Modeling Racial Disparities in Juvenile Justice

Researchers: Darcely Peña and Elsa Hrtiz

Introduction

This project expands on research initiated through an internship with the Cumberland County Juvenile Probation Office which revealed that youth from racially marginalized groups are disproportionately referred to the county's juvenile probation system and are subject to harsher disciplinary outcomes.

This study aims to identify the root causes of disproportionate referral rates in Cumberland County and examine their implications on structural bias within the U.S. juvenile justice system.

Objectives:

- Estimate referral rates on the census-tract level to understand key drivers of juvenile justice involvement.
- Predict the race of juveniles to determine which demographic and case related factors are strongly correlated with race within the juvenile justice system.

Methods

Datasets:

- **Pennsylvania Case Management System:** case records for juveniles from 2012 to 2022, along with information on charges and dispositions received
- **US Census - ACS Survey:** demographic, social, and economic characteristics for all census tracts in Cumberland County from 2012 to 2022

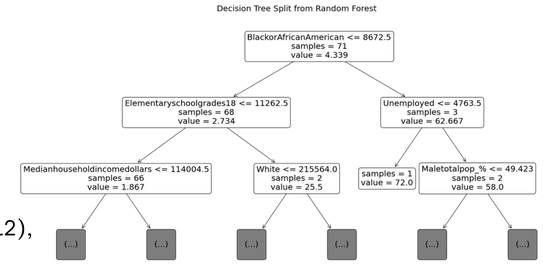
Models:

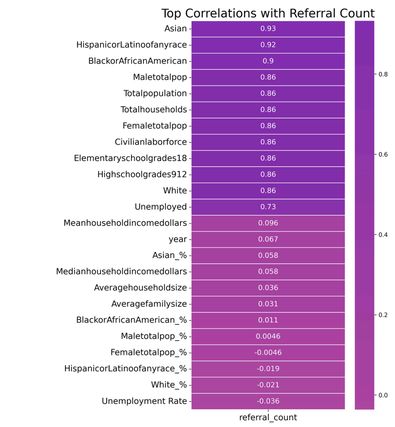
- A Random Forest regression model was employed to predict juvenile referrals at the census tract level in Cumberland County.
- A Random Forest model was trained to classify a referred recipient's race. The most influential variables were then examined to seek out patterns between the predictors for each race.

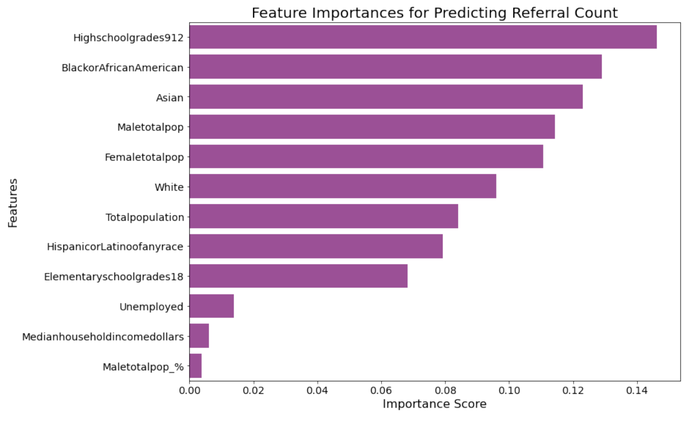
Results

Predicting Referral Counts:

- **Accuracy:** 94%
- **Mean Squared Error:** 1.28
- **Primary Split Features**: Black or African American, Count of Elementary Students, Unemployed Count
- **Top Feature Importances:** Count of HS Students (9-12), Black or African American, Asian
- **Top Correlations:** Asian, Hispanic, Black

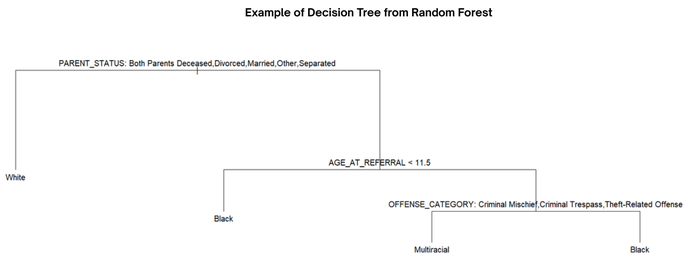


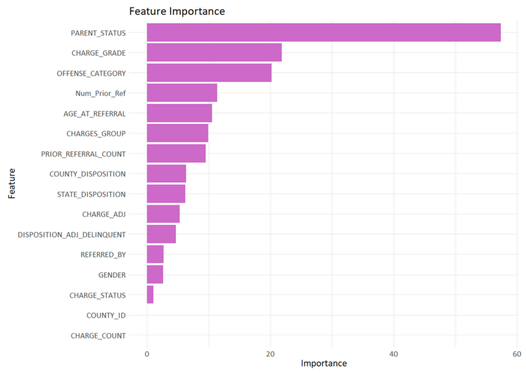


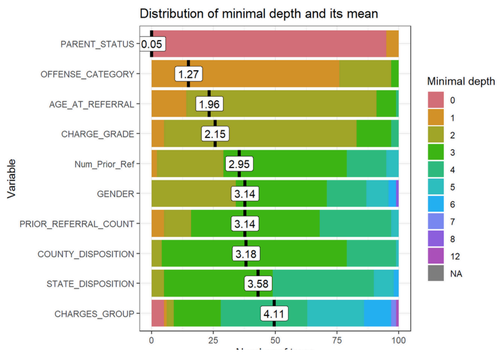


Predicting The Referred Recipient's Race:

- **Accuracy:** 83.6%
- **Primary Split Features:** parent status-- signifies whether the referral recipient comes from a home with married, divorced, separated, or single parent homes.
  - offense category
  - age at referral
- **Least Influential:** charge group,
  - state and county dispositions







Conclusions

Strong predictors of referrals counts are race, youth population density, and unemployment. This model highlights potential structural biases in the juvenile justice system as race is consistently the most important feature. The role of economic factors are secondary suggesting other systematic influences in demographic makeups.

Some of the strongest predictors of race in this analysis were factors not included in the judicial system, but instead in the recipient's homes. It was found that white referral recipients came from homes with both parents deceased, divorced, married, or separated, with black and multi-race recipients coming from mainly single parent homes. From there, age and type of offense were used to differentiate between the other races-- indicating certain races follow patterns in their age and type of crime committed.

Implications

Social:
This model indicates that areas with higher youth populations and certain racial demographics have higher referral rates. Social services and community organizations may use this data to target interventions and offer resources to at risk-youth. This includes recognizing youth of various races could be considered "at risk" for different reasons.

Legal:
The strong influence of race in predicting referral rates could indicate discriminatory practices within the juvenile justice system. This would call for legal reform aimed at reducing racial disparities.

Ethical:
If the model identifies a particular racial group as a key factor in higher referral rates or certain types of crimes, it can inadvertently stigmatize that community and perpetuate harmful stereotypes.