Seattle Terry Stops Analysis

Spencer Hadel

Summary

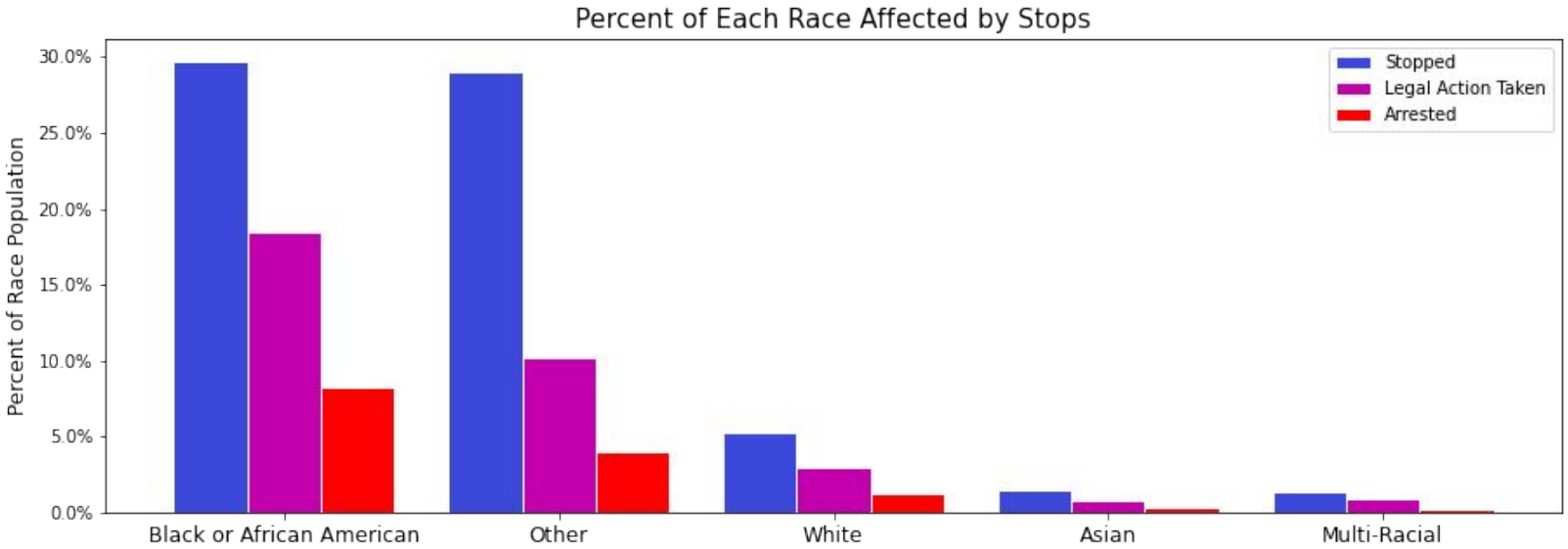
- The new mayor of Seattle has promise the public that they will investigate racial disparity in the Police Department.
- This analysis seeks to identify key factors of this disparity in Terry Stops, a.k.a. Stop-and-Frisks.

Outline

- Business Problem
- Data
- Methods
- Results
- Conclusions

The Problem

This analysis will investigate the recorded Terry Stop data of Seattle.gov in an attempt to identify racial disparity in the outcomes of theses stops.



Data

- This analysis utilizes Terry_Stops.csv, a dataset of over 52,000 entries of Terry Stops in Seattle from 2015 to the present (2022).
- Data was cleaned to increase model performance.
- Notable and important features include Subject Perceived Race, Officer Race, Arrest Flags, Frisk Flags, and Stop Resolution.

