

Magnifying the Inequities of Justice: An AI Problem

BY: DONALD COOPER, FALL DATA SCIENCE COHORT 2020

Yesterday's hotspots (08/26/2018)

Plan for today (08/27/2018)

Machine Learning to Predict Crime

It is no mystery that crime is a fact of life in any society. What is often a mystery, however, is the modern data-driven methods employed to reduce crime and track repeat offenders.

In the age of AI Algorithms, it only makes sense for Law Enforcement and the American Judicial System to stay cutting-edge and utilize Statistics and Data Science to reduce crime and the potential for repeated crime.

Private and Public organizations have gotten creative to tackle crime. Take for example the backdrop. The hotspot map of crime was provided by the business management consultancy Spunk Inc. It's innovative and made to be helpful for keeping communities safe.



Recidivism

But can these methods help to prevent crime? There is certainly the potential. But what it ultimately comes down to are those responsible for deploying the data driven solutions.

Take for example, predicting Recidivism, or the tendency of a convicted criminal to reoffend. An algorithm labeled COMPAS is responsible for streamlining the judicial process of determining sentencing based on predictive scores associated with the potential to reoffend.

What is COMPAS?

COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) is a popular commercial algorithm, designed by the Software Company Northpointe Inc, used by judges and parole officers for scoring criminal defendant's likelihood of reoffending (recidivism).

However, it has been shown that the algorithm is biased in favor of white defendants, and against black inmates, based on a 2 year follow up study (i.e who actually committed crimes or violent crimes after 2 years). The pattern of mistakes, as measured by precision/sensitivity is notable.



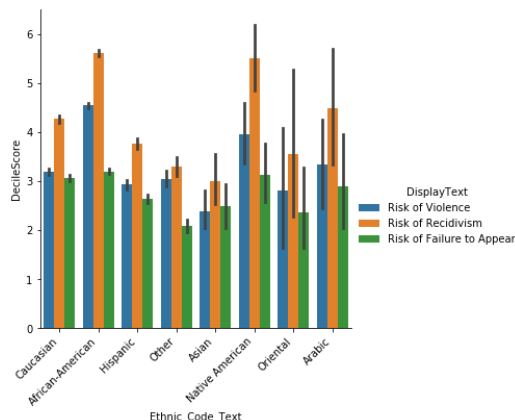
Why is this Important?

ProPublica, an organization specializing in investigative journalism who conducted the initial analysis mentions that “black defendants were often predicted to be at a higher risk of recidivism than they actually were”.

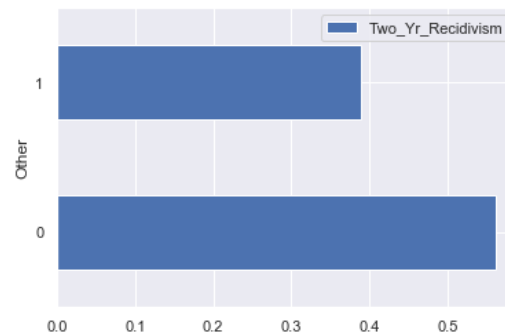
Their analysis found that black defendants who did not recidivate over a two-year period were nearly “twice as likely to be misclassified as higher risk compared to their white counterparts”. This is disconcerting. The algorithm essentially segregates based on race and fails to serve equal justice.



