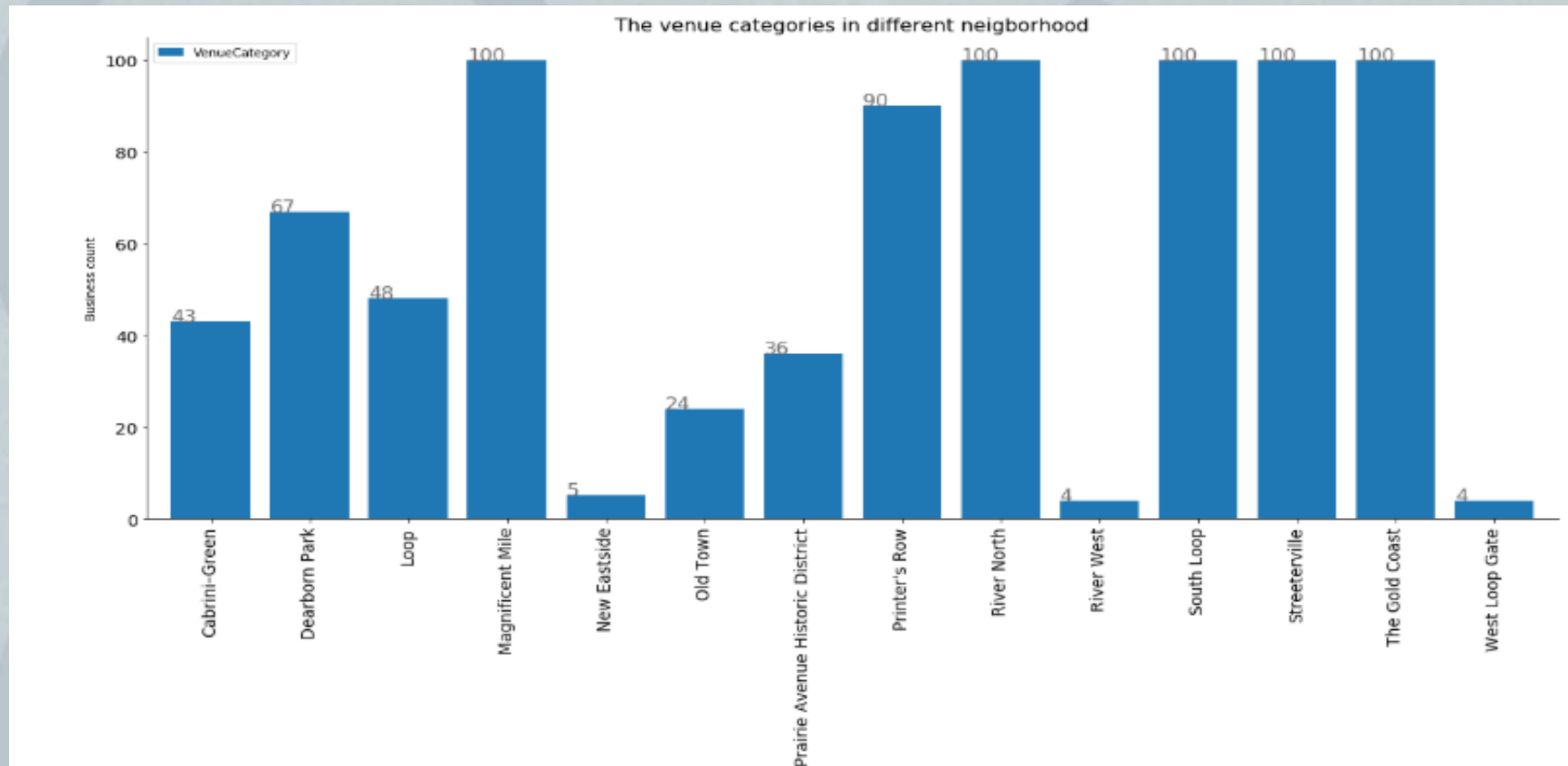
- Business problem: If I want to open an Italian restaurant in Chicago, which neighbourhood will be a good choice?
- Target audience: business owner/ someone who is interested in opening a new business.

DATA

- The 16 neighborhood names in Chicago central area:
- https://en.wikipedia.org/wiki/Community_areas_in_Chicago
- The coordinates of these neighborhoods from Geocoder.
- The venues in 500m radius of neighbourhood centre from Foursquare API.
- The venue information of Toronto which has a similar population size with Chicago.