Seattle

Seattle Open Data

Methods

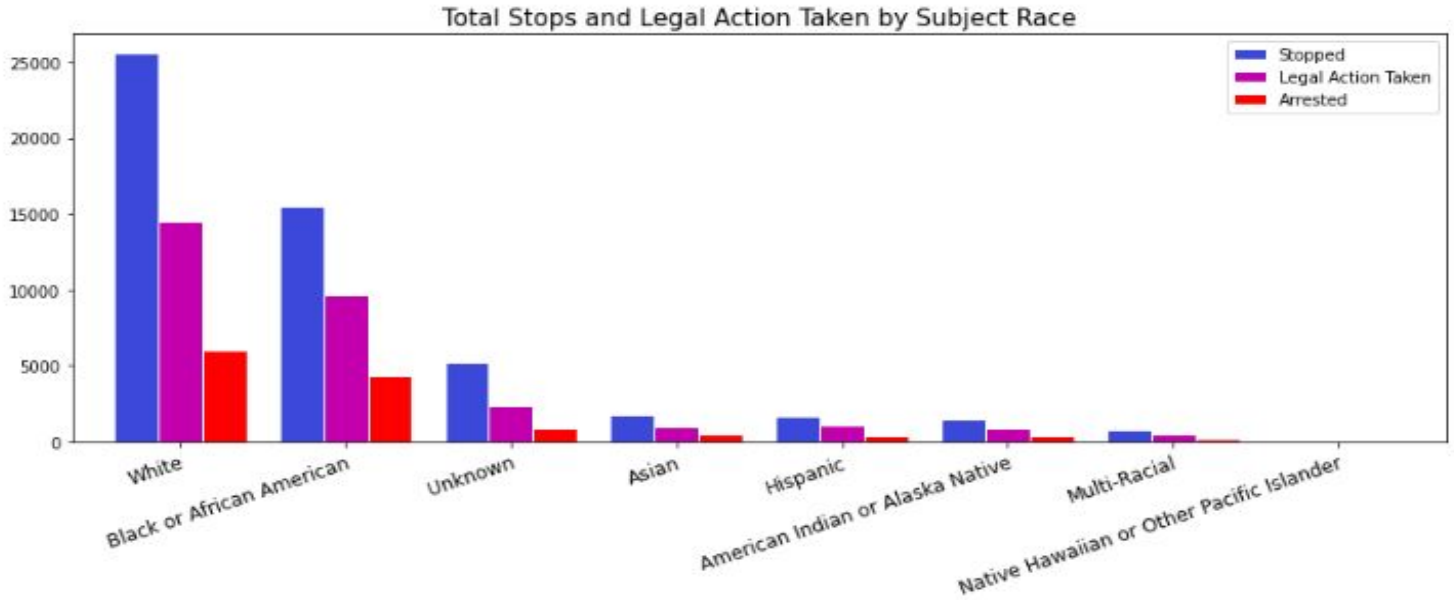
- Iterative approach to model selection and tuning
- Measured classification models' **Performance Scores**
- Different models for different **Target Variables**

Pre-Modeling:

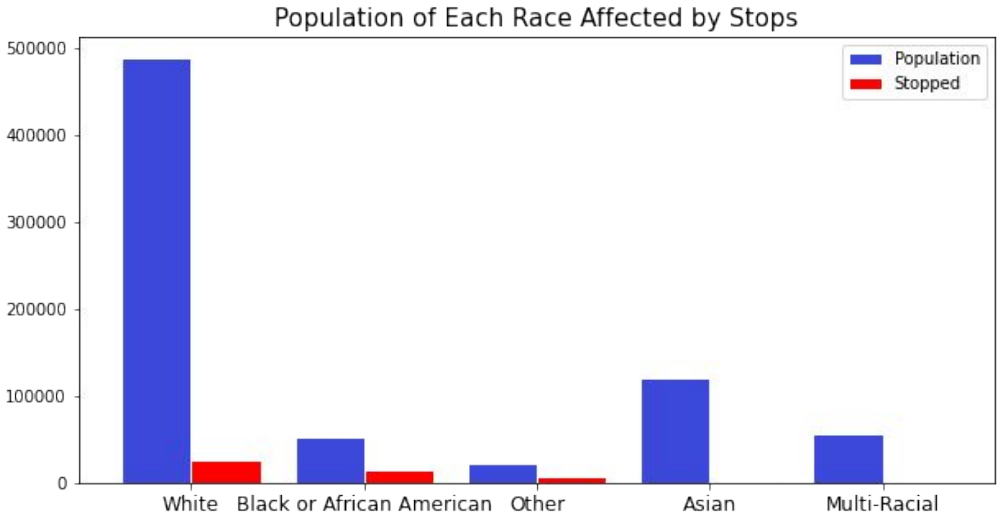
Data Visualization

Results - Data Visualization

- Initial data analysis showed some basic findings about arrest data at Traffic Stops
- This shows total population stopped and the outcome of those stops.

