

HW 5

Luca Buchoux

12/29/2023

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 and 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 philosophical evidence cited.

While COMPAS has the potential to be a useful tool, there are several statistical and ethical concerns that should be considered before fully integrating it into your decision making process. Despite the argument that perfect fairness is impossible in algorithmic models, the use of COMPAS in its current form is unjustifiable due to its potential to perpetuate racial disparities and worsen systemic biases. Also, relying too heavily on the algorithm's outputs without addressing its issues can lead to unjust outcomes.

One of the main assessments for fairness in an algorithm is equalized odds, which requires that the performance be consistent across different groups, particularly in terms of false positives and false negatives. It has been shown that COMPAS does not satisfy this criteria, specifically when considering racial groups. Black defendants are often assigned higher risk scores compared to white defendants, even when accounting for similar criminal histories and other relevant factors. This bias raises a critical concern since it implies that certain racial groups are more likely to be deemed a higher risk of recidivism based on the same set of facts.

From a statistical standpoint, it is important to recognize that achieving perfect fairness across all demographic groups is impossible. No algorithm can be entirely free of bias, and the concept of fairness is not cut and dry, so there must be trade-offs between different types of fairness metrics. However, when an algorithm systematically disadvantages one group at the expense of another, these shortcomings cannot be dismissed. The real world impact of these unfair outcomes is significant, as predictions with higher risk scores may influence your decision on parole, which can lead to longer periods of incarceration and perpetuate existing inequalities in the criminal justice system.

While it is true that no algorithm can be perfectly fair in every context, the ethical responsibility to minimize harm should be the priority. The use of COMPAS without addressing its statistical bias can lead to negative real world consequences, particularly for marginalized groups. In this case, black individuals are more likely to be incorrectly labeled as high risk offenders. These outcomes are not just results from a statistical model. They have real consequences on people's lives and contribute to the ongoing overrepresentation of black individuals in the U.S. criminal justice system.

You should not rely on the argument that achieving perfect fairness is impossible as a reason to accept these disparities. Instead, the focus should be on improving the model to reduce these biases and better align the algorithm's predictions with the ethical standards expected in the criminal justice system. The goal should

¹<https://www.propublica.org/datastore/dataset/compas-recidivism-risk-score-data-and-analysis>

²knit to a pdf to ensure page count

not be to accept flawed tools but to enhance them, ensuring they are as fair and accurate as possible given the data and available resources.

Another concern about the use of COMPAS is the potential for overreliance on the algorithm's outputs. Even when used as an aid in decision making, there is a risk that you may place too much trust in the algorithm's results, especially since those results appear to be objective. This risk of cognitive bias is well known. Individuals tend to lean towards algorithmic outputs because they assume that they are more accurate or objective than their own judgment. This effect could inadvertently result in you placing too much weight on the predictions generated by COMPAS, potentially leading to decisions that do not fully consider the specific circumstances of a case.

These decisions have life-altering consequences for defendants. If you rely too heavily on an algorithm that is known to have significant racial biases, there is a risk that you will reinforce the same biases that COMPAS is supposed to help mitigate. Also, if you fail to consider the broader historical context of the data through which these predictions are made, you may inadvertently contribute to further deepening the disparities already present in the justice system.

I recommend that you not treat COMPAS as a final decision making tool. Instead, it could be used as one of many factors to be considered in the context of an individual's case. The algorithm's predictions should be evaluated alongside other forms of evidence and your own professional judgment. In other words, COMPAS should not replace your discretion but could complement it, while you retain the ultimate responsibility for decisions.

In conclusion, while COMPAS has the potential to be a valuable tool in assessing the risk of recidivism, its current state raises significant ethical and statistical concerns. The racial disparities in its predictions, the risk of overreliance, and the potential for reinforcing systemic biases all point towards the need for significant improvements before its use can be justified. Until these issues are addressed, the use of COMPAS as a primary decision making tool is not justifiable because it risks perpetuating harm rather than advancing fairness in the criminal justice system.