

Chicago Car Crash

Predicting the severity of a car crash

Business Problem

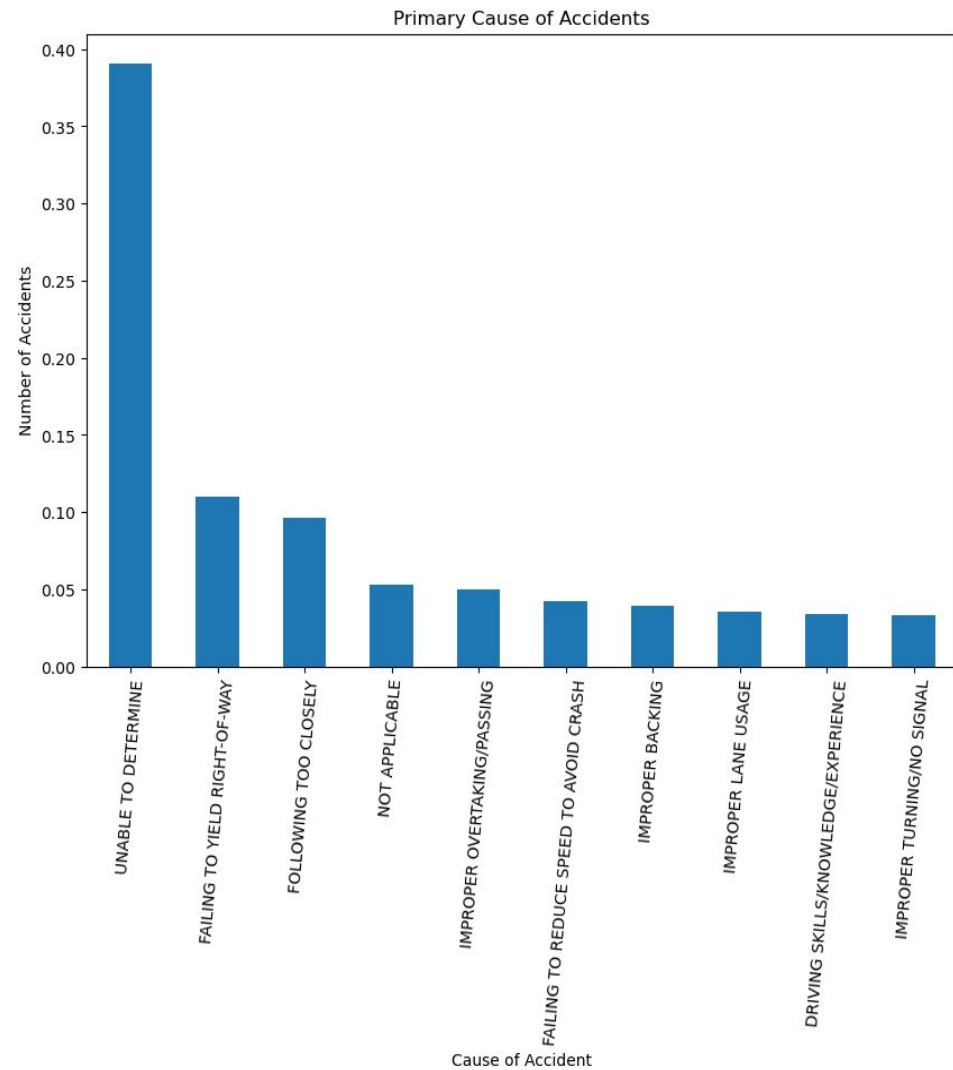
- The City of Chicago faces challenges in identifying the root causes of traffic accidents. This project aims to develop a machine learning model that can predict accident causes based on environmental conditions to inform prevention strategies and improve public safety.

The Data Set

- The data acquired is from Chicago Police Department crash data set link below:
<https://data.cityofchicago.org/Transportation/Traffic-Crashes-Crashes/85ca-t3if>
- This data set has more than half a million records of car crashes
- It has the details of the date, location, time, number of units and cause of accident
- It also provides road conditions, weather, traffic, work zone and the severity

