HW 5

Serena Zhang

3/21/2024

The COMPAS algorithm should not be used in the decision-making process for granting parole. COMPAS raises issues of inaccuracy, statistical fairness, and accountability. Due to these factors, its potential for overriding a judge's discretion in the parole decision is a significant cause for concern.

COMPAS is an algorithm created by Northpointe, Inc. that predicts recidivism probabilities for criminal defendants. Inmates answer a questionnaire when booked into jail, and their responses are fed into COMPAS to generate recidivism risk scores. The algorithm is black box, meaning it is so complex that there is no way for any human to understand how the variables fed into the algorithm are used to make predictions, even its creators. Therefore, when COMPAS makes incorrect predictions, how exactly it came to that conclusion is a mystery. This raises issues of accountability: if a COMPAS prediction for an inmate is ultimately incorrect, and COMPAS was used by a judge in making a parole decision for that inmate, who is to be held accountable for that decision?

In fact, COMPAS makes incorrect recidivism predictions often – a ProPublica study found that COMPAS is only 65% accurate, an accuracy rate only marginally higher than a coin flip. Another salient issue arises out of the fact that COMPAS is wrong in different ways for different groups of people: it is almost twice as likely to wrongly label Black defendants as high risk recidivists, and almost twice as likely to wrongly label White defendants as low risk.

In another world where parole decisions are entirely dictated by COMPAS, Black defendants at low risk of recidivism would be far more likely to be wrongly kept in prison, and White defendants at high risk of recidivism would be far more likely to be prematurely released.

Therefore, it is unsurprising that COMPAS fails multiple measures of statistical fairness. All algorithms will have some degree of error, and statistical fairness measures whether these errors affect certain groups disparately. Northpointe has argued that race is not directly used in the COMPAS algorithm, which means it cannot be unfairly biased against Black defendants. However, zip code, which serves as a proxy for race, is used. Because one's zip code is a strong predictor of one's race, the inclusion of zip code makes it possible for COMPAS to be unfairly biased against defendants of different races, as explained above. In this way, use of COMPAS in parole decisions potentially exacerbates issues of systemic racism already present in the justice system.

From a utilitarian standpoint, the costs of using the COMPAS algorithm far outweigh the benefits. The above issues of inaccuracy and fairness have harmful, far-reaching implications for the justice system, especially when COMPAS is only 65% accurate in its predictions. With such a high level of inaccuracy, the amount of good that COMPAS could do in protecting the public from reoffenders is dubious. Ultimately, a judge's discretion remains the most important in making parole decisions. While COMPAS is just one potential tool in a judge's toolbox, it is easily misunderstood, often inaccurate, and could potentially skew or overpower a judge's decision. Ultimately, human discretion is what matters when making decisions about human lives. Therefore, COMPAS should not be used in making parole decisions.