PREDICTING CAR COLLISON SEVERITY

CAR ACCIDENTS- IMPROVING RESPONSE

- Seattle like all cities have car accidents that happen on a regular basis
- Car accident response is one of the biggest drain on first responder resources
- Car accidents can range from minor accidents where this is little to no property damage to accidents with fatalities
- The current procedure is for Seattle Police Department (SPD) to respond first, assess the accident and call in EMT or fire accordingly.
- This procedure leads to delay in treatment for those that are severely hurt and need immediate attention or possible extraction from the vehicle.

CAR ACCIDENTS- IMPROVING RESPONSE

- The city of Seattle wants fire and EMT to improve response time to accidents so that victims are treated quickly.
- Unfortunately, it is not feasible for fire and EMT to respond to every accident.
- The city would like to create a model that would allow them to predict the type of car accidents based on the characteristics of the accident as reported by bystanders or other non-first responder call in.

CAR ACCIDENTS- IMPROVING RESPONSE

DATA

- SPD maintains all accident reports in a log in CSV
- Log has characteristics of accidents as recorded by reporting police officer
- Data set has 38 characteristics
- 2 of them are Severity codes as 1 or 2. 1 not severe 2 is severe. The data is redundant so one can be removed.
- Some data is administrative
- Most of the data is in text but is uniform so the text does not vary for example. Rain is not described is rainy or described as showers. Road conditions are not described as wet and slippery, only as wet.

- The entire data set contained 194, 673 accidents.

CAR ACCIDENTS- IMPROVING RESPONSE

DATA CLEANING

- Redundant data is removed-duplicate data
- Text Data and date weekday is coded-as integers (1through n)
- Administrative data is also removed as it is not characteristic of the accident
- Any accidents with missing, Nan, unknown or other is dropped
- Final data set was made up of 1 label-Severity code and 7 characteristics
- Total size of 166,212 accidents

	SEVERITYCODE	X	Y	W_code	RDcond_code	ADDR_code	date_code	LT_code
0	2	-122.323148	47.703140	1	1	1	3	1
1	1	-122.347294	47.647172	2	1	2	3	2
2	1	-122.334540	47.607871	1	2	2	4	1
3	1	-122.334803	47.604803	3	2	2	5	1
4	2	-122.306426	47.545739	2	1	1	3	1

CAR ACCIDENTS- IMPROVING RESPONSE

- Not Severe to severe is 3:1
- Weather snowing- 80.2% not severe, Sleet- 74.5% not severe

Date- no noticeable difference

Road conditions are similar to Weather- snow and ice more likely to be not severe accidents

Weather

Severity	1- overcast	2- raining	3-clear	4- snowing	5-fog	6-sleet	7-blow sand	8-high wind
1 (not severe)	67.8%	65.8%	67.2%	80.2%	66.5%	74.5%	71.4%	70.8%
2 (severe)	32.2%	34.2%	32.8%	19.8%	33.5%	25.5%	28.6%	29.2%

Date

Severity	1- Monday	2- Tuesday	3- Wednesday	4- Thursday	5- Friday	6- Saturday	7- Sunday
1 (not severe)	66.4%	66.4%	66.5%	66.1%	67.5%	67.7%	69.2%
2 (severe)	33.5%	33.5%	33.5%	33.9%	32.5%	32.3%	30.8%

Road Conditions

Severity	1- Wet	2- Dry	3- snow/slush	4- ice	5- Sand/mud dirt	6- Standing water	7- Oil (recently paved road or oil spill)
1 (not severe)	66.3%	67.2%	81.2%	75.7%	62.5%	73.4%	59.2%
2 (severe)	33.7%	32.3%	18.8%	24.2%	37.5%	26.6%	40.8%

CAR ACCIDENTS- IMPROVING RESPONSE

- Not Severe to severe is 3:1
- Lighting – Dark with no streetlights show less severity

Location- intersections have higher instances of sever accidents

Lighting conditions

Severity	1- Daylight	2- Dark- light on	3-Dark No streetlights	4- Dusk	5- Dawn	6- Dark Streetlights off	7-Dark Unknown lighting
1 (not severe)	66.0%	69.3%	77.0%	66.1%	66.1%	72.1%	62.5%
2 (severe)	34.0%	30.7%	23.0%	33.9%	33.9%	27.9%	37.5%

Location type

Severity	1- intersections	2- Block
1 (not severe)	56.2%	73.4%
2 (severe)	43.8%	26.6%

CAR ACCIDENTS- IMPROVING RESPONSE

Data set was balanced and plotted on a heat map.

