

Minimizing Risk Of Car Accidents

Geoffrey Gacuca

Road Safety

- Road Safety is an incredibly important issue that is undervalued in society today
- Understanding where, when and how Road Traffic Accidents (RTAs) happen is a game changer.

Problem

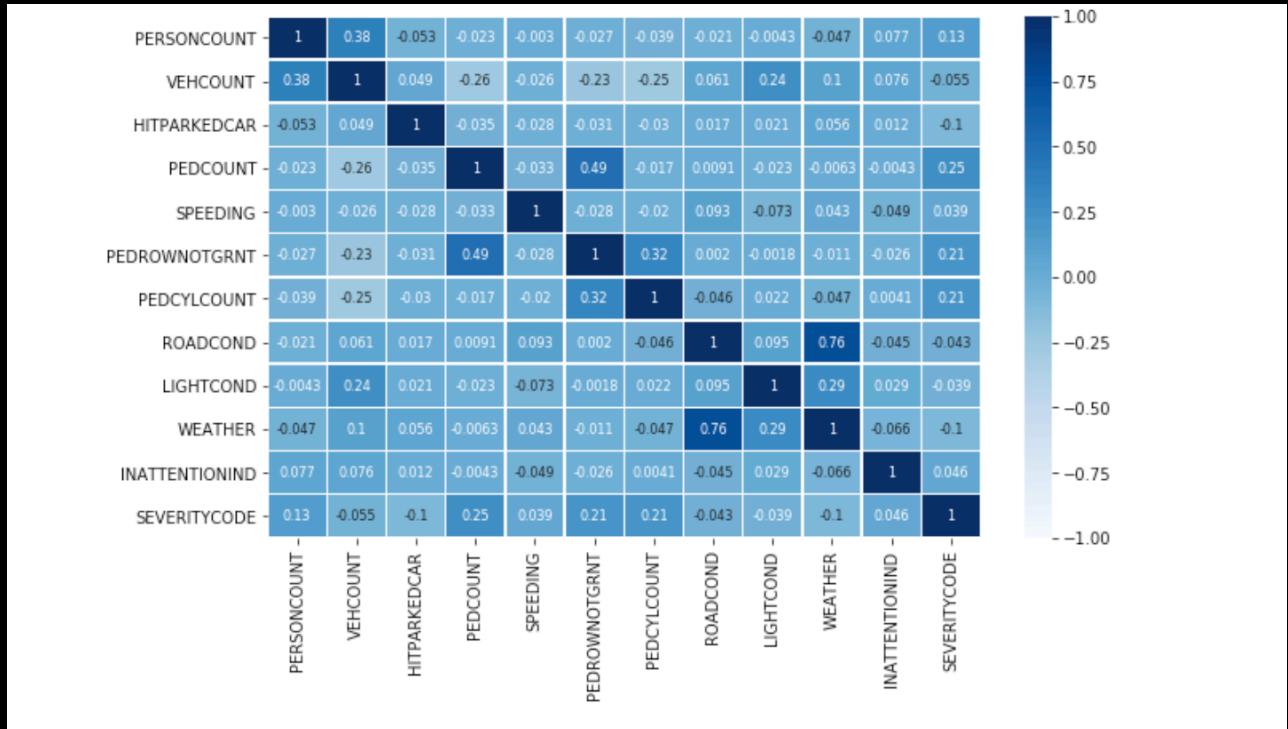
- There are multiple factors that contribute to RTAs and because of the large amount of information on accidents it can be difficult to determine at times which data is important and has a strong relationship to the severity of a car accident and which data is simply extra information

Data Acquisition & Cleaning

- The dataset used in this project comes from the Seattle Department of Transportation and provides RTA data from 1st January 2004 and 20th May 2020.
- The dataset has a total of 194,673 collision incidents with 37 different features such as COLLISIONTYPE , PERSONCOUNT, WEATHER, ROADCOND, LIGHTCOND and more

Exploratory Data Analysis

- This is a visualization of the variables chosen for feature selection and their correlation to one another. Darker color indicates a stronger correlation.



Classification Models Performance

Algorithm	Jaccard Similarity Score	F1-score	Logloss
KNN	0.721279	0.679636	NA
Decision Tree	0.739283	0.662702	NA
Logistic Regression	0.739181	0.663551	0.562489

Conclusion & Future Directions

- Analyzed the relationship between the level of severity in car accidents compared to several features such as weather, light conditions & road conditions
- More information provided on a focus on the age, quality and brand of the vehicle they are driving there might be more improvements to the model in finding how these can affect severity levels of car accidents as different car makers may have different standards on safety.