Car Accident Severity

Introduction

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.

Data of the project

The data is based in Seatle 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.

Methodology

From preliminary analysis, a single factor doesn't have much influence on the severity. The accident severity should be based on all factors combined.

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.

Story for illustration purposes only

Modeling

- 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.

4. Result

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

The Jaccard index and the F1 score of the two models were as above.
From the evaluation, the two models generate similar accuracy at around 60%. Both models could be utilized to make prediction for the accident severity.

Conclusion

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.