Classification of Road Accident Severity

Value of Classification of Road Accident Severity

- Road accident can occur anytime and anywhere and it may cause heavy casualties. It is good for public to understand the potential risk factors of road traffic factors

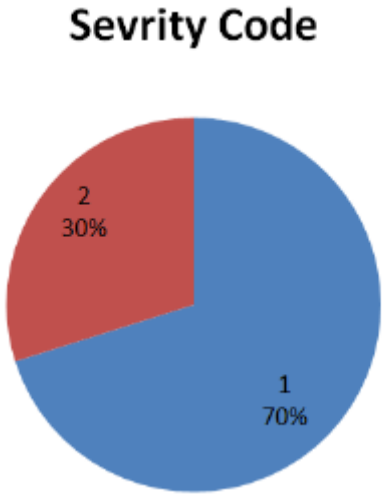
- Transportation bureau can take necessary precautions to minimize the occurrence of road accidents

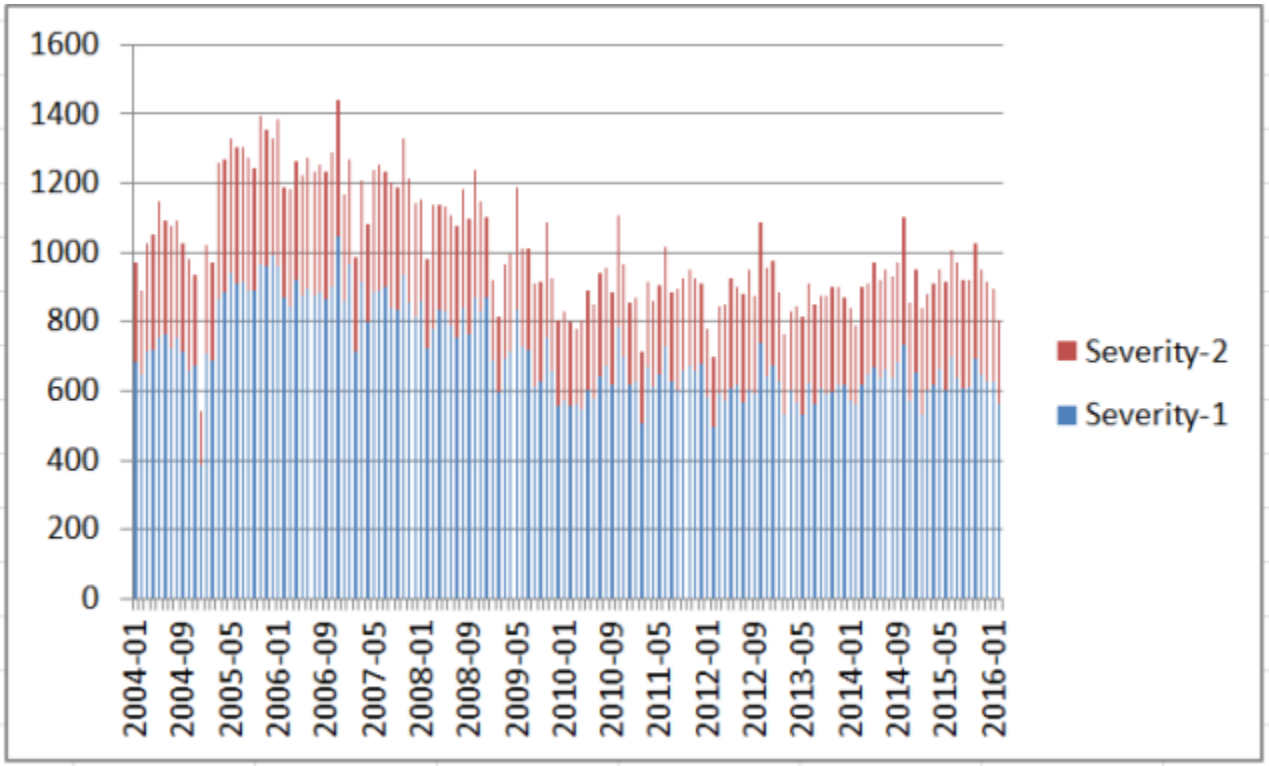
- Maths , Statistic major students and data scientist would also be interested in the development of classification model

Data Handling

- The data is source by public transportation data for Seattle city
- Total 194,673 records, 581,188 SEVERITYCODE=1 records and the rest are SEVERITYCODER=2
- 38 data columns, after data cleansing, 3 columns would be used in model development.

