

Predicting Traffic Citation Outcomes from Traffic Stop Data

By Kevin Diaz and Jeremiah Liao

Project Goal and Motivation

Goal: We want to create a model that can predict if a traffic stop resulted in a citation or a warning using data recorded by the officers during the stop.

- Created a model predicting the outcome with 75% accuracy

Key Questions:

- Which features have the greatest impact on the effectiveness of the model?
- Can we detect any biases in how citations are issued based on the importance of those features?

Traffic stops are one of the most common forms of police-citizen interaction. Therefore, it is important to analyze what information an officer may be relying on when deciding to issue a citation.

Models Evaluated

Primary model: **Histogram Gradient Boosting Classifier**

- Demographic information (race, gender)
- Traffic stop information (charge, time of day, day of week, presence of alcohol, etc)
- Vehicle information (vehicle type)

Other models compared:

- **Baselines:**
 - Random
 - Always Citation / Always Warning
- **Heuristic:** Predict the most common outcome per charge (e.g., speeding, DUI, etc).
- **HGB (Histogram Gradient Boosting) Classifier**
 - No demographic information (race, gender)

By removing or including features like race, gender, time, or location, we wanted to isolate the contribution of each factor and test whether demographic information is meaningfully affecting predictions.

Evaluation and Results

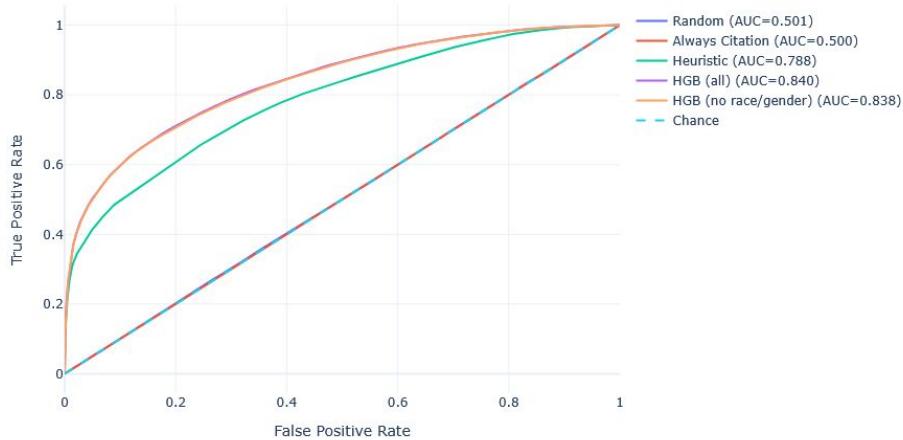
We were surprised by the performance of the heuristic algorithm. It performs fairly well, with an F1 Score of 0.696.

Our best model only slightly improved performance by including time, location, and other information about the arrest.

Removing or including race and gender as features seems to only marginally affect performance.

This seems to suggest that the charge field itself already captures most decision-making that the officer is doing.