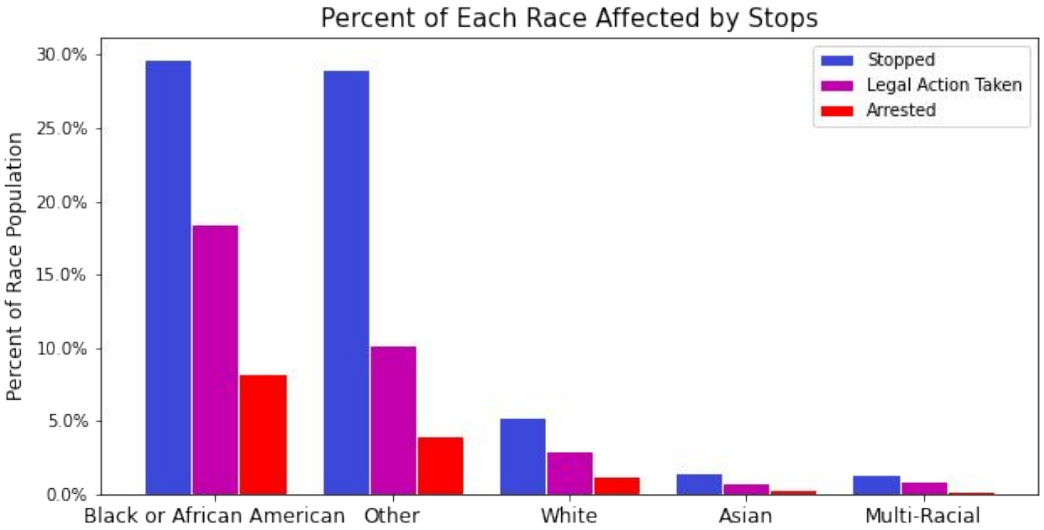


Results - Data Visualization



- The prior results seem relatively unbiased until we learn that Seattle's population is **65% White**, showing that our Terry Stop Data has been drastically affected by racial bias.

Results - Data Visualization



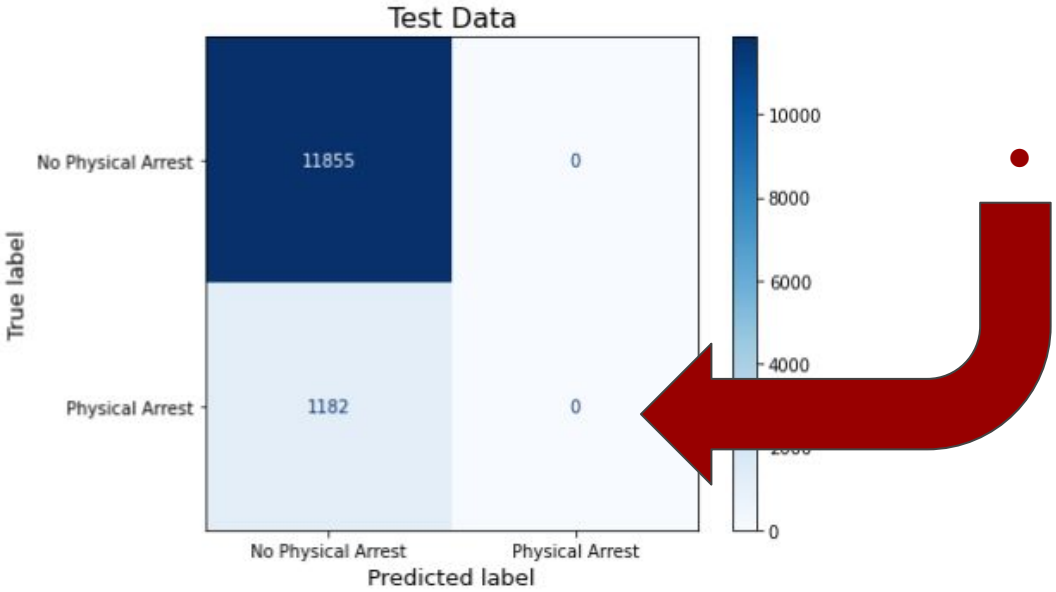
- For example, if Black or African American subjects only make up **7.1%** of the population, why do they make up **nearly 31%** of the Terry Stop Data?

