Predicting Traffic Accident Severity

- By Mohammad Hossein Saeedian

- September 2020

Predicting traffic accident severity is valuable for local government officials

- Seattle, is one of the largest seaport towns in the state of Washington, with over 700,000 residents.
- In 2019, there was a car crash every 4.5 minutes, with several hundred people dying in these accidents per year. *
- Local Seattle government officials are concerned with the overall traffic fatalities in the city. These stakeholders want to understand what the leading factors are that are correlated to these fatal traffic accidents and provide ways to prevent such occurrences.
- Predicting accident severity can lead to:
  - Create different policies related to local traffic laws
  - Adjustments on roadways or traffic flow patterns

Data Collection and Cleaning

Used the “Data Collisions.csv” sample file from IBM Watson that was for Seattle, WA.

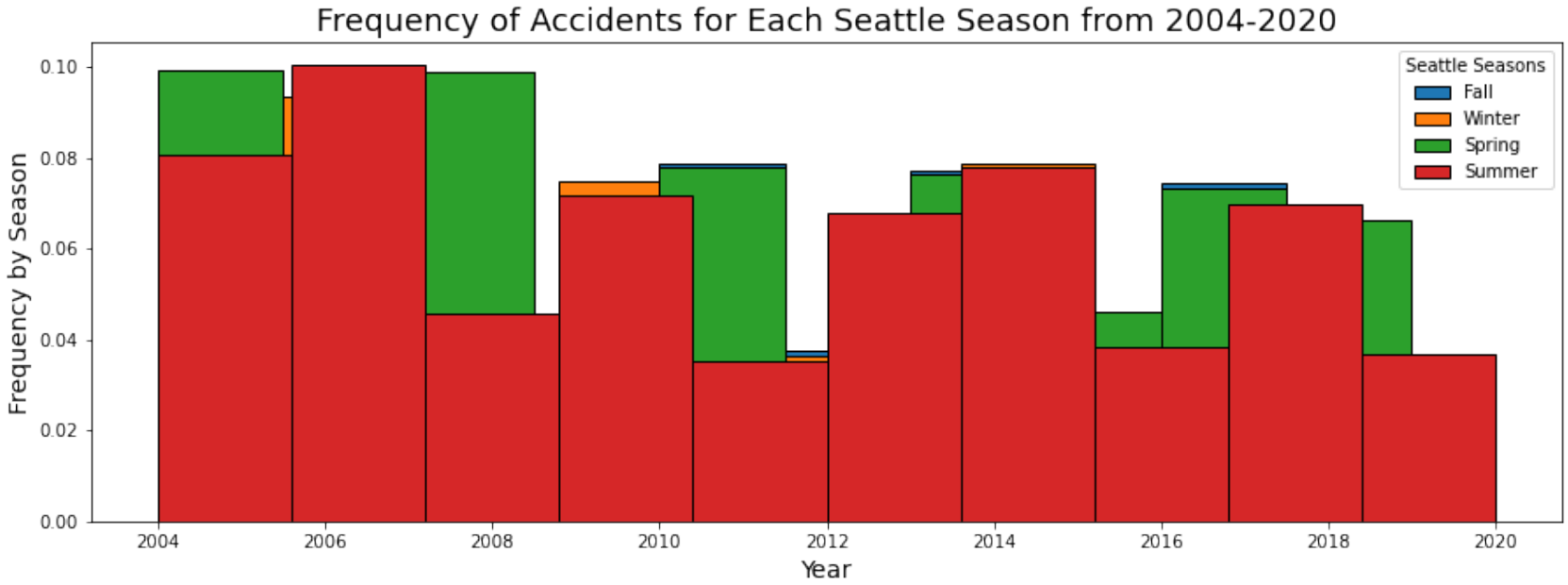
Accident occurrence data ranged from Date range from 01January2004 to 02May2020. In the data set, there are 194673 events

The following attributes in the data set that appeared to have more predictability of the target column “SEVERITYCODE” than others are:

