Prediction Model to Measure Crowdedness Within Amsterdam

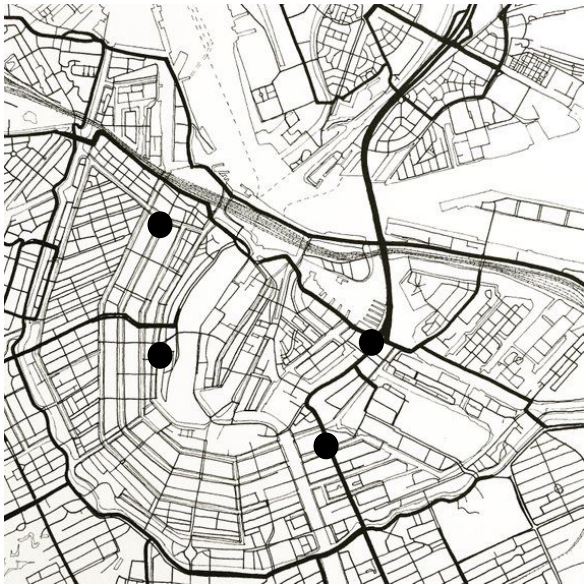
Don de Lange

Universiteit van Amsterdam

City of Amsterdam

Problem Statement

- No method to predict city-wide crowdedness
- Crowdedness unevenly measured within the city



How can a prediction of the level of crowdedness within the city of Amsterdam be given, based on input from city-wide available data sources?

Data

Data

- GVB Public transport data

	Sensor	Date	Hour	SensorLongitude	SensorLatitude	CrowdednessCount
0	GAWW-02	2018-03-11	0	4.898903	52.373786	0

- Crowdedness data localized data gathering sources

	Date	Event	Latitude	Longtitude
0	2018-04-20	Springsnow Festival	52.372638	4.894106

- Amsterdam Event Data

	Date	Hour	NieuwmarktCode	NieuwmarktLat	NieuwmarktLon	NieuwmarktArrivals
0	2018-04-01	0	NMT	52.371942	4.901239	31.0

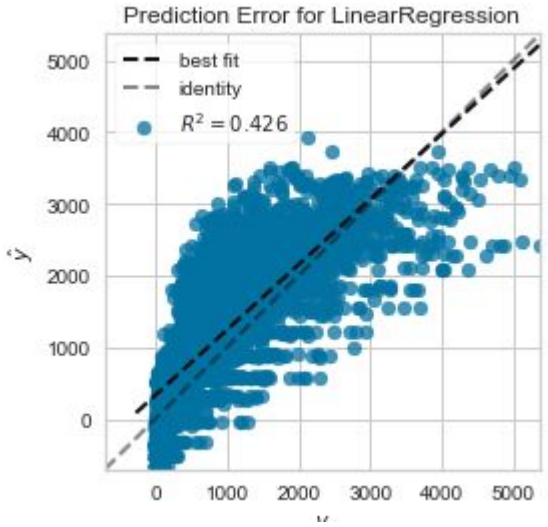
Time

- Time made circular with sinus and cosine
  - Makes 23:00 close to 01:00

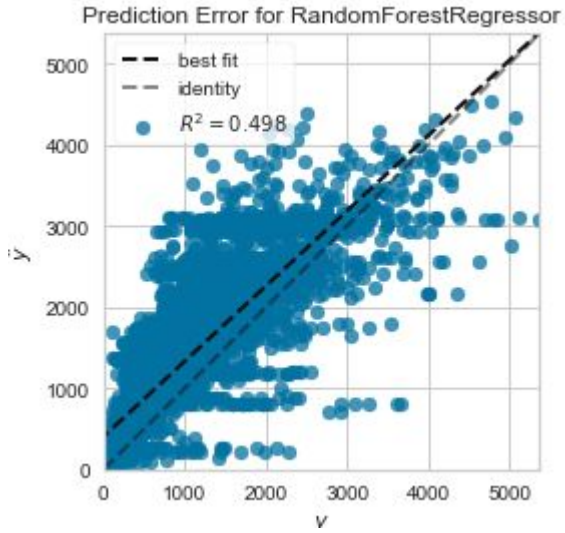
```
v["hour_sin"] = np.sin(2 * np.pi * v["Hour"] / 2400)
v["hour_cos"] = np.cos(2 * np.pi * v["Hour"] / 2400)
```

Models

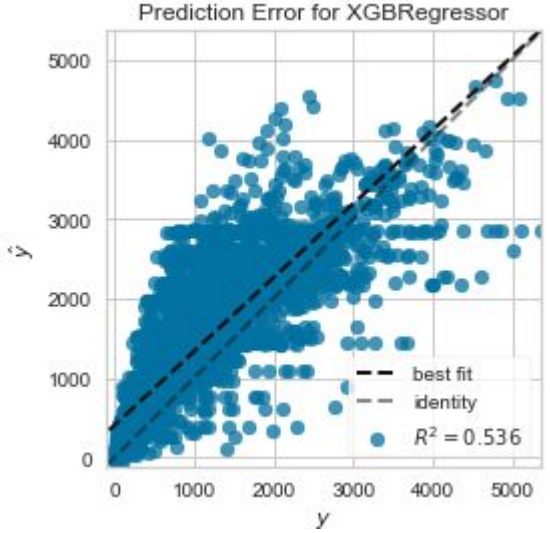
Regression - Evaluation



R^2 Score: 0.42610179006098087
MSE: 474492.7948214673
RMSE: 688.8343740126993



R^2 Score: 0.49838245282978344
MSE Test: 414731.9294018637
RMSE Test: 643.9968395899655



R^2 Score: 0.5361123190873034
MSE Test: 383537.2865562742
RMSE Test: 619.3038725506843

Classification - Evaluation

Majority Class

Accuracy Score: 0.07314814814814814

	precision	recall	f1-score	support
1	0.00	0.00	0.00	1429
2	0.00	0.00	0.00	1716
3	0.00	0.00	0.00	859
4	0.07	1.00	0.14	316
micro avg	0.07	0.07	0.07	4320
macro avg	0.02	0.25	0.03	4320
weighted avg	0.01	0.07	0.01	4320

Random Forest Classification

Accuracy Score: 0.5525462962962963

	precision	recall	f1-score	support
1	0.54	0.87	0.67	1429
2	0.75	0.36	0.49	1716
3	0.42	0.39	0.41	859
4	0.46	0.57	0.51	316
micro avg	0.55	0.55	0.55	4320
macro avg	0.54	0.55	0.52	4320
weighted avg	0.60	0.55	0.53	4320

XGBoost

Accuracy Score: 0.5833333333333334

	precision	recall	f1-score	support
1	0.43	1.00	0.60	3
2	0.50	0.75	0.60	4
3	0.70	0.64	0.67	11
4	1.00	0.17	0.29	6
micro avg	0.58	0.58	0.58	24
macro avg	0.66	0.64	0.54	24
weighted avg	0.71	0.58	0.55	24

Future

Future

- Increase Performance Models
- Add more data
- Add spatial dimension

Thank you for your time