

# Fairness of Machine Learning Applications in Criminal Justice: Insights from Evaluation of COMPAS

Yue Hao

yqh5470@psu.edu



PennState



## Motivation

- On going debate on Correctional Defendant Management Profiling for Alternative Sanctions (COMPAS) -> **Based on Random Forest Classifier**
- Different evaluation metrics on "fairness"

Notion	Metirc
<i>equalized odds</i>	FPR and FNR
<i>calibration</i>	PPV

# Research Objectives and Methods

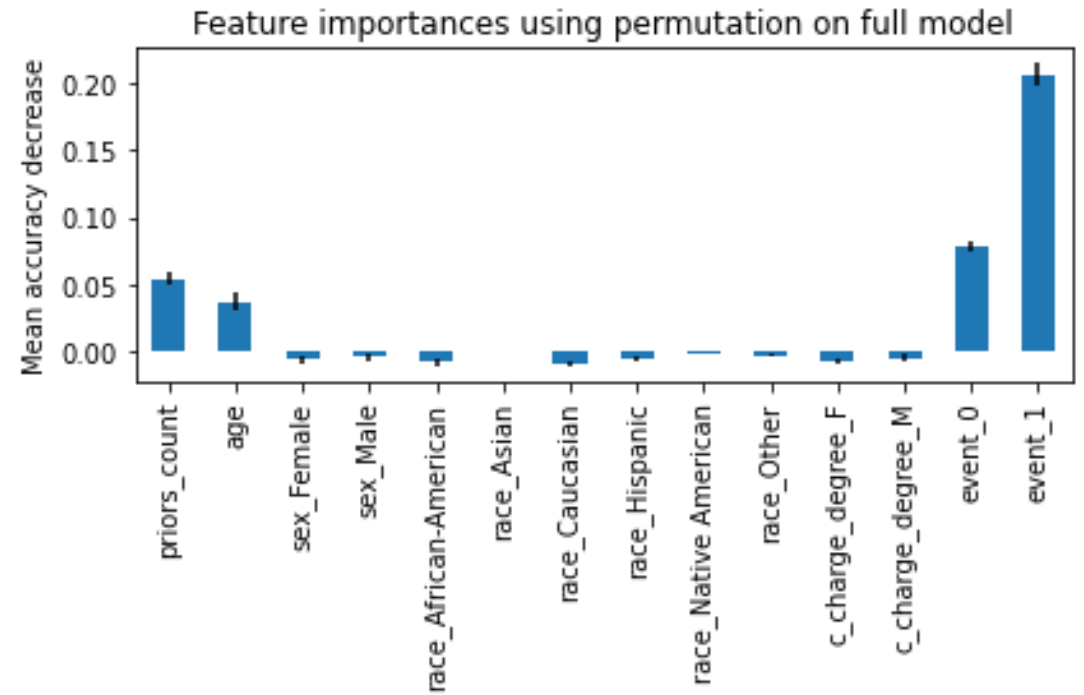
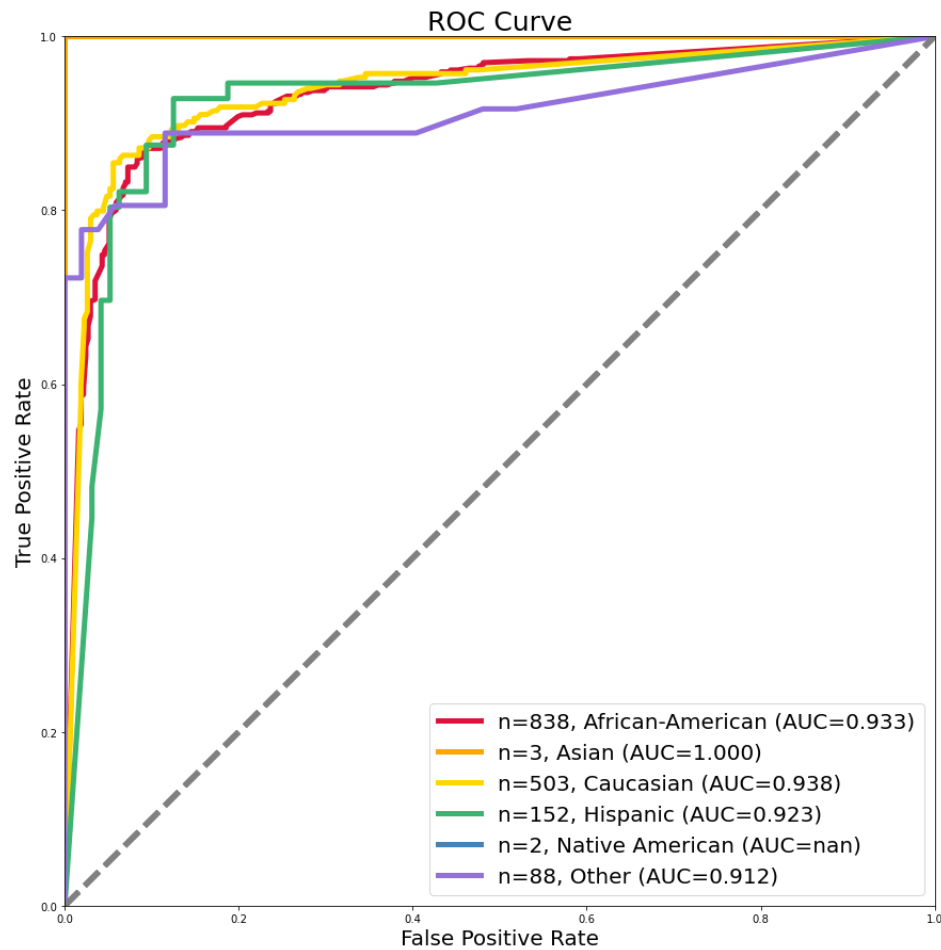
- Re-evaluating fairness of COMPAS based on different metrics
  - -> Suitableness of Random Forest model in criminal justice context

Train a RF classifier

evaluation

Racial Groups	FPR	FNR	PPV
<b>Native American</b>	0.0	0.0	1.0
<b>Asian</b>	0.0	0.0	1.0
<b>African-American</b>	0.0583	0.0657	0.9577
<b>Caucasian</b>	0.0246	0.0681	0.9654
<b>Hispanic</b>	0.0761	0.0489	0.9824
<b>Other</b>	0.0156	0.0465	0.9847
<b>Overall</b>	0.0388	0.0640	0.9602

# Results



# Insights

Social justice in different levels :

**social good vs. personal interest**

- Assumption of RF
  - Sampling is representative
- Use of features
- Dynamics of data inputs and impacts on outputs