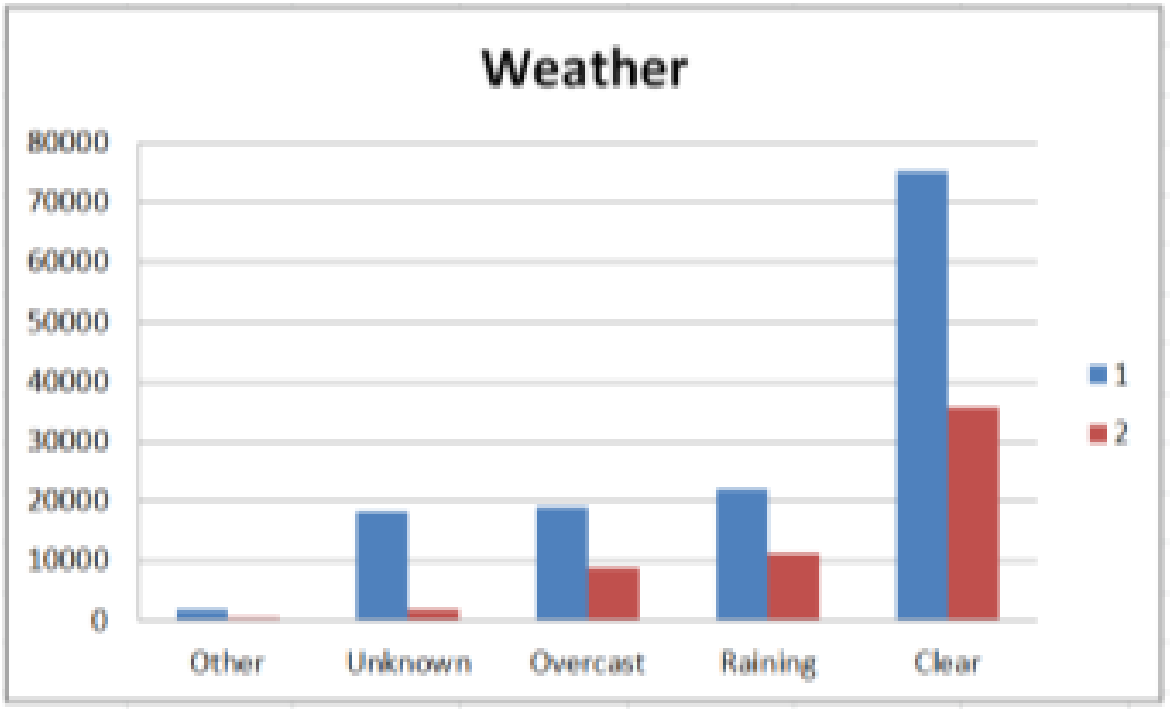
70% road accidents were classified as Severity 1





