Analyzing the COMPAS Recidivism Algorithm

Raechel Walker[1], Matthew Taylor[2], Olivia Dias[3], Zeynep Yalcin[4], Cynthia Breazeal[5]

[1]Massachusetts Institute of Technology
raechelw@media.mit.edu

[2]Massachusetts Institute of Technology
mewtaylor@gmail.com

[3]Massachusetts Institute of Technology
omdias77@mit.edu

[4]Wellesley College
zy1@wellesley.edu

[5]Massachusetts Institute of Technology
cynthiab@media.mit.edu

Abstract

Data activism involves using data science to challenge power inequalities, such as racism. While we have created a five part data activism curriculum, we will discuss the few portions of the curriculum that emphasize AI fairness. In this set of activities, students learn how to use data science to recognize, mitigate, and advocate for people that are disproportionately impacted by systemic inequality. Students execute the entire data science development process (data mining, data cleaning, data visualizations, and modeling). First, students will learn about the historical bias in arrest data. It is essential that students understand the nuances of systemic racism in the criminal justice system, so they can understand how AI can amplify racism. In the final assignment, students will analyze an infamous AI algorithm, called the Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) recidivism algorithm. The project will teach them the critical thinking and programming skills required to challenge social injustice. This assignment aims to enlighten students about the importance of analyzing algorithms throughout the entire data science development process. These skills can be applied to a data set students are passionate about, such as predictive policing and bias facial recognition.