HW 5

Logan Schmitt

03/27/2024

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 philosophical evidence cited.

Logan Input:

Your honor: I strongly urge you to reconsider using the COMPAS algorithm to evaluate a criminal defendant's risk of reoffending and their suitability for parole release. There are significant concerns regarding not only its statistical reliability but also its ethical nature as well.

In analyzing the statistical concerns, it's important to take into consideration the confusion matrices for White and Black defendants. These depict the accuracy of the COMPAS algorithm by showing the true positive, false positive, false negative, and true negative rates for each group. In this context, it shows how often the algorithm correctly predicts a defendant's probability of reoffending, therefore leading to the decision to approve or deny probation. Below, the first table represents the confusion matrix for White defendants and the second table represents the confusion matrix for Black defendants:

```
##                   Did Not Reoffend Did Reoffend
## Will Not Reoffend             1139          461
## Will  Reoffend                 349          505

##                   Did Not Reoffend Did Reoffend
## Will Not Reoffend              990          532
## Will Reoffend                  805         1369
```

Based on these results, some disparities exist among the classification rates between the classes of defendants. The COMPAS algorithm classifies Black defendants as being at a much higher risk of reoffending than White defendants. This already signals the potentially problematic nature of these results. As a result, we can further analyze these under the two notions of statistical fairness to test whether disparate impact, that is, unintentional discrimination towards a certain protected group, exists.

```
##                                                    Measure    Result
## 1                Overall accuracy for White defendants 0.6699267
## 2                Overall accuracy for Black defendants 0.6382576
## 3    Disparate impact between White and Black defendants 0.5916376
## 4  Statistical parity between White and Black defendants 0.2402002
## 5 Predictive equality between White and Black defendants 0.2139250
```

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

6   Equal opportunity between White and Black defendants 0.1973730

Both these notions, using a given threshold (i.e., $\epsilon = 0.2$), test whether there are roughly comparable rates of classification into the two class labels across protected and unprotected sets, in this case, Black and White. The COMPAS algorithm, despite having a relatively high overall accuracy, still violates both notions. Using the specified threshold, to test whether these results are significant, all, except for predictive equality, are violated. Therefore, it indicates that the COMPAS algorithm demonstrates disparities in treatment between protected and unprotected groups— specifically Black and White criminal defendants. This discrepancy raises doubts about the fairness and impartiality of the COMPAS algorithm, particularly in how it might perpetuate preexisting biases within the criminal justice system.

Aside from having statistical concerns, the COMPAS algorithm contains several moral and philosophical issues as well.

Firstly, we must consider the fact that race is not directly used as a variable within the algorithm. Instead, a proxy for race, zip code, is used. This alludes to the idea that using race itself is entirely problematic and can lead to many issues, such as further perpetuating biases and creating negative feedback loops. However, zip code is highly correlated with race. As a result, this proxy has the potential to yield the same results as using race, thereby introducing indirect racial biases into the system. This indirect approach of incorporating race into the COMPAS algorithm obscures sources of biases, making them harder to identify and address.

Secondly, there is a complete lack of transparency in how the algorithm works. This makes it incredibly difficult to inspect the algorithm for fairness, understand how a specific risk score was calculated, or challenge potentially biased results. As a judge, you presume all accountability in the decisions you create; however, this is entirely not possible when using an algorithm that you not only fail to understand but also did not create. This effectively fails to ensure that decisions are made based on just and equitable grounds.

Finally, analyzing the COMPAS algorithm from a deontological perspective violates the second categorical imperative. Reducing defendants to data points takes away certain moral rights that we as humans have, such as human dignity. Such an approach neglects the moral worth of individuals and their capacity for change, treating them instead as sources of study defined solely by their past actions. Furthermore, the algorithm has already been shown to display certain biases that affect an individual's opportunity to gain freedom. At a minimum, we are obligated to ensure that the algorithm has perfect accuracy, however, this may not be achievable. Therefore, we must strive for fairness, which the COMPAS algorithm appears to not adequately accomplish. These factors combined effectively use these people as mere means to an end further using them as ways to improve on an already biased classification algorithm.

Taking into consideration these statistical and ethical issues surrounding the COMPAS algorithm, I encourage you to avoid using it in the decision making process for granting parole.