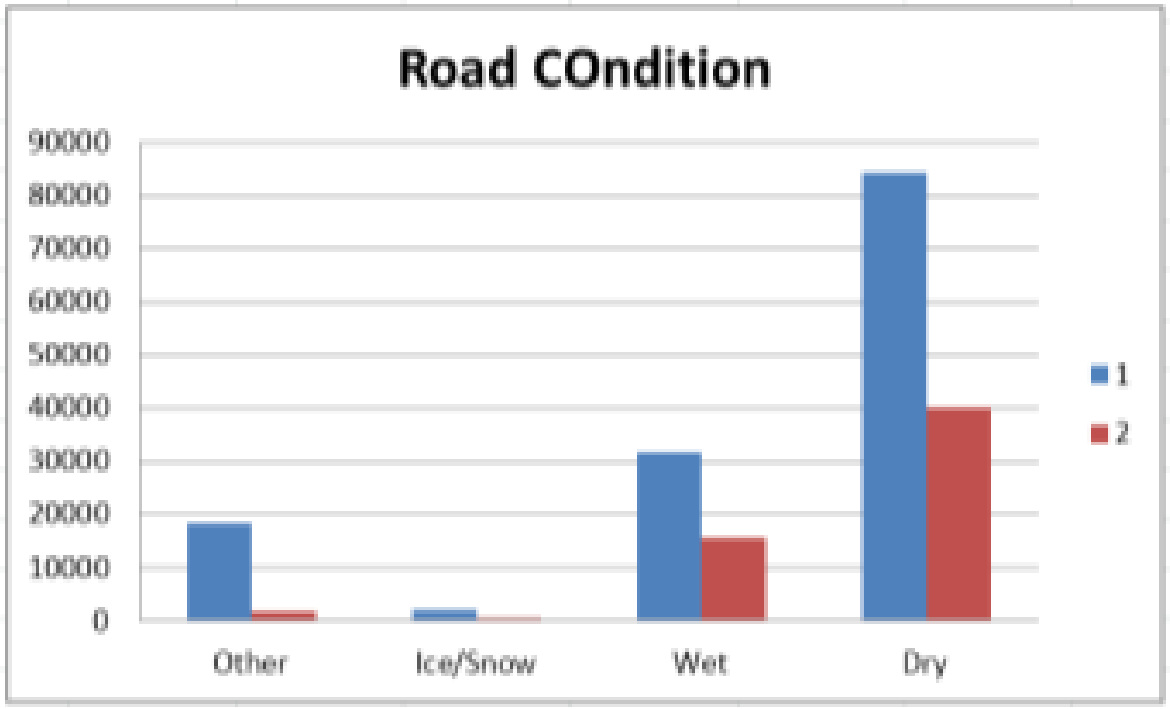
Distribution of the features against accident severity

Weather
Severity 1 happened most even it is clear weather



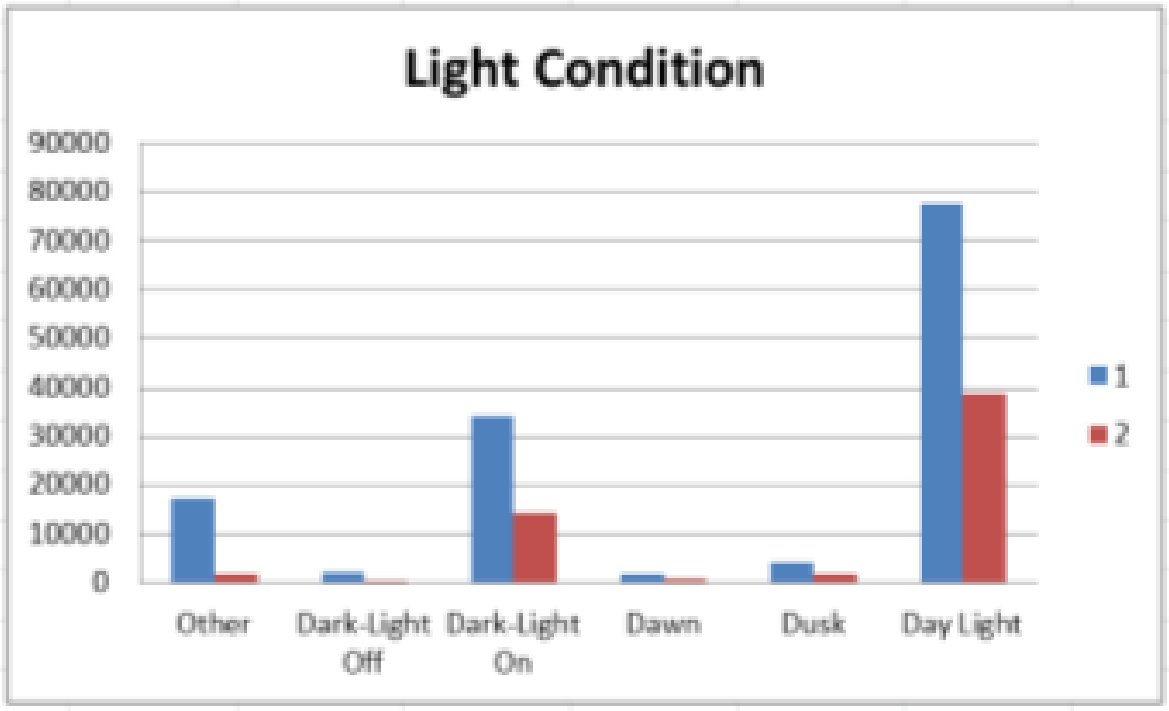
Distribution of the features against accident severity

