Capstone Project

Accident Severity Prediction

Introduction

- A model to Predict Accident Severity

  Road accidents are extremely common, and they often lead to loss of property and even life.
  Hence it is good to have a tool that can alert the drivers to be more careful depending on the weather and road conditions.

- Target Users : Drivers, Police Department and Emergency Services

  If the severity is high the driver can decide whether to be extra cautious or delay the trip if possible.
  Emergency Services and Police Department can then plan ahead and implement extra safety protocols to prevent future accidents.

Data acquisition and cleaning

- Data Set - the collision data recorded by the Seattle Department of Transportation(SDOT) - https://s3.us.cloud-object-storage.appdomain.cloud/cf-courses-data/CognitiveClass/DP0701EN/version-2/Data-Collisions.csv

- Raw Data had 194673 rows and 38 columns

- After removing records with missing values and removing irrelevant data
  166510 rows and 12 features

Data Description

- Target Variable

  SEVERITY CODE

  1 - Property Damage only Collision

  2 – Injury Collision

- Features

  WEATHER – Clear, Overcast, Raining, Snowing

  ROAD – Dry,Ice,Snow/Slush,Wet

  LIGHT – Dark-Street Lights On, Dawn, Daylight, Dusk

Models and Results

Model	Accuracy	Jaccard Score	F1 Score	Log Loss
KNN	0.67	0.67	0.54	
Decision Tree	0.67	0.67	0.54	
Logistic Regression	0.67	0.67	0.54	0.65

The best K value for K Nearest Neighbour Model - 8

Conclusion

- Built models to predict the severity of an accident that may occur given certain condition
- Performance is consistent for all 3 models
- Recommend to use Logistic Regression because of speed and ease of implementation
- Accuracy has room for improvement