Data Modeling

Target: ‘Physical Arrest’

Results - Baseline Model

Baseline Model: Logistic Regression

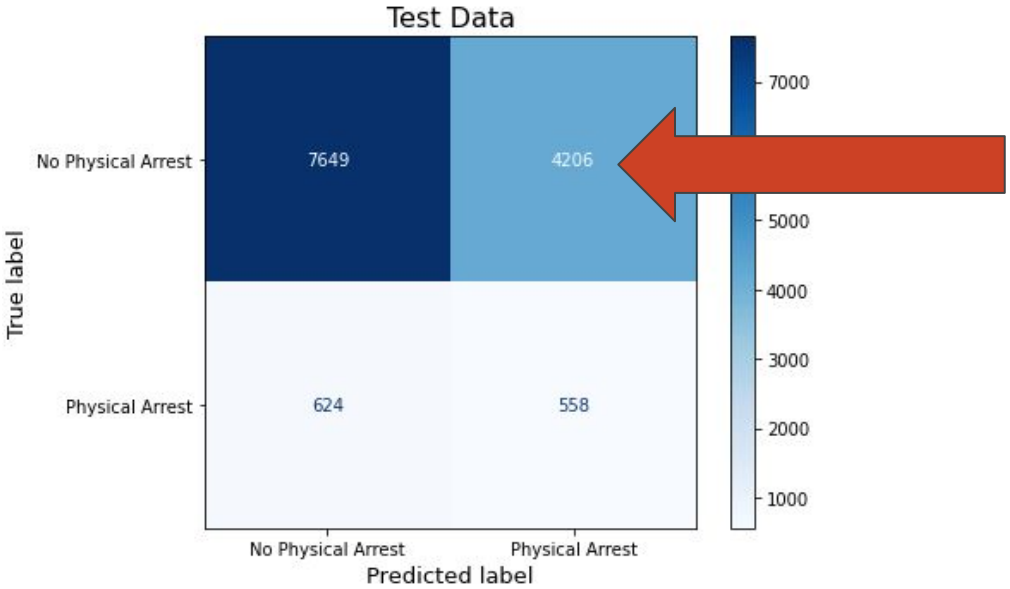


Score: 0.0

- Initial model cannot successfully predict any instance of ‘Physical Arrest’

Results - Baseline Model with Synthetic Data

Logistic Regression (Synthetic Data)