METHODOLOGY: EXPLORATORY DATA ANALYSIS

- The number of venues in each neighbourhood:



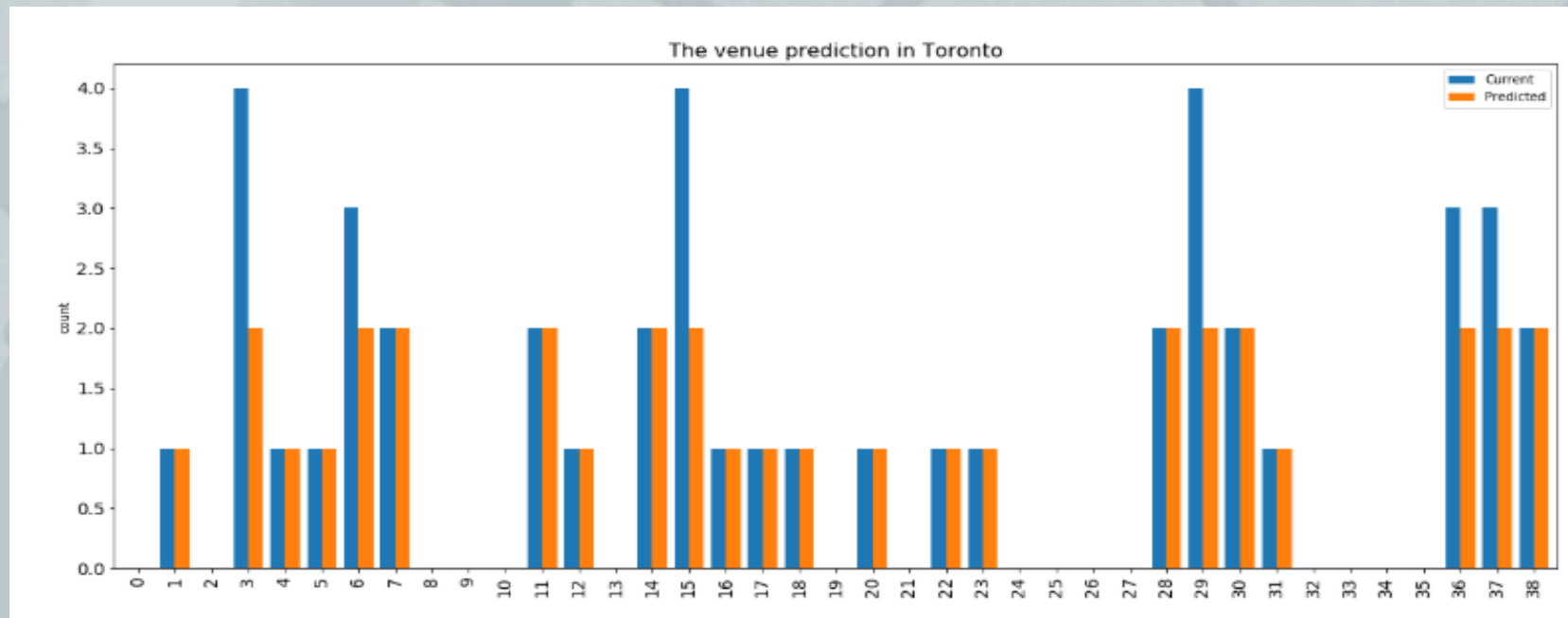
METHODOLOGY: EXPLORATORY DATA ANALYSIS

Top 10 venues in each neighbourhood. Italian restaurants is of a great popularity.