ROC Curve Comparison



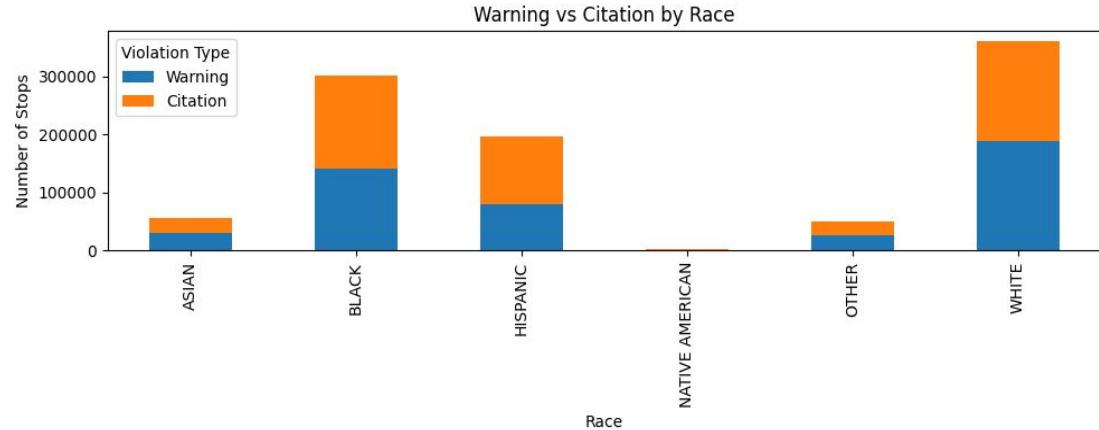
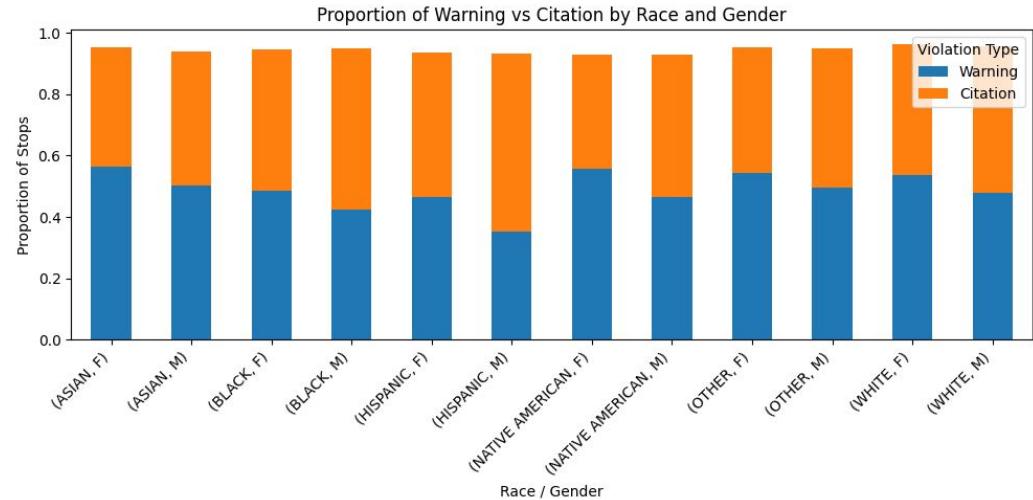
Model	Accuracy	F1 Score	ROC AUC
Heuristic (Charge-only)	70.4%	0.696	0.788
HGB (All Features)	75.6%	0.745	0.841
HGB (No Race/Gender)	75.4%	0.743	0.839

EDA Findings & Effect on model

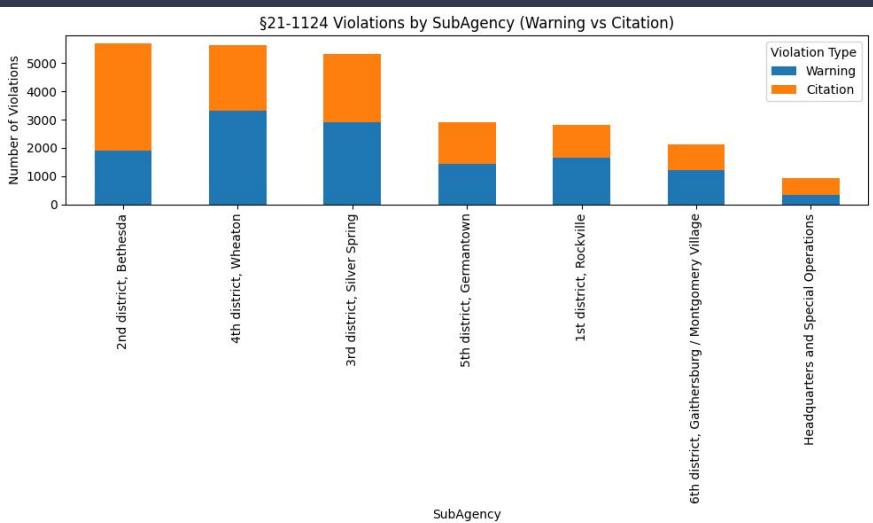
During EDA, we saw noticeable disparities in traffic stop outcomes based on race/gender.

Hispanic men have a citation rate of 58% compared to 38% for asian women.

Minority drivers are stopped at disproportionately higher rates than their share of the population of Montgomery county.



Effect of Features on Model Performance



- Charge is the primary predictor of citation outcomes
 - A simple heuristic model based only on the charge is already reasonably effective
- Other contextual boolean fields (Alcohol, property damage, belts, etc) were rarely populated, limiting our insight into the decision making process.
- Time and location differences resulted in the biggest performance gains. Enforcement seems to vary by location.
- Due to the lack of contextual features, we are unable to get many valuable insights into the reasoning of the officer when deciding on a charge.

Conclusion

Our original plan was to judge, based on feature importance, whether citations were being unfairly issued.

Our model was modestly effective as a classifier, but the lack of signal from any feature other than charge shows us that our dataset can't provide much insight into *why* an officer issued a citation, or if the charge was even justified. Especially since important contextual details are often left blank.

Our results suggest the need for stronger oversight and more comprehensive documentation of traffic stop reporting. Officers should be required to properly fill out all fields, and provide some form of standardized justification when choosing a charge.