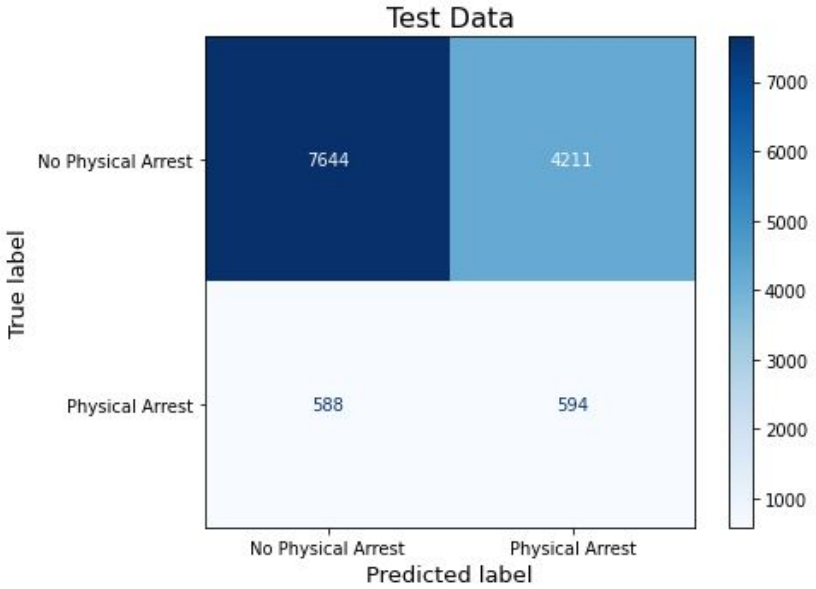


Score: 0.187

- False Positives:
  Model predicts Physical Arrest when test data does not indicate one

Results - Best Model for Target: ‘Physical Arrest’

XGBoost (with Synthetic Data)

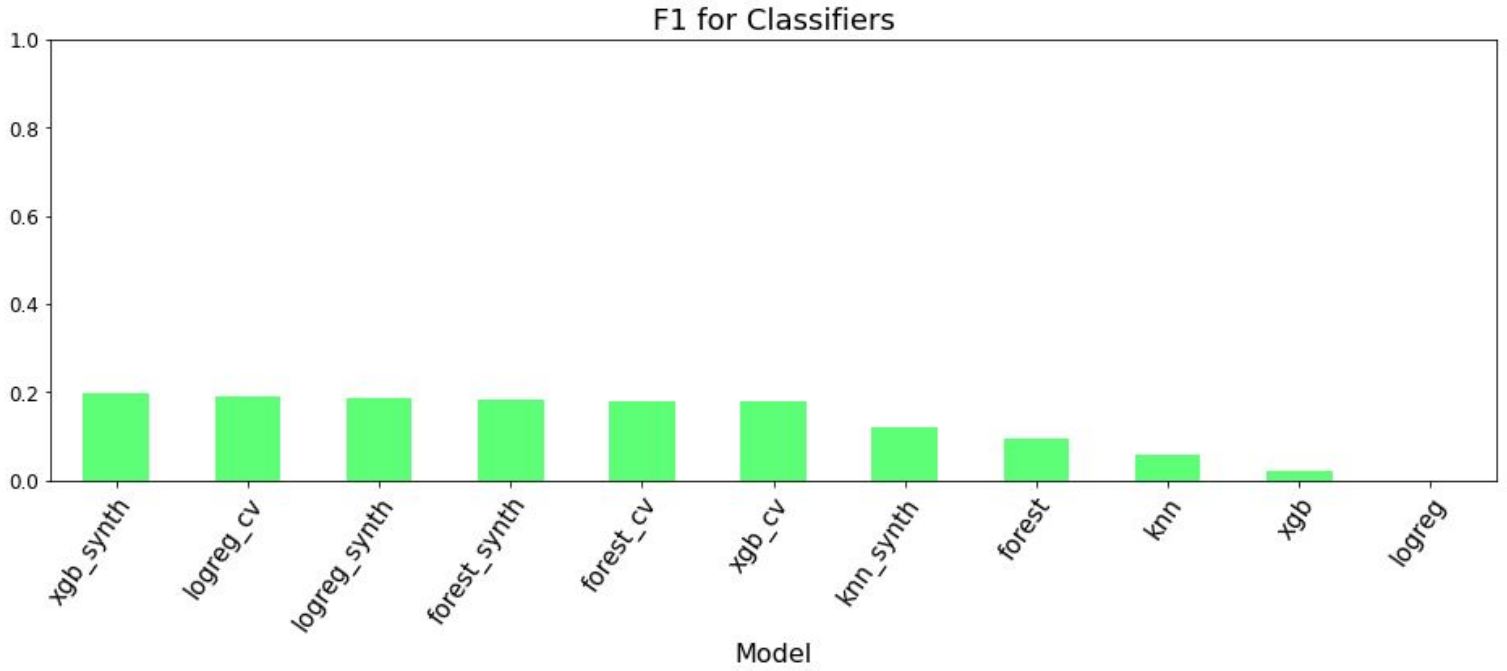


Score: 0.198

- Almost identical model performance with only very small improvement in F1 Score