Assc. Risks by Ethnicity



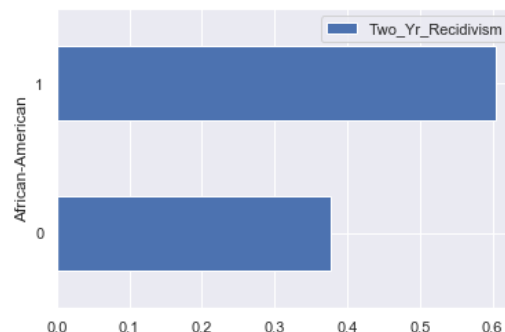
Recidivism Other Races



Recd. Risk by Ethnicity

| RecSupervisionLevelText | Low | Medium | Medium with Override | Consideration | High |
|-------------------------|------|--------|----------------------|---------------|------|
| Ethnic_Code_Text | | | | | |
| African-American | 9993 | 5328 | | 4626 | 2961 |
| Arabic | 36 | 15 | | 6 | 3 |
| Asian | 195 | 15 | | 9 | 9 |
| Caucasian | 9888 | 2652 | | 1659 | 669 |
| Hispanic | 5115 | 978 | | 507 | 273 |
| Native American | 84 | 45 | | 36 | 18 |
| Oriental | 30 | 0 | | 0 | 3 |
| Other | 1500 | 297 | | 117 | 57 |

Black Recidivism



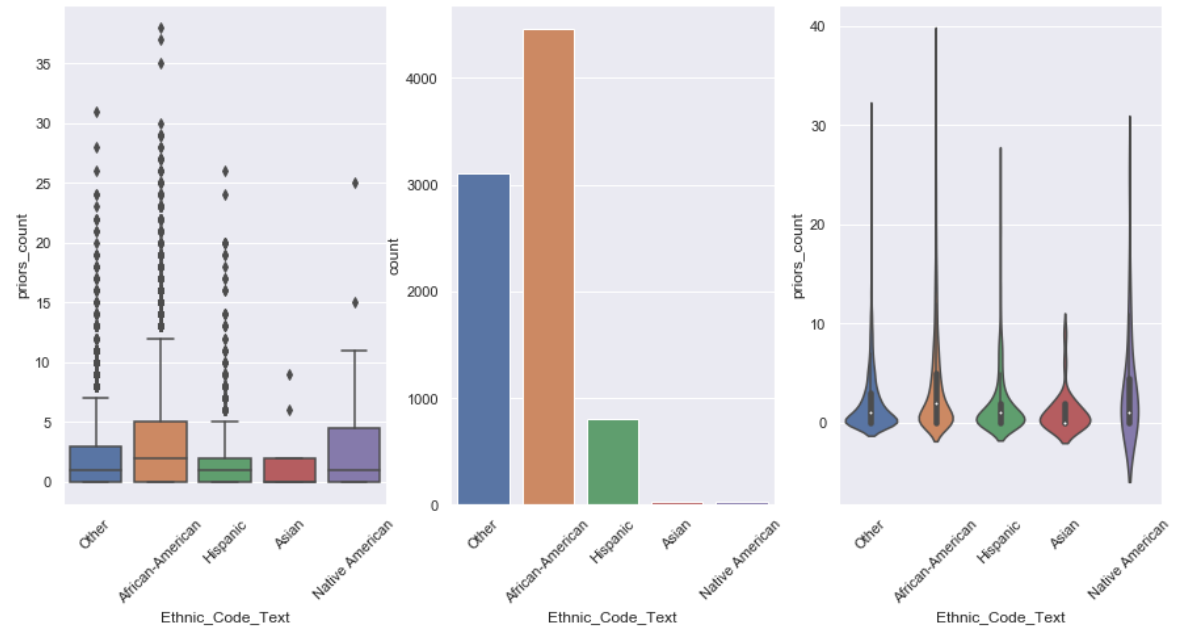
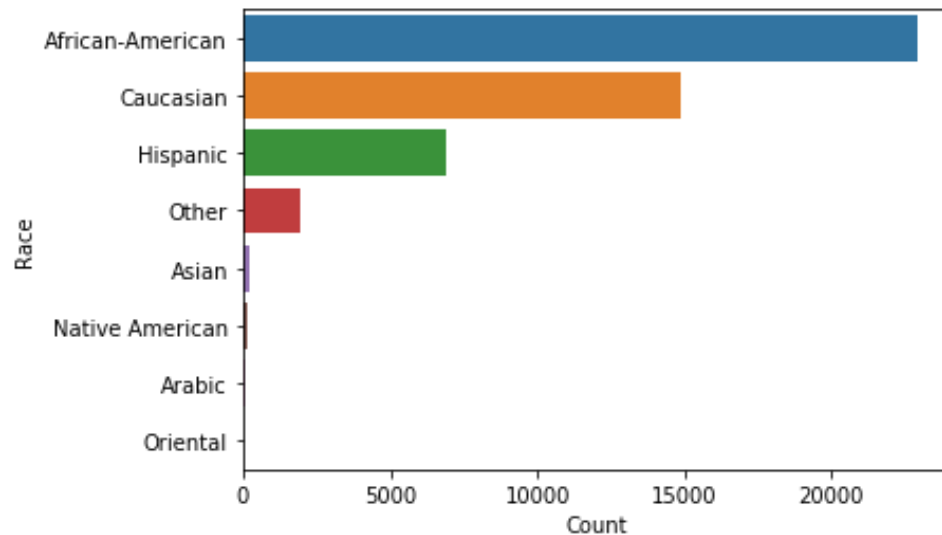
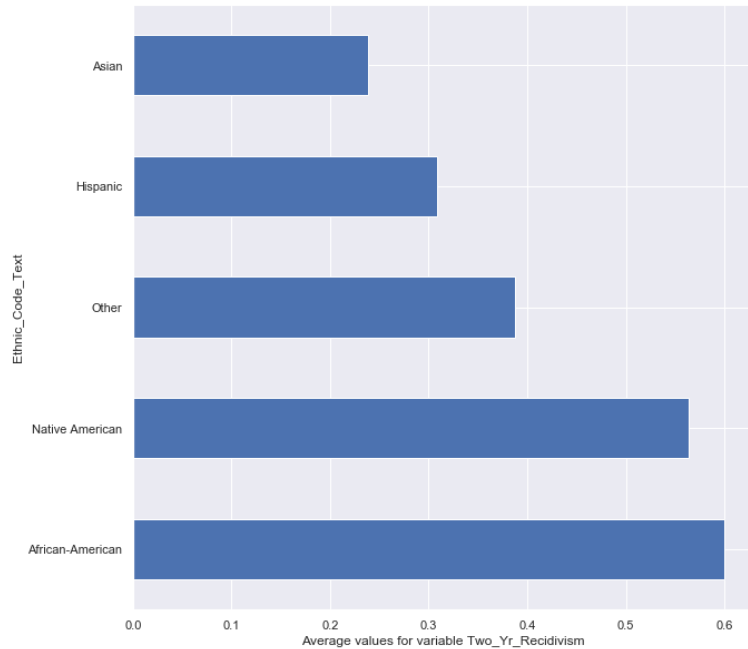
A Confirmation Study

