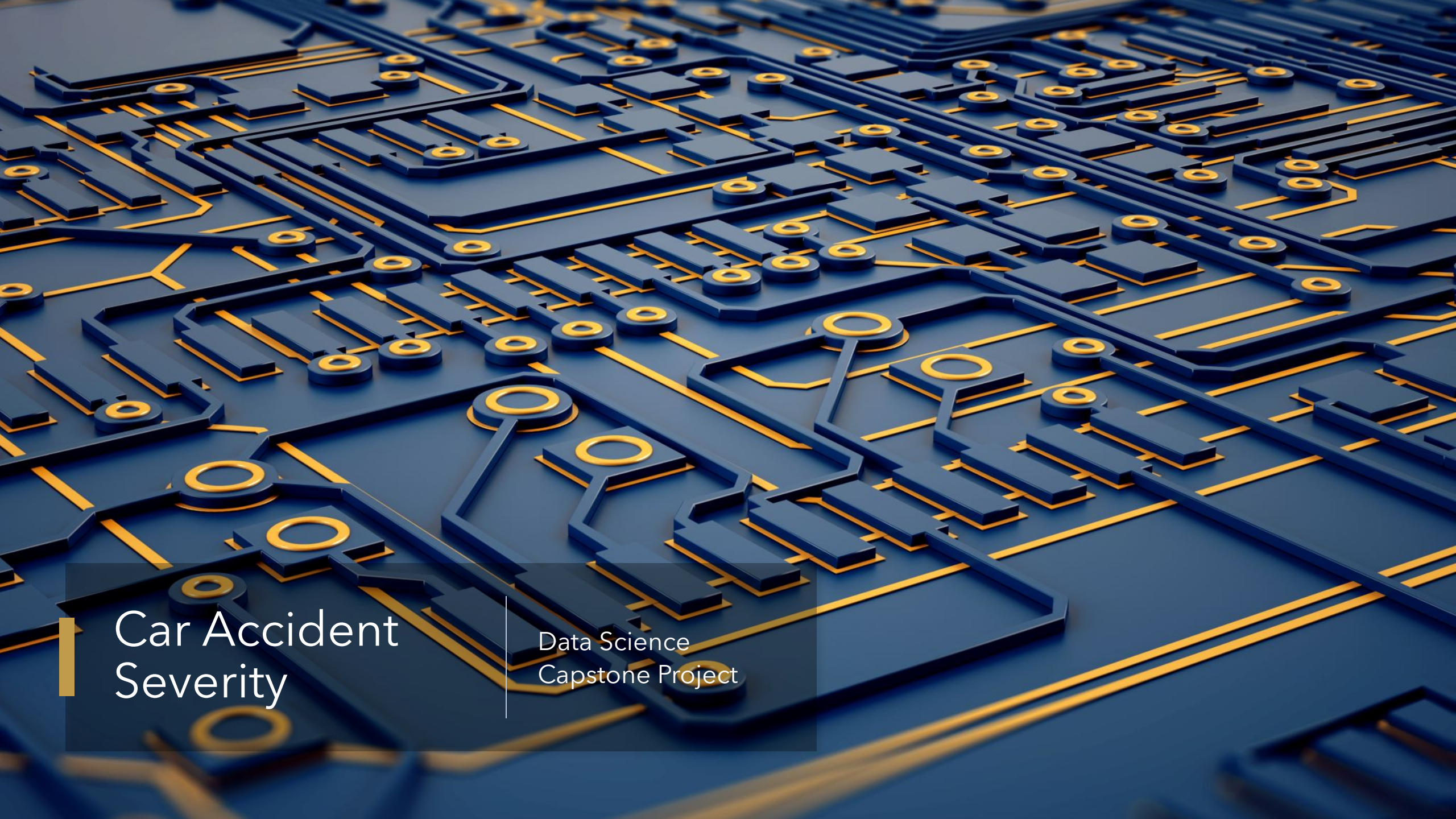


Car Accident Severity

Data Science
Capstone Project

Introduction

- Car accidents can be caused due to various reasons. Some of which include; road conditions, weather and visibility conditions. The extent of these factors also affects the severity of the accidents.

- It is the aim of this project to determine and predict what the severity of an accident can be based on some conditions.