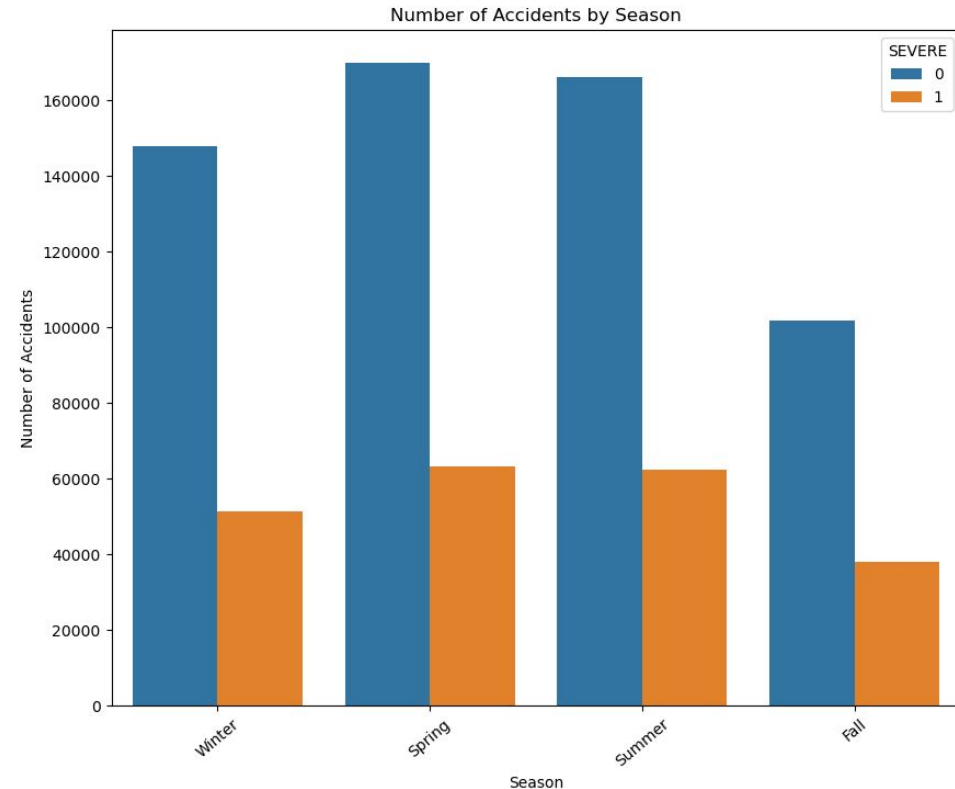
Primary Cause of Accidents

The chart shows the top 10 primary causes of car crashes



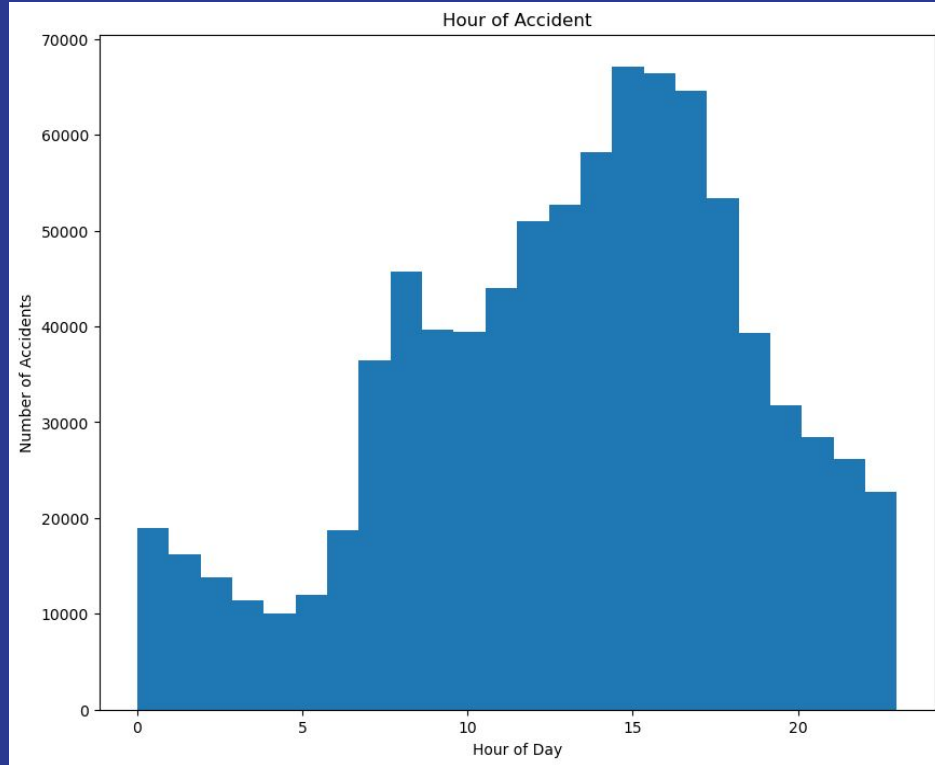
Number of Accidents by Seasons

- The chart shows the number of accidents that occurred in each season.
- The blue bar indicates the non severe car accidents.
- The orange bar indicates the severe accidents