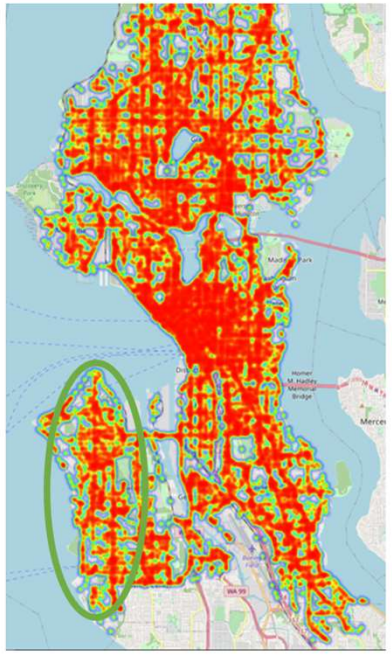


Figure 1 severe

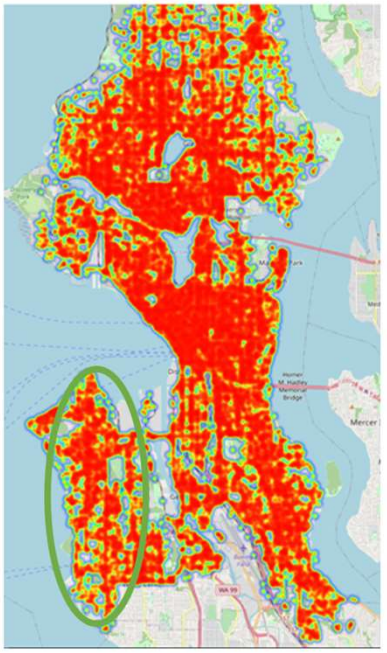
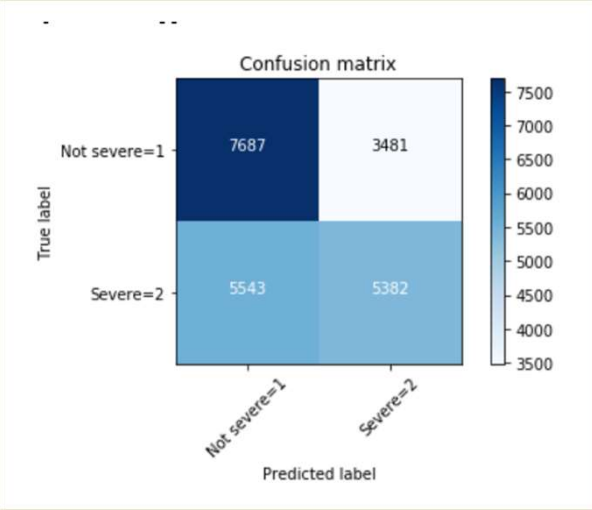


Figure 2 not Severe

CAR ACCIDENTS- IMPROVING RESPONSE

Logistical Regression Model

Jaccard Index	.592
Log Loss	.676

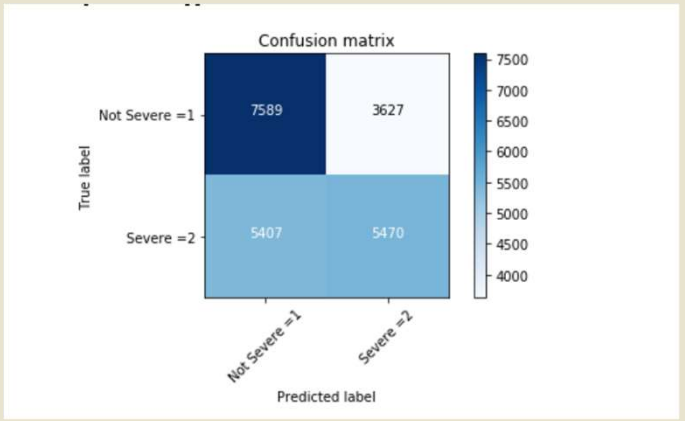


	precision	recall	f1-score	support
1	0.58	0.69	0.63	11281
2	0.60	0.48	0.54	10812
micro avg	0.59	0.59	0.59	22093
macro avg	0.59	0.59	0.59	22093
weighted avg	0.59	0.59	0.59	22093

CAR ACCIDENTS- IMPROVING RESPONSE

SVM Model

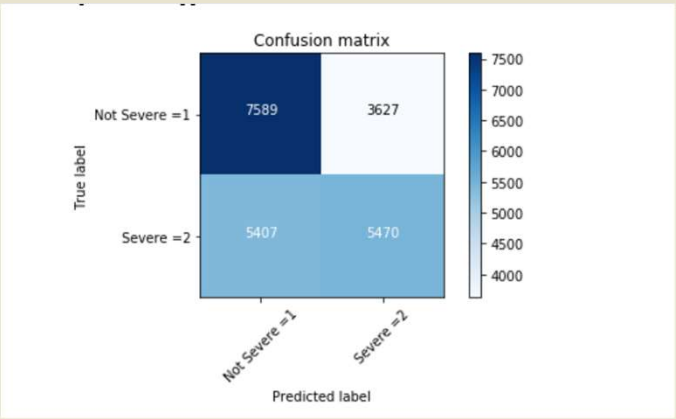
Jaccard Index	.591



	precision	recall	f1-score	support
1	0.58	0.68	0.63	11216
2	0.60	0.50	0.55	10877
micro avg	0.59	0.59	0.59	22093
macro avg	0.59	0.59	0.59	22093
weighted avg	0.59	0.59	0.59	22093

CAR ACCIDENTS- IMPROVING RESPONSE

SVM Model



Jaccard Index	.591

CAR ACCIDENTS- IMPROVING RESPONSE

Conclusion

- Model predictability was relatively low. Target was 75%+
- Model can be used to predict when a police officer is only required. Accuracy for that is 70%.
- Model can be deployed with the understanding that in many cases where a sever accident is predicted it will not be severe. This will be acceptable depending on resources available and will improve response time.

- Model can be improved by increasing the labels for severity. Example: 1 not severe, 2 bodily injury but can walk, 3 bodily injury cannot walk, 4- vehicle extraction required, 5 death.