Data

- The data used here has been provided by the Seattle Police Department and recorded by Traffic Records. This data includes all types of collisions. Collisions will display at the intersection or mid-block of a segment. The time frame of the data is from 2004 to present day. The data had been updated on a weekly basis.

- The “SEVERITYCODE” column has some numbers and represent road accident severity in the following way:

  3 – fatality

  2b – serious injury

  2 – injury

  1 – prop damage

  0 – unknown

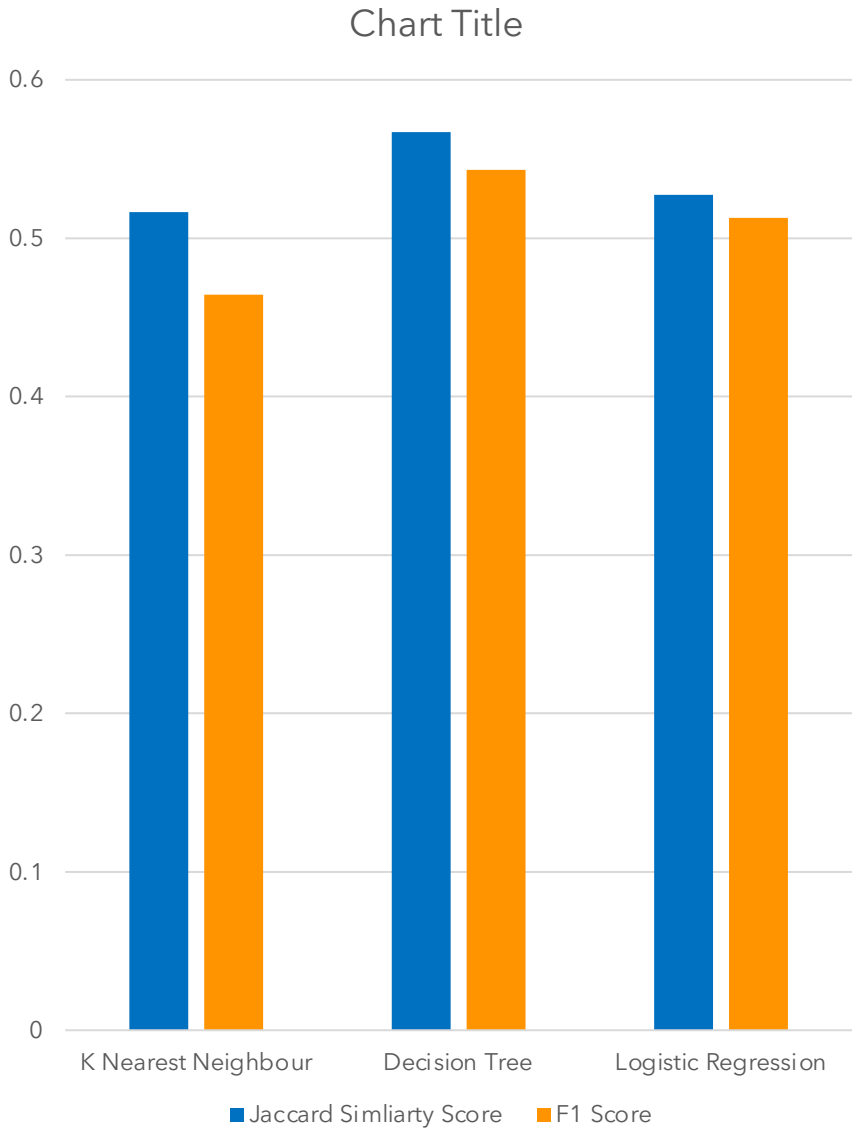
Pre-Processing the Data

- Looking at the metadata, it had been inferred that three columns could be considered as the independent variables and these were; WEATHER, ROADCOND, and LIGHTCOND.

- The data types of the three variables were categorical and had to be converted to their respective categorical codes (in int64).

Training and Testing

- The model was trained using three differenet machine learning techniques. There respective errors are shown.
- It can be infered that the Decision Tree model outputs the most accurate model



Conclusion

- The Decision Tree model outputs the most accurate model and hence is used to model the dataset.

- This project has been done to reduce the overall severity of accidents by alerting drivers of the dangers in different road conditions.