Road Condition
Severity 1 happened most even the road is dry



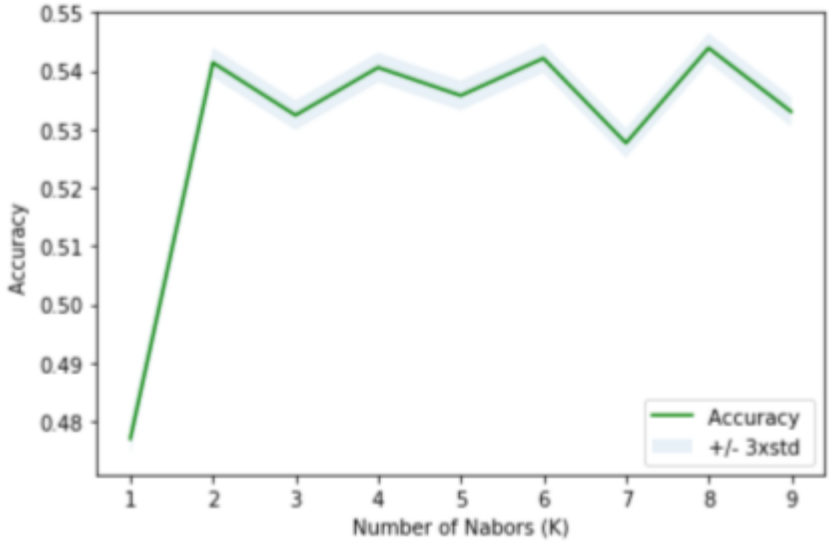
Distribution of the features against accident severity

Road Condition
Severity 1 happened most even there is day light



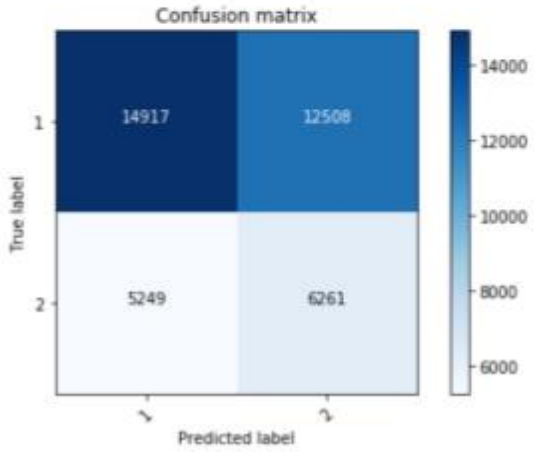
Model performance

Model accuracy reached the highest if number of K =8



	precision	recall	f1-score	support
1	0.74	0.54	0.63	27425
2	0.33	0.54	0.41	11510
micro avg	0.54	0.54	0.54	38935
macro avg	0.54	0.54	0.52	38935
weighted avg	0.62	0.54	0.56	38935

```
Confusion matrix, without normalization
[[14917 12508]
 [ 5249  6261]]
```



Conclusion and future directions

- Closely monitor the model performance
- Capture more diversified data features such as status during driving, model of the cars
- Engage more data scientists and transportation bureau officers in
  model development to share experience