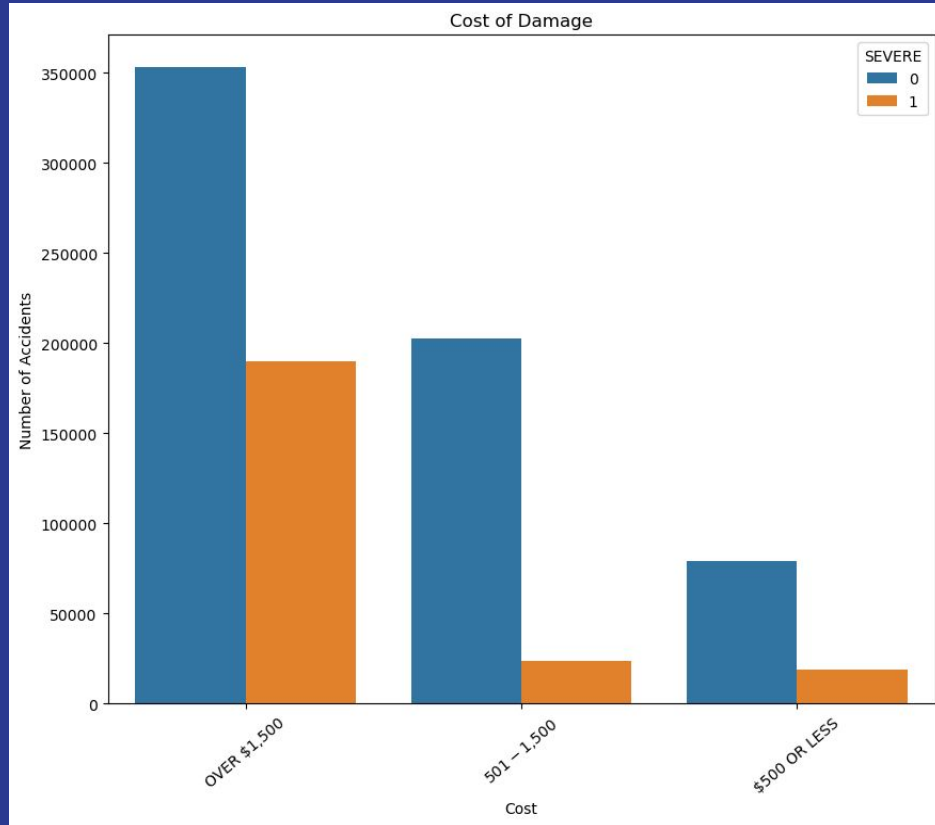
Accidents per Hour

- The most accidents occurred in morning and evening rush hours



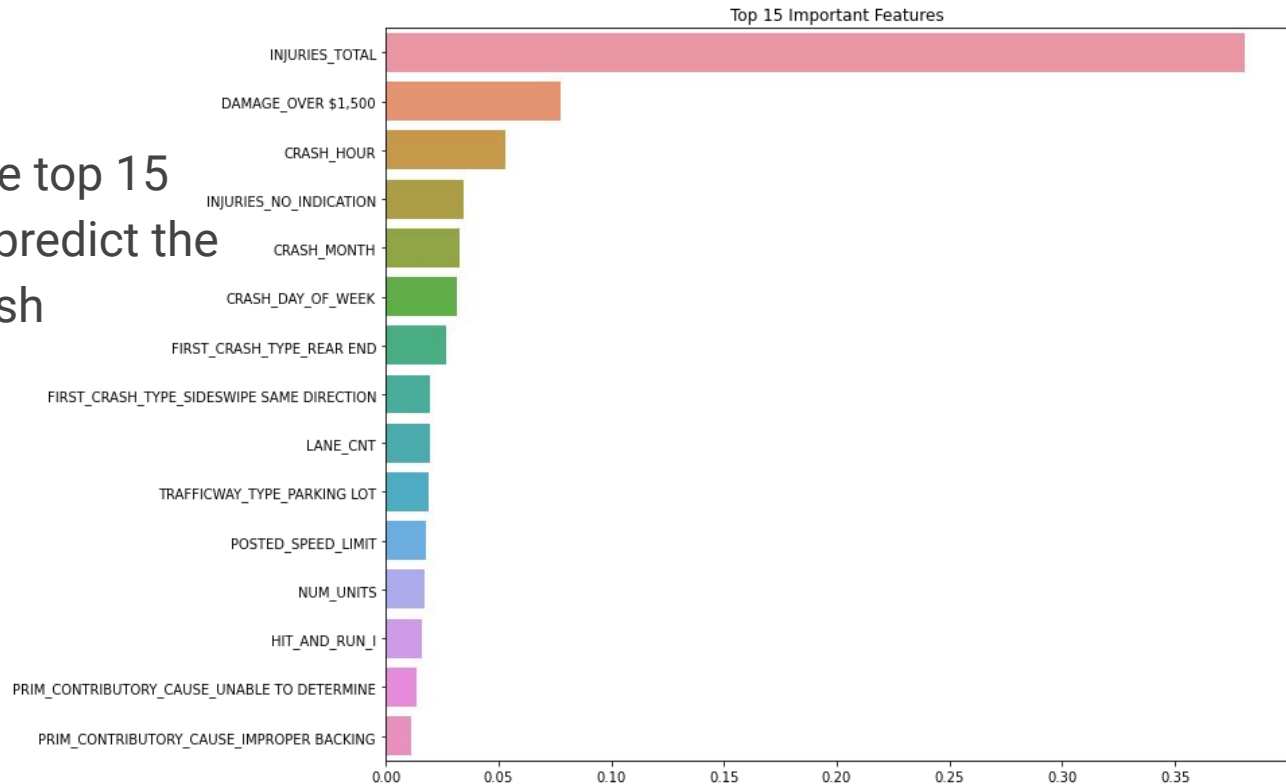
Cost of Damage

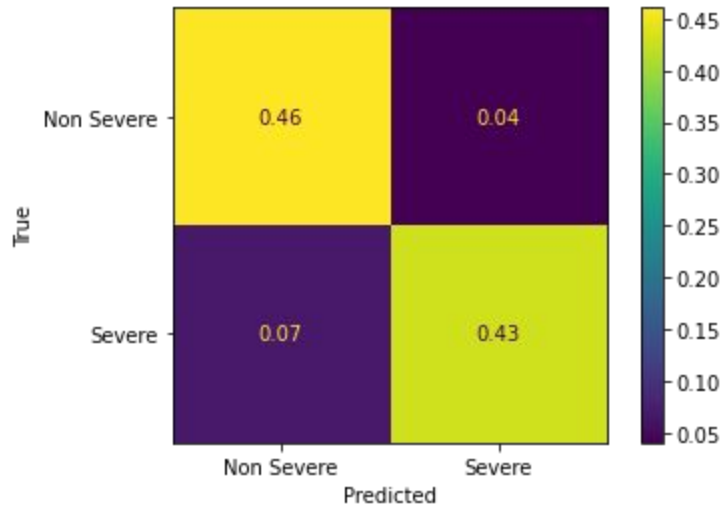
- Most severe and non severe accidents damage cost is more than \$1500



Important Features

- This chart shows the top 15 features which can predict the severity of a car crash





Model	Accuracy
Logistic Regression	0.9
Decision Tree	0.85
Random Forest	0.89

Final Model

- Logistics regression model Accuracy - 90%
- Confusion matrix of the Logistics Regression model

Recommendations

- Improving road conditions:
Addressing and improving conditions in areas where poor roadway surfaces or defects have been identified.
 - Accident prevention campaigns:
Educating the public on safe driving behaviors, particularly around yielding to pedestrians and complying with traffic control devices.
-

Recommendations

- Improved law enforcement: Increased enforcement of traffic laws and road safety measures, particularly in high-risk areas.
 - Advanced traffic management systems: Installing sensors and signals to alert drivers of upcoming traffic jams, especially during peak hours.
-

Conclusion

- Final models show that certain columns have more of an effect on the severity of a car accident than others.
- Accidents involving pedestrians are more likely to result in severe injury.
- Accidents where there is just a side swipe leads to non severe.

Future Work

- Incorporate more features to improve the model's predictive power.
 - Use additional data sources to validate model accuracy.
 - Develop a live dashboard for monitoring and predicting crash occurrences.
 - Collaborate with city planners to use the model's results for real-world applications.
-

Thank You



[Linkedin](#)



[Website](#)



Samwel Waithaka