Results - Model Comparison

(Ideal F1 score should approach 1.0)

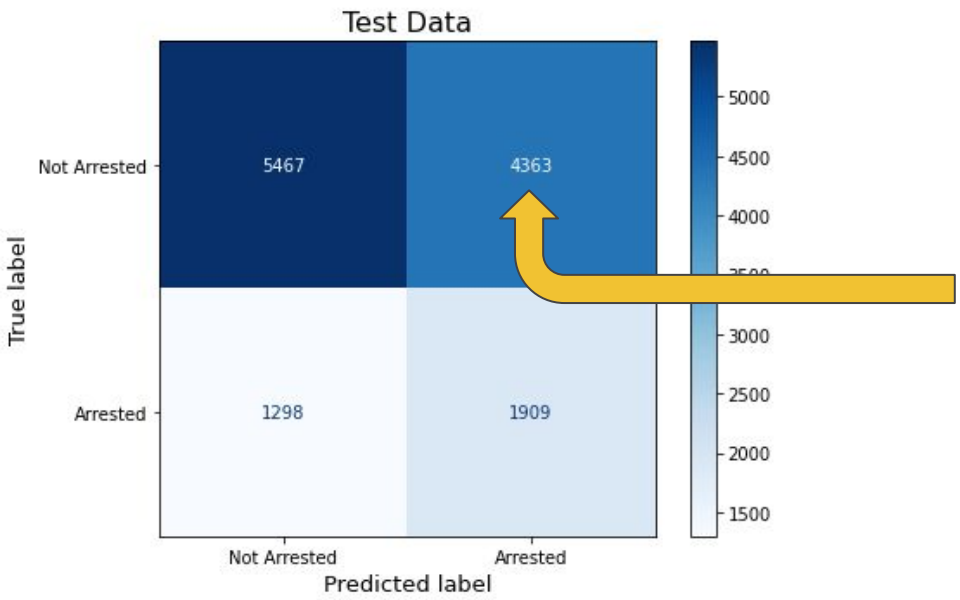


New Targets:

- ‘Arrested’
- ‘Legal Action Taken’

Results - New Targets

Best Model for Target: ‘Arrested’

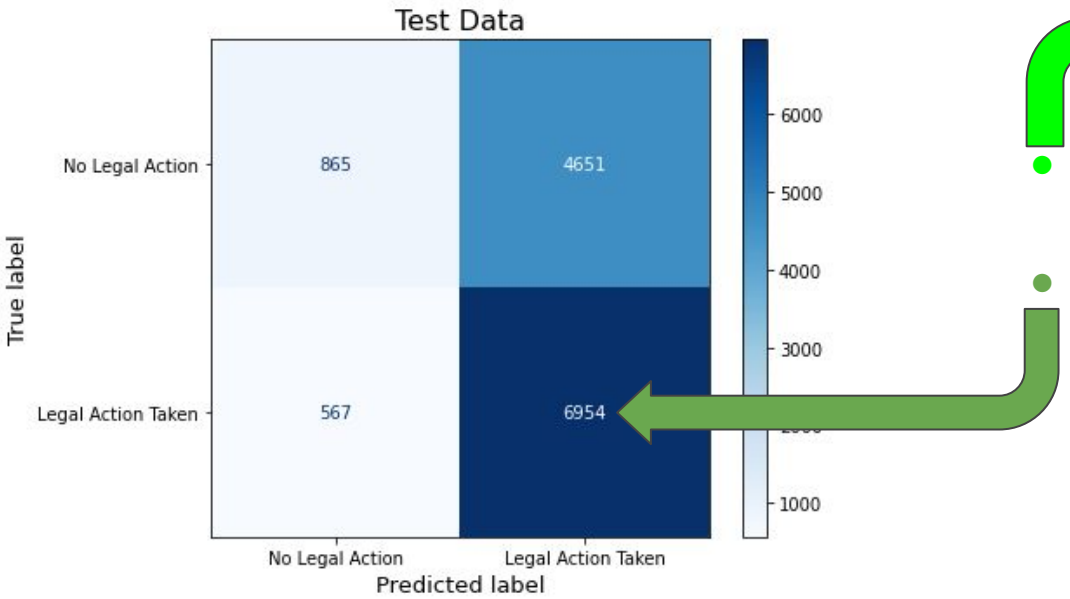


Score: 0.403

- Better Performance
- Still predicts large amount of False Positives

Results - New Targets

Best Model for Target: ‘Legal Action Taken’

