Terry Stop Arrests

A Binary Classification Model

Carlos Garza

Business Problem

- Terry vs. Ohio set the legality of random “stop and frisk” on the grounds of reasonable suspicion
- Given subject and officer data, can we predict what stops will result in an actual arrest?
- What are the factors that most highly indicate a stop ending in an arrest?

Data Utilized

- Data originates from Seattle Police Records available at data.seattle.gov
- Dependent Variable: “Arrested”

Independent Variables

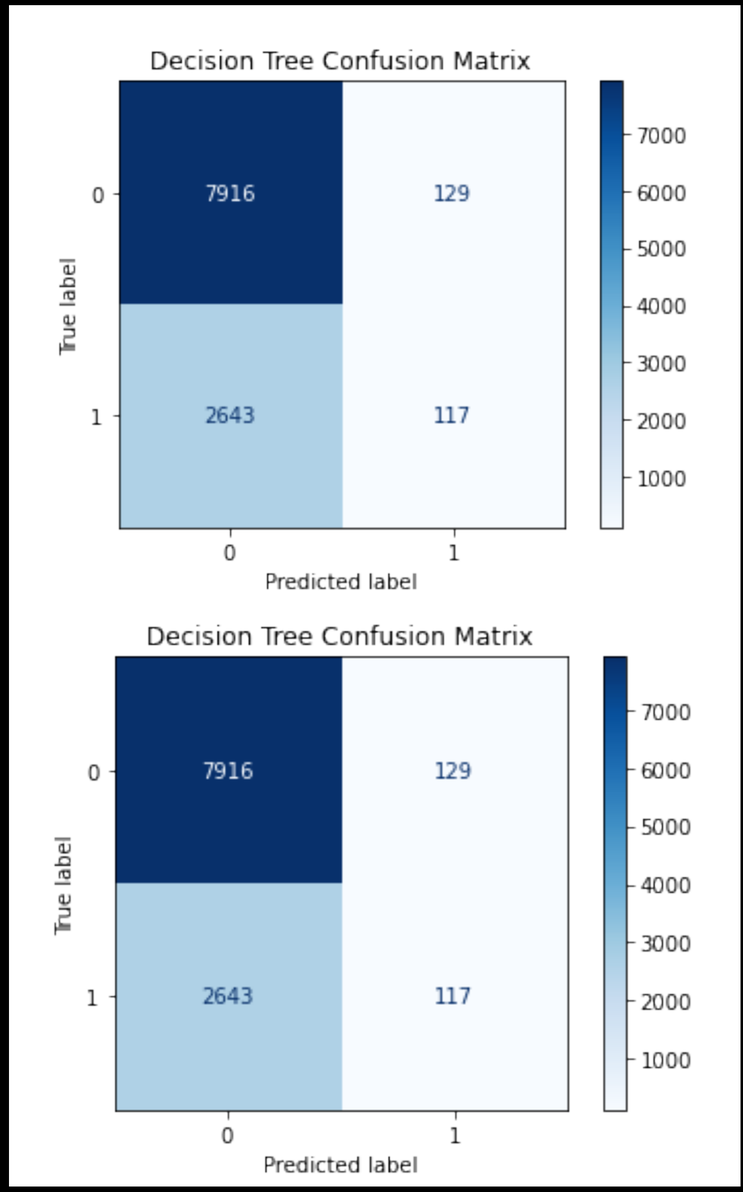
Subject Age Group	Weapon Type
Officer YOB	Officer Race
Subject Perceived Race	Subject Perceived Gender
Reported Date	Reported Month
Frisk Flag	

