Washington D.C. Traffic Accident Dashboard

Sam Clark and Joanne Choi

8/18/2022
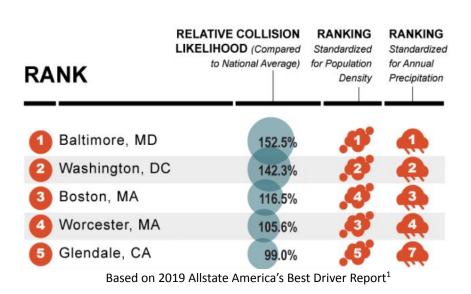
Overview

Problem

- Washington DC has the 2nd highest vehicle collision rate in the U.S.

- Compared to the average commute time of 27 minutes, DC commuters spend an average of 43 minutes

Solution

- We developed an interactive dashboard that contains machine learning components that can help inform D.C. commuters on the risk of commuting on a given day

RANK		RELATIVE COLLISION LIKELIHOOD (Compared to National Average)	RANKING Standardized for Population Density	RANKING Standardized for Annual Precipitation
1	Baltimore, MD	152.5%	1	1
2	Washington, DC	142.3%	2	2
3	Boston, MA	116.5%	4	3
4	Worcester, MA	105.6%	3	4
5	Glendale, CA	99.0%	5	7

Based on 2019 Allstate America's Best Driver Report[1]

Machine Learning Models

- Regression model predicting the number of car accidents likely to occur
  - ~82% accuracy

- Classification model predicting the level of severity for a given accident
  - ~97% accuracy
  - ~95% recall for each class
  - ~97% precision for each class

- XGBoost performed the best for both regression and classification models

Dashboard Demo

Future Work and Improvements

- Regression Model:
    - Investigate model stacking to improve performance
    - Add more data including traffic volume

- Classification Model:
    - Concerns for overfitting
    - "No Injuries" class missing from data
    - Oversampling method (SMOTE)

- Include additional accident data from MD and VA

- Continuously integrating accident data to keep the models up to date

Takeaways

- Most number of car accidents occur in:
  - May
  - Wednesday
  - 5AM

- Car accidents are highly correlated with precipitation

Questions?