

Predicting Severity of an accident.

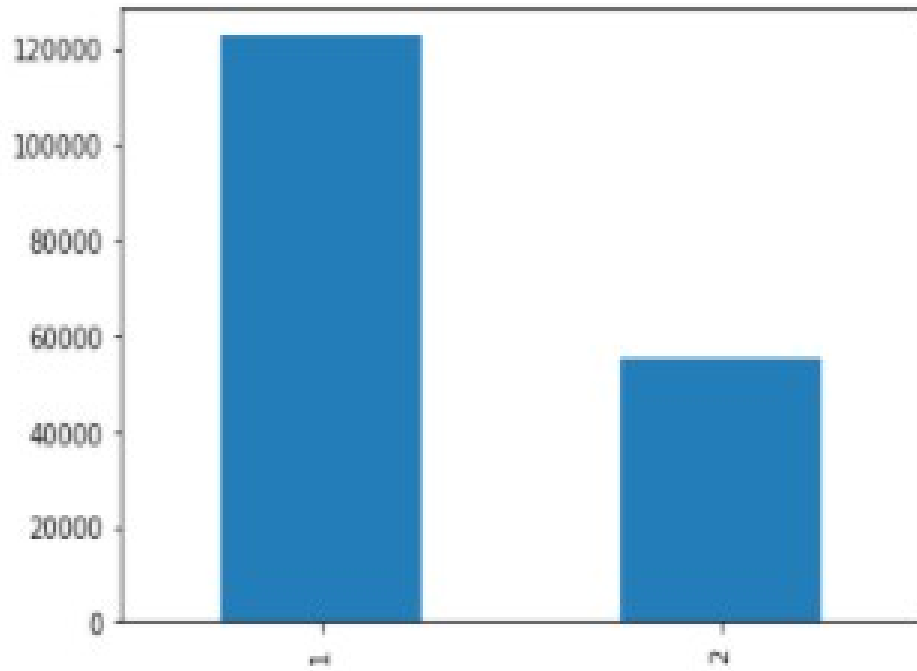
Predicting severity

- Predicting severity of an accident on a specific road would alert the authorities that this road needs to be fixed.
- Sign could be put at every road to warn drivers to be careful while driving.
- Accidents caused by light or road condition could be reduced by fixing the roads and increasing light stands.
- Accidents caused by natural causes such as rain or storms could be reduced by closing the roads which proved to be more dangerous during this conditions.

Data acquisition and cleaning

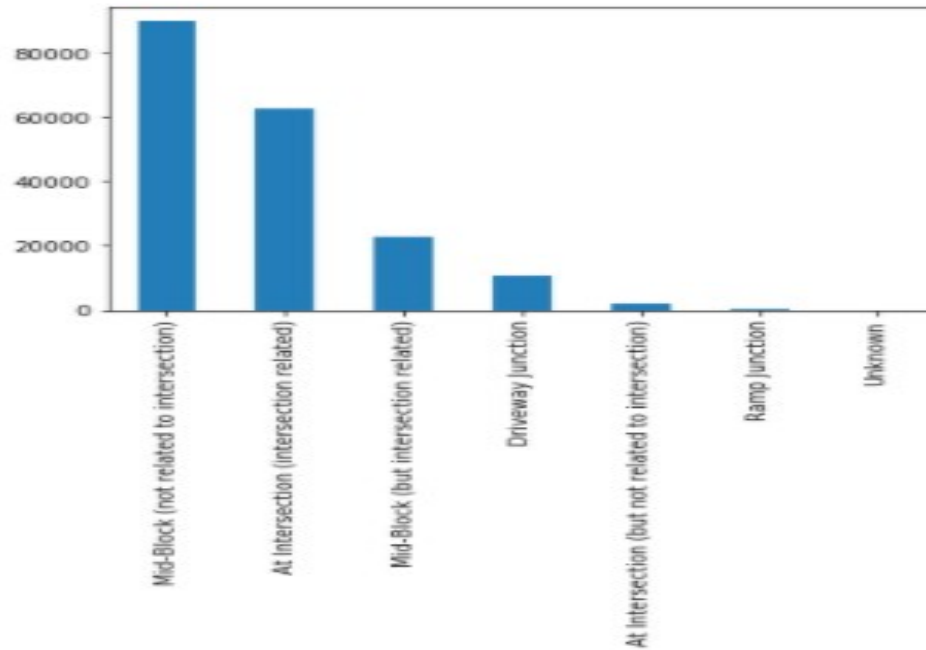
- This dataset is offered by ibm cloud.
- This data is collected by seattle police department
- It contains accidents from 2004 till present.
- It contains 194673 row and 38 Column
- Duplicate, highly similar or highly correlated features were dropped.
- Cleaned data contains 10 feature.

Severity values



From the given data, the severity code values are mostly 1 or 2

Junction types



Junction type plays a crucial part and it is a feature which will help to predict the severity of an accident.

Models

- There are many type of model that could be used such as Regression, Clustering, Classification etc.
- For this problem, the more suitable model would be the Classification model.
- The first used machine learning model is the Logistic Regression model.
- The second model is the Decision Tree model.

Logistic Regression

- For evaluation, Jaccard index and F1-score have been used as metrics for accuracy and score.

	precision	recall	f1-score	support	
1	0.73	0.98	0.84	24610	
2	0.81	0.21	0.33	11041	
micro avg	0.74	0.74	0.74	35651	DT Jaccard index: 0.74
macro avg	0.77	0.59	0.59	35651	DT F1-score: 0.68
weighted avg	0.76	0.74	0.68	35651	

Decision Tree

- Same evaluation metrics as Logistic Regression and the results came slightly better than the logistic Regression model

	precision	recall	f1-score	support	
1	0.73	0.99	0.84	24610	
2	0.88	0.19	0.31	11041	
micro avg	0.74	0.74	0.74	35651	DT Jaccard index: 0.74
macro avg	0.81	0.59	0.57	35651	DT F1-score: 0.68
weighted avg	0.78	0.74	0.68	35651	

Conclusion

- Built useful models to predict the severity of an accident.
- Accuracy of the models has room for improvement.
- The models should be used to help in reducing car accidents.