Baseline Model

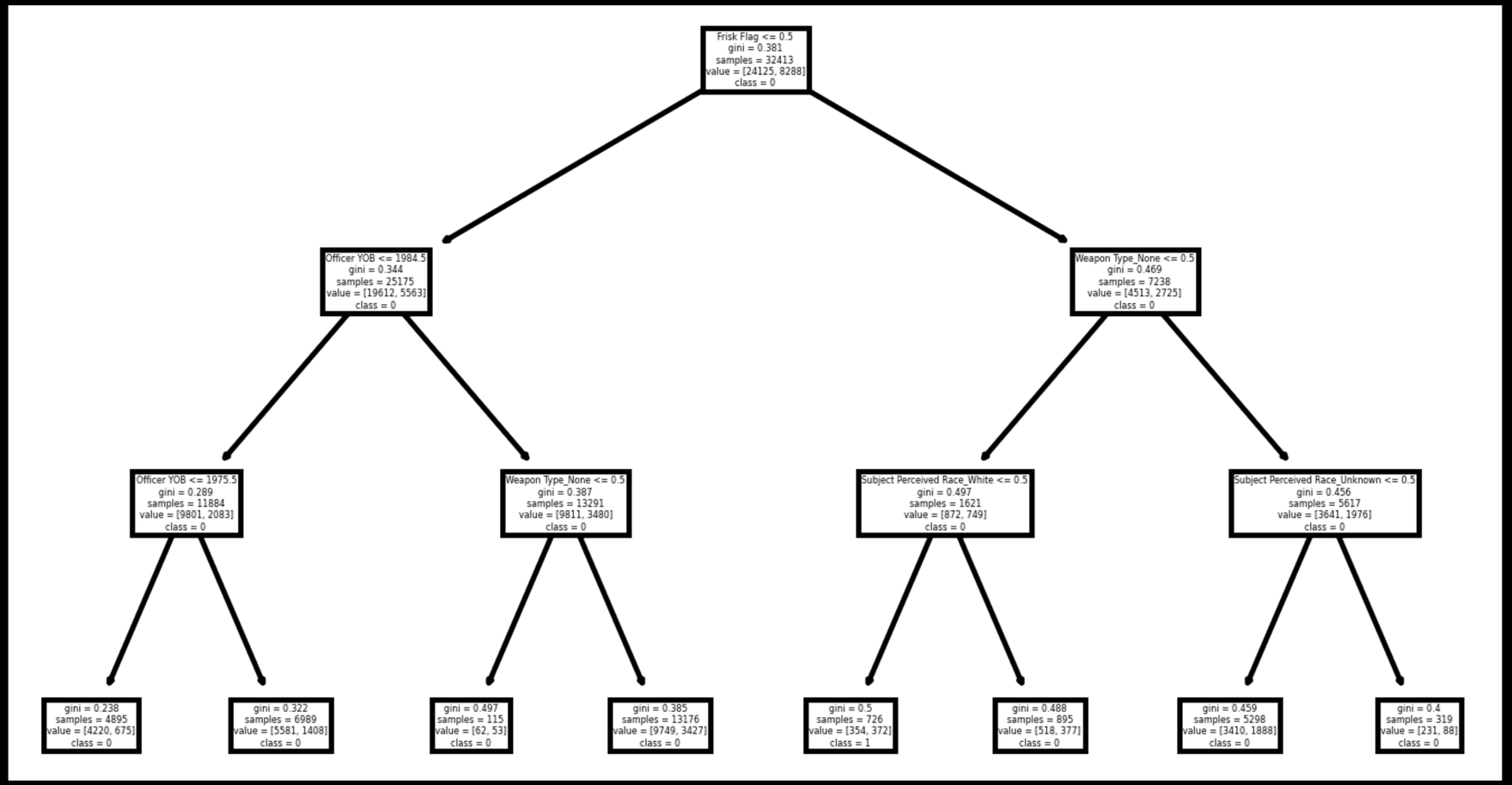
- K Nearest Neighbor
  - Accuracy: 0.707
- Decision Tree
  - Accuracy: 0.743
- Random Forest
  - Accuracy: 0.745