	Neighborhood	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	Cabrini–Green	Sandwich Place	Italian Restaurant	Pizza Place	Chinese Restaurant	Thai Restaurant	Bar	Salon / Barbershop	Bowling Alley	Fast Food Restaurant	Shipping Store
1	Dearborn Park	Pizza Place	Gym / Fitness Center	American Restaurant	Hotel	Tapas Restaurant	Coffee Shop	Furniture / Home Store	Indian Restaurant	Dance Studio	Cosmetics Shop
2	Loop	Harbor / Marina	Park	Boat or Ferry	Sushi Restaurant	Garden	Fountain	Museum	Music Venue	Concert Hall	Coffee Shop
3	Magnificent Mile	American Restaurant	Hotel	Boutique	Clothing Store	Italian Restaurant	Steakhouse	Restaurant	New American Restaurant	Coffee Shop	Salon / Barbershop
4	New Eastside	Pizza Place	Italian Restaurant	Shipping Store	Train Station	Beach	Yoga Studio	Donut Shop	Forest	Food Court	Fast Food Restaurant
5	Old Town	Park	Bar	Salon / Barbershop	Martial Arts Dojo	Middle Eastern Restaurant	Seafood Restaurant	Shipping Store	Coffee Shop	Bike Rental / Bike Share	Big Box Store
6	Prairie Avenue Historic District	Historic Site	American Restaurant	Pizza Place	Café	Park	Hotel	Candy Store	Restaurant	Seafood Restaurant	Rental Car Location
7	Printer's Row	Pizza Place	Coffee Shop	American Restaurant	Gym / Fitness Center	Hotel	Sandwich Place	Ice Cream Shop	Taco Place	Lounge	Pub
8	River North	Hotel	Italian Restaurant	American Restaurant	New American Restaurant	Pizza Place	Cosmetics Shop	Steakhouse	Coffee Shop	Seafood Restaurant	Gastropub
9	River West	Gym	Other Repair Shop	Forest	Shipping Store	Food Court	Fast Food Restaurant	Farmers Market	Entertainment Service	Dry Cleaner	Donut Shop
11	Streeterville	American Restaurant	Hotel	Clothing Store	Restaurant	Salon / Barbershop	Tea Room	Italian Restaurant	Women's Store	New American Restaurant	Sporting Goods Shop
12	The Gold Coast	Italian Restaurant	Clothing Store	American Restaurant	Gym	Café	Coffee Shop	Boutique	Steakhouse	Bar	Shoe Store
13	West Loop Gate	Bookstore	Fast Food Restaurant	Park	Dog Run	Dry Cleaner	French Restaurant	Fountain	Forest	Food Court	Farmers Market

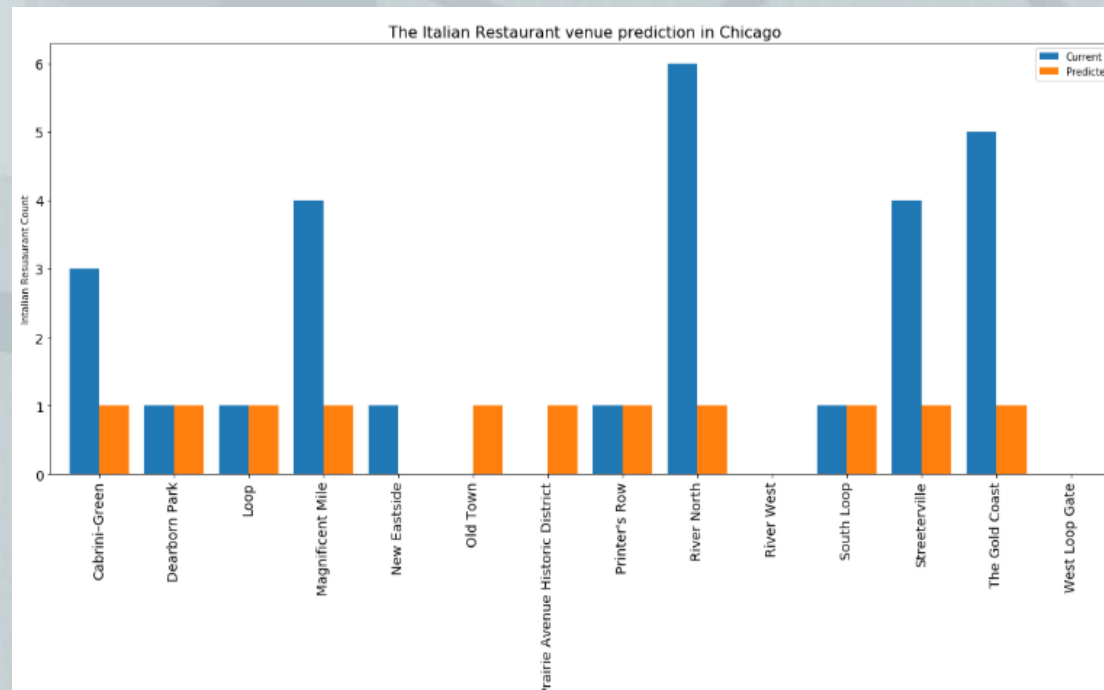
METHODOLOGY: REGRESSION

- The Support Vector Regression: Cross-validation score 0.29
- Grid search: best parameter combination $C=1.0$ and $\gamma=0.1$
- Evaluation: 37 neighborhoods in Toronto