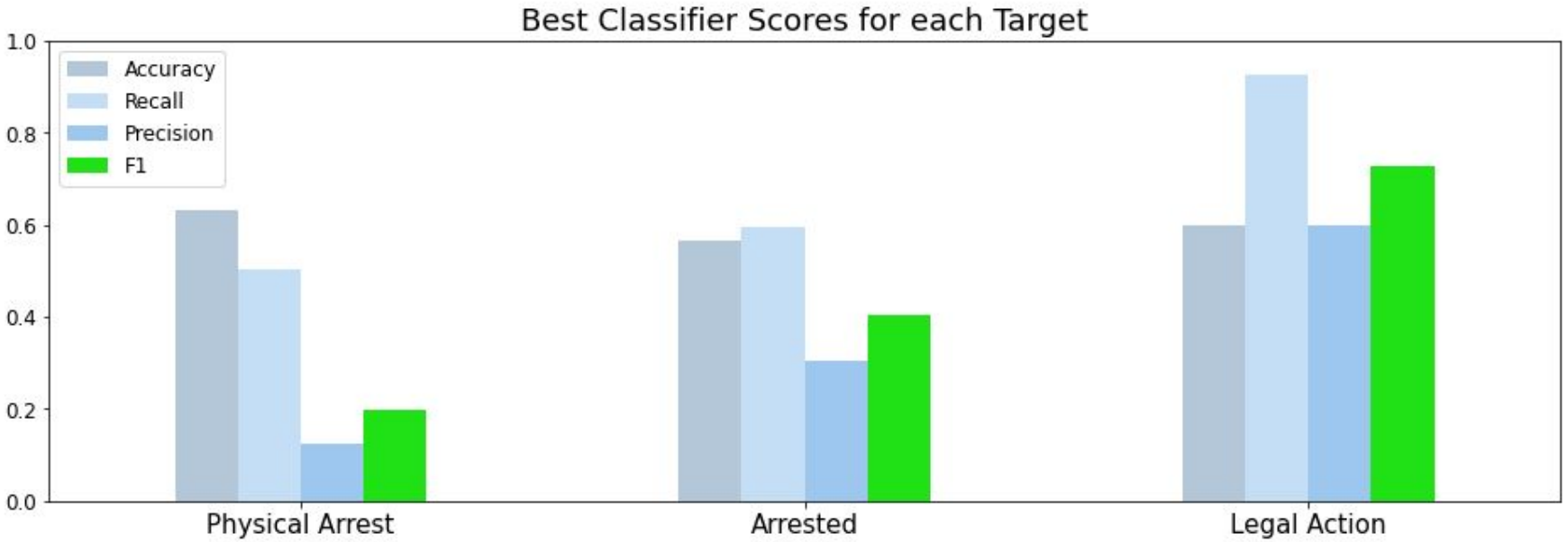


Score: 0.727

- Significant increase in performance
- Predicts True Positives More than False Positives

Results - Final Model Comparison

Significant Performance Boost When Choosing New Targets



Final Observations

- Data needs to be investigated and manipulated differently to achieve higher performance, and draw more actionable insights.

- Different techniques may need to be used, such as multi-target classifiers or unsupervised learning models.

- Continued and more varied analysis is required.

Next Steps

- Political Action
  - Standardize Police Force Record Keeping Techniques
  - Collect and make public more data on Terry Stop instances and outcomes

- Future Analyses of Terry Stop Data:
  - Officer Demographics and Terry Stop Statistics
  - Multi-Target Classification Models

Thank you

Email: shadel96@gmail.com
GitHub: @shadel96
LinkedIn: linkedin.com/in/spencer-hadel/