Car Accident Severity

Applied Data Science Capstone

Introduction

- **description of problem, discussion of background**
- **road accidents are common leading to losses**
- **need to understand most common causes**
- **algorithm has to be designed**

- **WHY?**

  predict the severity of a possible accident

Business Problem

- **Conditions:**
  - current weather
  - road condition
  - visibility condition
- **understanding of the factors and their correlations**
- **Beneficial**
  - alert drivers to drive safe
  - inform police to enforce protocols

Data

- **dataset file Data-Collisions.csv**
  - size : ~ 1,90,000 collisions
  - collisions at Seattle
  - from 2004 to present, updated weekly
- **37 attributes**
  - ADDRTYPE
  - SEVERITYCODE
  - WEATHER, ROADCOND, LIGHTCOND

Methodology

- **some features do not have numerical type**
- **label encoding**
- **normalizing the dataset**
- **splitting dataset:**
  - 70% training
  - 30% testing

Modeling and Prediction

- **Modeling**
    1. k-Nearest Neighbours
    2. Decision Tree
    3. Logistic Regression
- **Prediction**
    1. Jaccard Index
    2. F-1 Score
    3. Logloss

Conclusion

- There is a correlation between the conditions studied and accident rate.
- IoT devices can be integrated with an APP for notifying drivers and also cops.