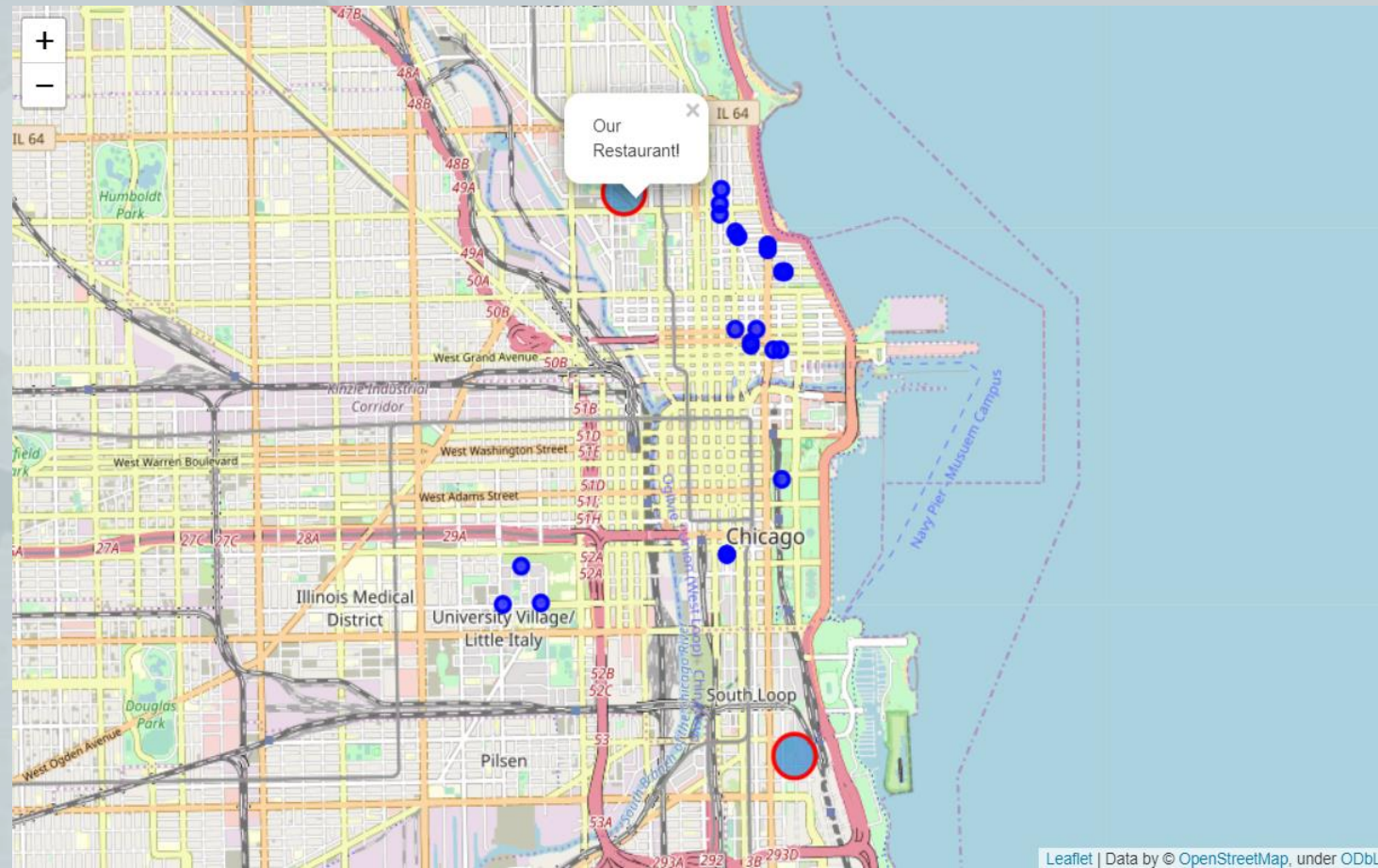
RESULT: REGRESSION

- Features: the number of all other categories in each neighbourhood
- Target: the number of Italian restaurant
- Test: the 235 category numbers of Chicago



RESULT: VISUALIZATION

- The existing Italian restaurants and two possible venues for new restaurant:



CONCLUSION

- Most neighborhoods have a larger number of Italian restaurant than the predicted number.
- The “Old Town” and “Prairie Avenue Historic District” are the only two neighborhoods which does not have an Italian restaurant as prediction. There are likely to be good locations to open a new restaurant.

POSSIBLE IMPROVEMENT

- A comparison of machine learning model trained by venues in different cities to derive the highest accuracy to predict the business environment in Chicago.
- The factors such as the rating of competitors, the public transportation, and parking may also have an influence for location selection.