Tuning/Model Selection

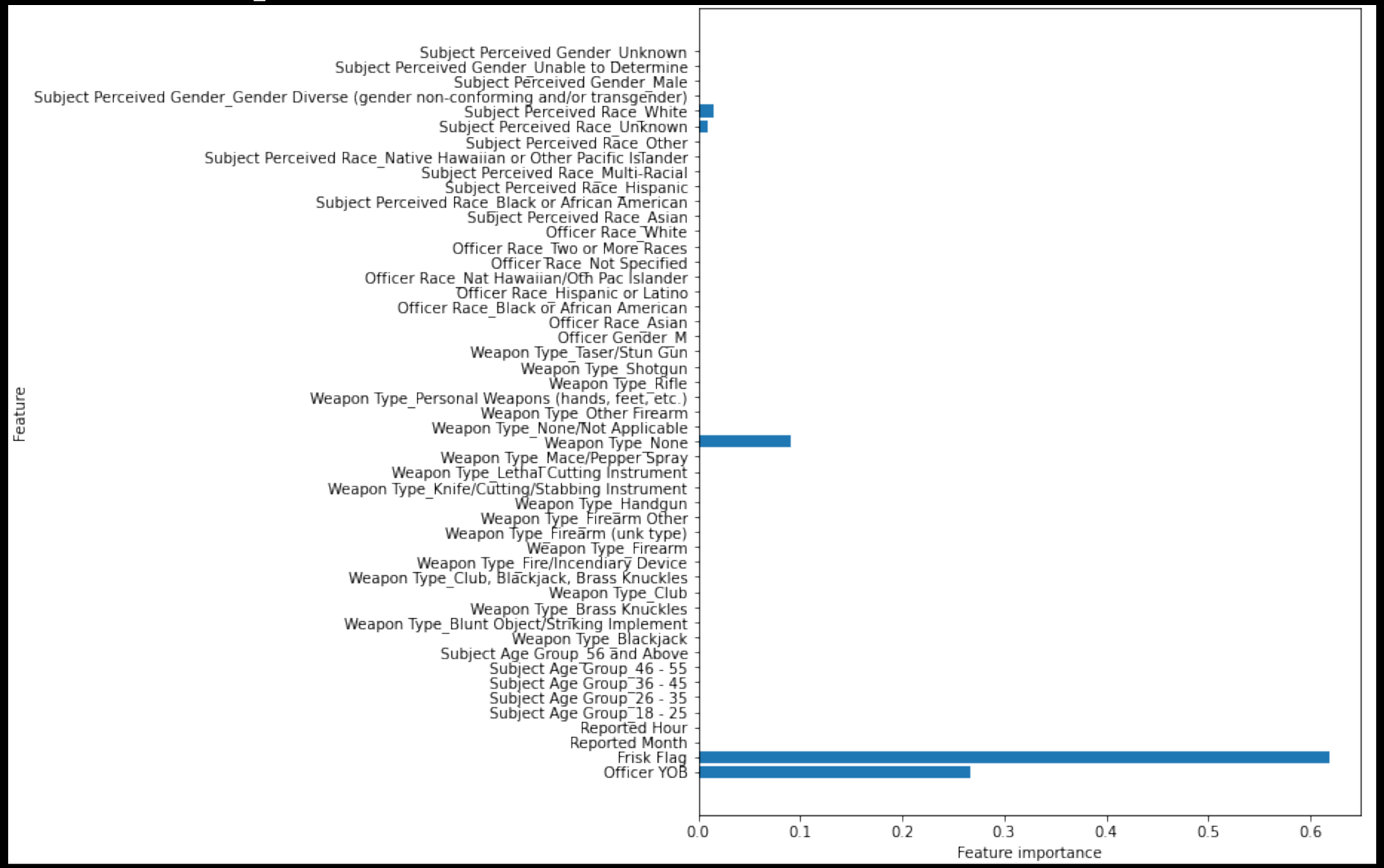
- Grid Search performed on decision tree and random forest to improve model performance
- Performances were infinitesimally close, so the decision tree was selected as the final model because of its lighter computational load



Final Decision Tree



Feature Importance



Conclusions

- Most important variable predicting arrest during a stop was if the subject was frisked
- Other important variables include the age of the officer and whether the suspect was carrying a weapon
- Type of weapon does not matter

Future Work

- Explore ways to decrease false positives in our model
- Incorporate data from a wider geographic area to create a more comprehensive model
- Explore relationship between officer age and arrest rate

Thank You

Carlos Garza