HW 5

Student Name

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 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.

Answer

As a statistical consultant, I would recommend that the COMPAS algorithm should NOT be used in parole decisions. While it is an impressive tool with great computational capabilities, its technical flaws and ethical implications make it an unsuitable tool as it introduces a lot of fairness and reliability concerns.

One of the biggest shortcomings of COMPAS is that it gives significantly different false positive rates for different racial groups. More specifically, Black individuals are labeled as high-risk far more often as compared to White individuals, with nearly twice the likelihood of being labeled as future offenders or repeat offenders. In this manner, the principle of equalized odds is violated, which states that an algorithm should have a consistent error rate across different groups. Since COMPAS succumbs to such disparities between groups, it participates in the bias that exists in the system instead of mitigating it. Philosophically, this unequal error rate goes against a fundamental concept of fairness, as it means that different groups experience different outcomes even when they exhibit similar behaviors or characteristics.

Rawls' Veil of Ignorance states that systems should be chosen as if behind a "veil of ignorance" where decision-makers don't know their social position or identity. If we were to design a fair justice system using this ideology, we would likely reject COMPAS. Such a tool would be unacceptable as it creates an unfair disadvantage for communities that are already marginalized. Implementing COMPAS would violate the principles of justice and fairness that Rawls' Veil aims to achieve.

Moreover, there are concerns regarding the accuracy of COMPAS, which sits at a mere 65%, which indicates that it gets about 2 out of 3 predictions correct. An untrained human can achieve similar accuracy rates (about 62%), which brings into question the value that COMPAS adds to the system, if any. This degree of uncertainty is very high considering we are talking about a person's life and/or freedom. We cannot simply overlook the incorrect decisions it makes since the stakes involve individual liberty and are therefore very high.

There are also methodological flaws, such as the use of proxy variables (like ZIP codes) that introduce multicollinearity with protected features and lead to indirect racial bias through these factors that seem inert otherwise.

Utilitarianism also states that actions should aim to maximize societal well-being. This means balancing public safety by correctly identifying high-risk repeat offenders but also providing those truly willing to undergo a change with rehabilitation in a fair manner. The use of COMPAS would prevent both of these. People seeing a biased

algorithm like COMPAS being used would lead them to question the authenticity of the judicial system. From a utilitarian perspective, it simply does not make sense to use a tool that does not reliably enhance predictive accuracy.

While we know that both judges and algorithms are susceptible to biases, there is a huge difference when it comes to accountability. Judges can be held accountable for their decisions, and more importantly, their decisions are more contextualized for individual cases. But an algorithm like COMPAS lacks transparency, which gives us very little insight into the reasoning behind a given output. This opacity makes it difficult to criticize the tool as the precise calculations are difficult to pinpoint versus the justification that a judge might provide.

Finally, the decision-making process should be grounded in human judgement, not an algorithm, especially when it comes to matters involving the liberty of people. Delegating a decision like this to an algorithm removes the empathy and context sensitivity that a human brings to a complex case. Judges, while they may have their own biases, are humans who have a nuanced understanding and ethical reasoning capacity that cannot be emulated in an algorithm as it cannot understand the human capacity for reformation and rehabilitation. While the desire to incorporate data-driven decision-making via the means of COMPAS is tempting, its statistical shortcomings and philosophical implications make it an insufficient tool.

---

1. https://www.propublica.org/datastore/dataset/compas-recidivism-risk-score-data-and-analysis (https://www.propublica.org/datastore/dataset/compas-recidivism-risk-score-data-and-analysis)↩

2. knit to a pdf to ensure page count↩