

Car accident severity

September 11, 2020 12:45 PM

I. Purpose of the project

The purpose of the project is to predict the severity of an accident, given a number of factors such as location, road condition, weather condition, car speeding, etc. The project will help drivers to understand the accident severity under certain conditions, so that they could drive with more care when necessary or even reschedule their trip to reduce the risk of car accidents. This project focus on the factors of weather condition, road condition and light condition.

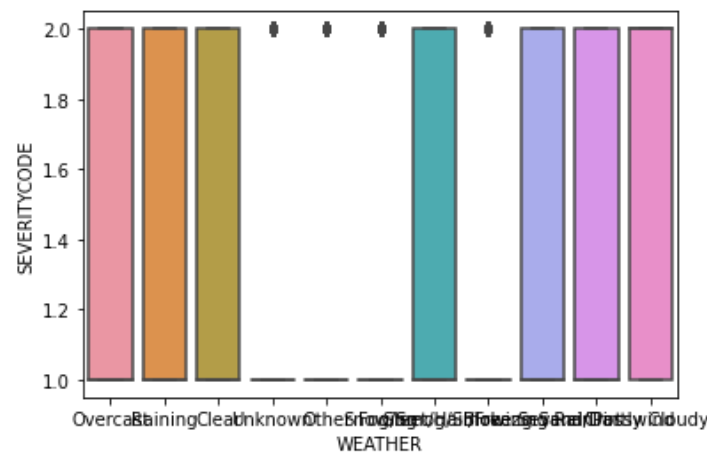
II. Data of the project

The data for the analysis is the car accident records based in Seattle city. The data were obtained through open source. It includes all types of collisions from 2004 to the present. The weather condition data includes clear, partly cloudy, overcast, raining, snowing, fog, sleet, blowing sand, severe crosswind. The road condition includes dry, wet, ice, snow, mud and standing water. The light condition includes daylight, streetlights on, dusk, dawn, dark-no street lights, etc.

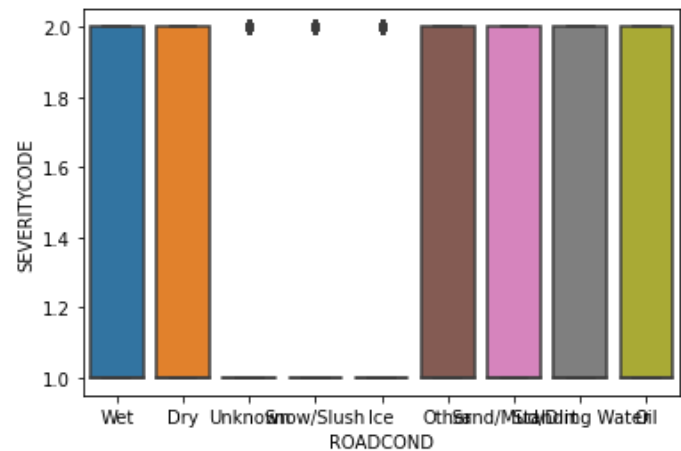
III. Methodology

The project started with exploring the severity of the accidents with a single factor. Box plots were generated as follows:

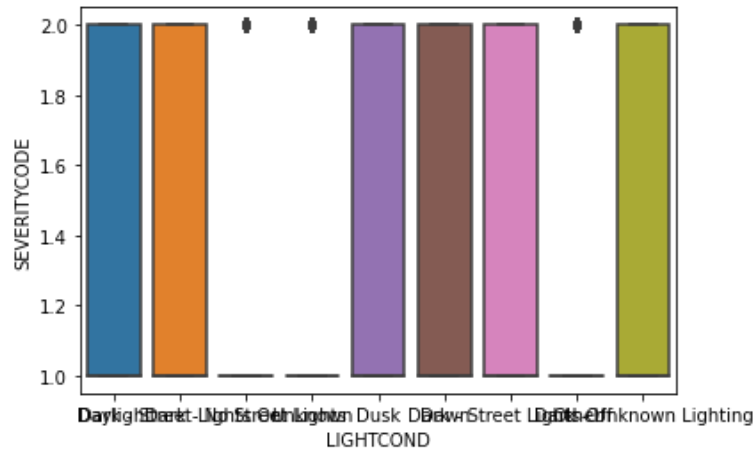
a. Weather condition's impact on severity



b. Road condition's impact on severity



c. Light condition's impact on severity



From the above plots, it seems that a single factor doesn't have much influence on the severity. The accident severity should be based on all factors combined.

Therefore, classification models are used to analyze the data.

- **KNN model**

The severity of an accident could be predicted by the accidents under similar condition. And the KNN model will help to find the closest comparison and generate the predicted severity.

- **Decision Tree**

The severity of an accident could be predicted based on past cases, where a decision tree could help the classification process.

We split the data into testing set and training set. After using the training set to train the model, we use testing set to generate predictions.

IV. Results

We use F1 score and Jaccard index to evaluate the models.

The Jaccard index and the F1 score of the two models were as follows

Model	Knn model	Decision tree
Jaccard	0.69	0.70
F1 score	0.59	0.58

From the evaluation, the two models generate similar accuracy at around 60%. Both models could be utilized to make prediction for the accident severity.

V. Discussion

In this study, it is found that the impact on severity of accidents is complicated, as it is based on all factors. The severity of the accident could be predicted based on the model. For example, based on the Knn model, when it is clear, it is dry on the road and the light condition is daylight, the severity level of an accident is 1.

I. Conclusion

In this study, I analyzed the relationship between accident severity and the weather condition, road condition and light condition. I built classification models to predict whether the accident severity will be affected by weather condition, road condition and light condition. These models can be very useful in helping drivers in a number of ways. For example, it could help identify the severity of an accidents given the weather condition, the road condition and the light condition. And the driver could take care accordingly.

