HW 5

Anna Fetter

11/08/2024

This homework is meant to give you practice in creating and defending a position with both statistical and philosophical evidence. We have now extensively talked about the COMPAS [1] data set, the flaws in applying it but also its potential upside if its shortcomings can be overlooked. We have also spent time in class verbally assessing positions both for an against applying this data set in real life. In no more than two pages [2] take the persona of a statistical consultant advising a judge as to whether they should include the results of the COMPAS algorithm in their decision making process for granting parole. First clearly articulate your position (whether the algorithm should be used or not) and then defend said position using both statistical and philosophical evidence. Your paper will be grade both on the merits of its persuasive appeal but also the applicability of the statistical and philosohpical evidence cited.

STUDENT RESPONSE

The COMPAS algorithm should not be used as a sole arbiter of sentencing or parole decisions. COMPAS is both morally and statistically troubling as the COMPAS algorithm performs differently across racial lines.

COMPAS, which stands for Correctional Offender Management Profiling for Alternative Sanctions, was developed by Northpointe (now known as Equivant) as a way to predict recidivism among defendants. In 2009, Northpointe defined recidivism as "a finger-printable arrest involving a charge and a filing for any uniform crime reporting (UCR) code." COMPAS came under fire when ProPublica published a 2016 study that revealed that COMPAS overpredicts recidivism among black defendants and underpredicts recidivism among white defendants.

ProPublica's 2016 study on COMPAS found that it was only 61% accurate at predicting recidivism over a two year period. This isn't much better than the results if this were a fair coin toss. Even worse, Propublica's analysis revealed COMPAS only correctly predicted violent recidivism 20% of the time.

The accuracy alone is troubling, not to mention the statistical fairness criteria COMPAS violates. COMPAS violates independence. Independence requires that classification rates between two binary class labels are comparable between sets. Independence criteria doesn't take the ground truth into consideration, so violating independence alone wouldn't be a reason to throw out an algorithm. As found by ProPublica, there is a difference in recidivism rates across racial and gender lines. COMPAS also violates separation and equalized odds. Equalized odds requires that the algorithm has equal false positive and false negative rates for each demographic group, ensuring that predictions are equally accurate across groups. In our calculations in class, we proved that COMPAS was above the epsilon of $=.2$, set by the four-fifths legal standard, showing that COMPAS violated equalized odds.

COMPAS does not explicitly consider race in its decision factors, but may consider proxy factors such as zip code, which correlates to race. This raises serious ethical concerns around the use of an algorithm like COMPAS. Given the United States' history of racialized laws, from slavery to Jim Crow laws to redlining, the use of an algorithm like COMPAS that shows algorithmic bias surrounding race violates the ideal of equality and is an irresponsible tool that could reinforce racial bias and harm minority communities.

The moral framework of deontology, which has Greek roots and roughly translates to "the study of what we owe" provides the strongest argument against the use of COMPAS. German philosopher Immanuel Kant's

---

[1] https://www.propublica.org/datastore/dataset/compas-recidivism-risk-score-data-and-analysis
[2] knit to a pdf to ensure page count

universal maxim formula states that an act can only be justifiable insofar that it can be universalized to all of humanity. COMPAS' opacity as a black box algorithm makes it close to impossible to ensure that COMPAS' powers are being applied universally and fairly across all demographic groups. The core problem of COMPAS, that it overpredicts recidivism among black defendants and underpredicts recidivism among white defendants, makes it impossible to argue the universal application. The ends-not-means formulation prohibits the use of instrumentalizing a moral agent. The use of an algorithm instrumentalized moral agents to predict recidivism based on group characteristics, ignoring the individualizing and "human factor" of each defendant.

The use of COMPAS also raises some concerns around due process. The due process clause found in the Constitution prohibits the deprivation of "life, liberty, or property" by the federal and state governments, without due process of law. In State vs. Loomis[3], the Wisconsin Court ruled the COMPAS did not violate due process, if COMPAS is employed properly and with awareness of COMPAS's limitations. The ruling instructed sentencing courts to use COMPAS scores as one of many factors to ensure that sentences remain individualized and not solely based on this algorithm. However, many critics of the ruling argue that since COMPAS operates as a "black box", meaning the proprietary algorithm isn't available or understandable. It may not be clear to judges of the defendant what exactly COMPAS did to determine the rate of recidivism. Critics say that the use of COMPAS makes it difficult for defendants and their legal representation to understand the precise line of reasoning that led to the increased sentencing.

Between COMPAS failing to statistical measures of fairness and the racialized algorithmic bias, COMPAS should not be used as a supplement in the sentencing process because the risk to minority populations and the integrity of the criminal justice system is not worth the potential gain in courtroom efficiency.

---

[3]https://studicata.com/case-briefs/case/state-v-loomis/