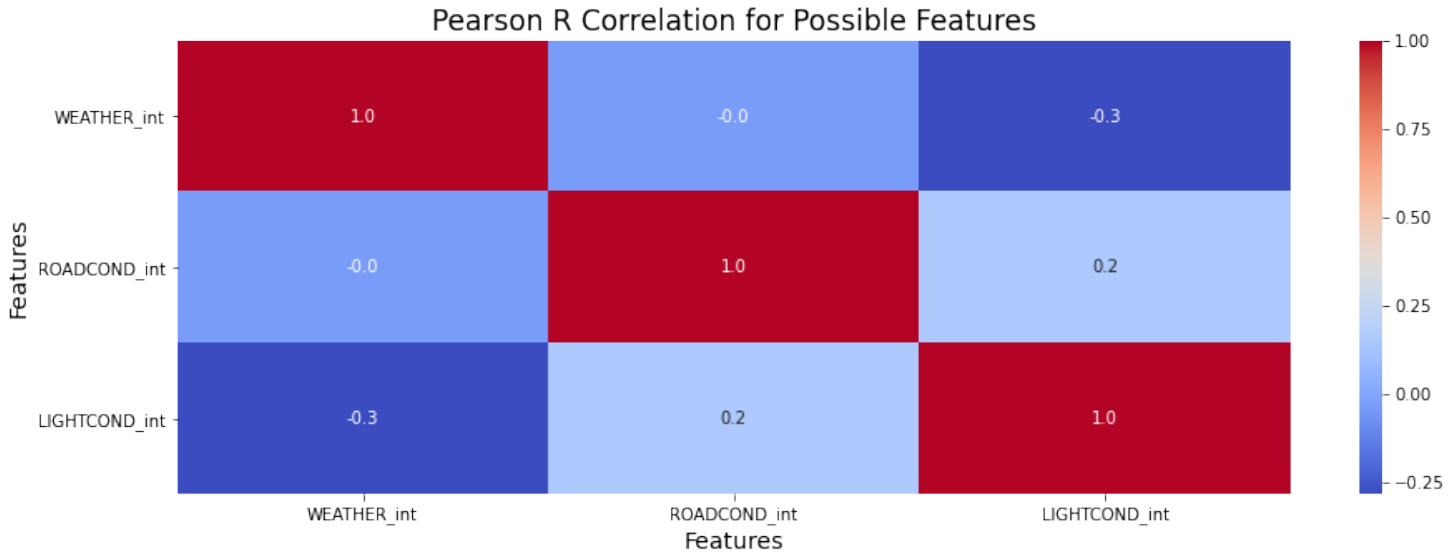
PEDCOUNT
PEDCYLCOUNT
VEHCOUNT
WEATHER

The WEATHER, ROADCOND, and LIGHTCOND were categorical and needed to be changed to integers for machine learning to occur.

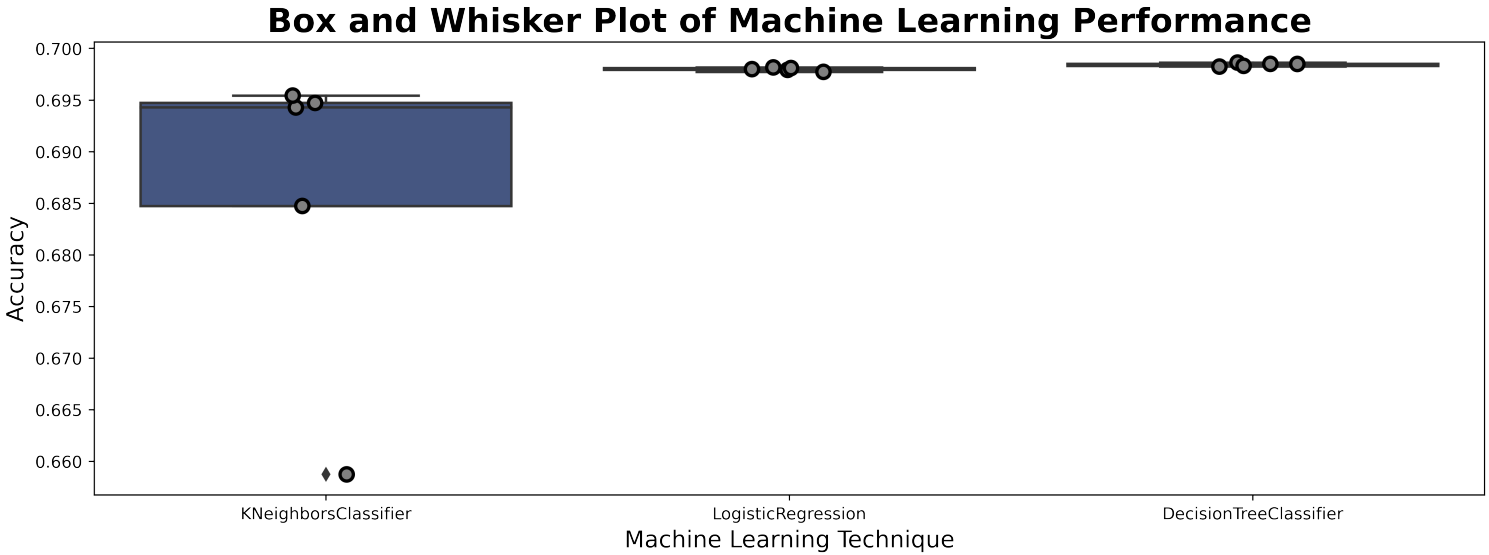
Seasonality of Accidents?



Pearson R Correlations for Factor Attributes



Prediction Plots of Model Accuracy Performance



Conclusion

- Developed useful machine learning models to predict accident severity in the city of Seattle.
- Accuracy and precision have some room for improvement.
- Additional data should be gathered for better use in machine learning classification models.