To cross-examine these findings, an additional study has been conducted using similar and different metrics to predict two-year recidivism across different ethnicities. This analysis aims to provide a similar take on ProPublica's analysis as well as validate their findings.

With some simple data cleaning, basic visuals, and some counting, it was immediately apparent that there was already issues in the sampling done for the analysis.

Additional EDA

Some additional points regarding the distribution between ethnic groups were the inconsistencies between the number of prior convictions and the number of different individuals who identify as a different race.



Basic Modeling

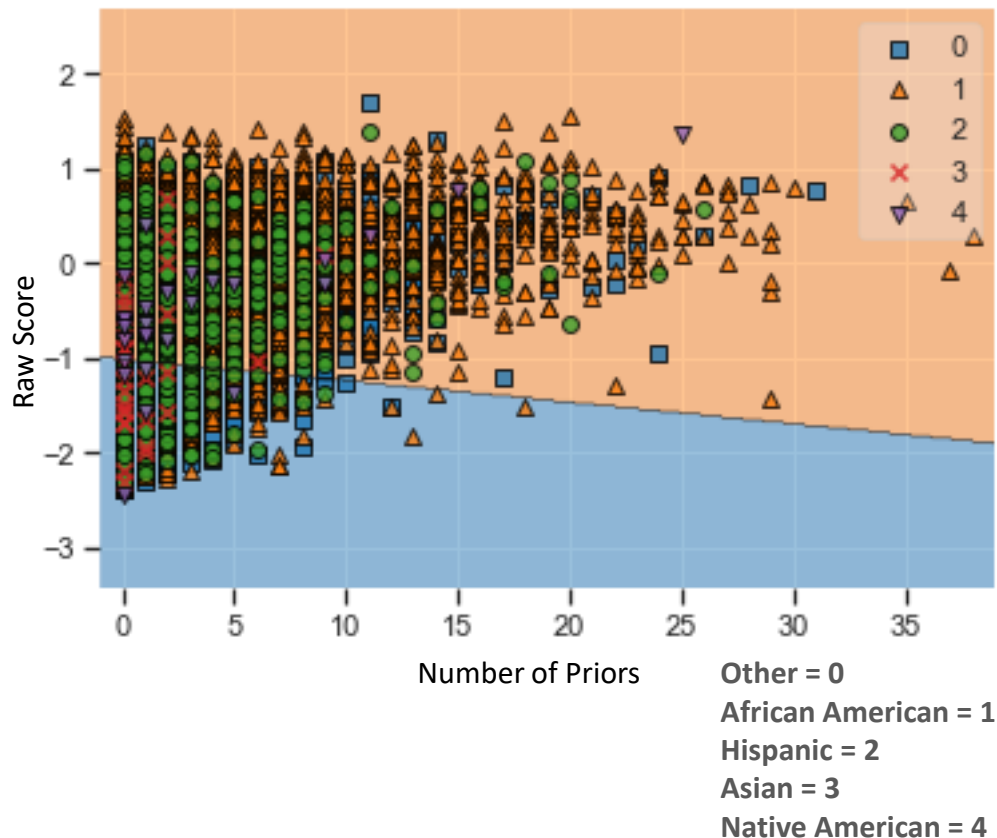
| Ethnicity | Logistic Regression | K-Nearest Neighbors Classification |
|------------------|---------------------|------------------------------------|
| African American | 0.7114 | 0.6981 |
| Asian | 0.8333 | 0.8333 |
| Hispanic | 0.8192 | 0.8192 |
| Native American | 0.5714 | 0.7143 |
| Other | 0.7286 | 0.7136 |

Modeling these predictors resulted in different predictive accuracies across all the ethnic groups. This is indicative of improper modeling technique across different subgroups. Additional modeling was conducted only to have low training and test scores, meaning that the predictive performance was low with the current data.

| | fit_time | score_time | test_score | train_score |
|---------------------|----------|------------|------------|-------------|
| Bayesian Ridge | 0.009642 | 0.001328 | 0.257562 | 0.260487 |
| LDA | 0.022938 | 0.000998 | 0.081649 | 0.074534 |
| KNN_Class | 0.009998 | 0.096077 | 0.242254 | 0.195149 |
| Tree | 0.010818 | 0.001995 | 0.118391 | 0.111280 |
| Logistic Regression | 0.012643 | 0.006851 | 0.053197 | 0.052012 |

Advanced Modeling

Linear SVM Plot



Using Support Vector Machines (SVM), it was worth looking at the dimensional space in which observations may reside by class-type. The Raw Score, or likeliness to reoffend, is plotted against number of priors.

It can be inferred from this plot that because of the uneven distribution between the different ethnic groups, there is an overemphasis in the SVM model in placing African Americans in the high risk for recidivism based on high priors and raw scores.

Conclusions and Thoughts

Based on the modeling from ProPublica, as well as additional data analysis and modeling conducted throughout this study, the COMPAS model does in fact segregate based on race factors and incorrectly predicts the risk of recidivism based on an unevenly distributed collection of data skewed heavily towards African Americans.

As mentioned in the beginning of this study, it is worth being methodical in how data is collected and analyzed. The COMPAS study not only should be revisited, but it should remind us all that there is a presence of inequity among our judicial process concerning the prediction of crime across select demographics.

