

STOR390HW5

Jillian Myler

2024-03-25

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 COMPAS¹ 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 pages² 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.

In recent years, the United States court system has seen an increase in utilizing algorithms in the process of setting bail for criminal offenses. These algorithms aim to provide the judge with a risk metric based on the offender's background quantifying the likelihood that the individual will repeat the crime that they are currently being charged with. One prominent algorithm being used this way is COMPAS, which stands for Correctional Offender Management Profiling for Alternative Sanctions, developed by Northpointe and currently used in the Broward County, Florida court system. While the intent of using these algorithms is to facilitate more informed decisions by judges, the COMPAS algorithm should not be used in any court system.

The COMPAS algorithm is not very accurate. When the offenders are placed in jail, they fill out a questionnaire to be fed into the algorithm. From this data, COMPAS assigns each offender a decile score on a scale from one to ten, where one is associated with a very low risk of recidivism and ten is associated with a high risk. However, when looking at the predicted recidivism scores of over 10,000 offenders, the algorithm was only 61% accurate. Further, the algorithm was only 20% accurate concerning its recidivism predictions in violent crimes. When thinking about using no algorithm at all, the assumed rate of accurately predicting that someone will repeat their offense would be 50%. Therefore, an algorithm with an accuracy rate of 61% has very little marginal improvement for informed decision-making. Moreover, it is worse than an uninformed guess with respect to its violent crime predictions.

While the COMPAS algorithm is quite inaccurate, it is also important to examine where, or with whom amongst the population, its inaccuracies in risk classification lie. When holding all else equal, black defendants were nearly twice as likely as their white counterparts to be incorrectly classified as high risk and expected to recidivate when they would go on not to. Further, white individuals were much more likely to be predicted not to recidivate when they would eventually re-offend than black individuals. Hence, the algorithm misclassified systematically in favor of white individuals. Thus, if a judge were to use this algorithm, it would not only be largely inaccurate but also arbitrarily discriminatory.

As a statistician, replicability is important, as is understanding whether there is a bias that may arise from the training-testing partition of a given model. To address replicability, this algorithm is a black box, meaning that its methods are not publicly available, only its predictions based on a given input. Concerning the training-testing partition, in this case, the training data would be the population data used to predict recidivism and the testing data would be the individuals whose scores were output by COMPAS and either

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

²knit to a pdf to ensure page count

repeat offenders or didn't. It is important now to realize that the trends of recidivism occurring in the training dataset may stem from deeply ingrained systematic issues within society. In the same manner that one would expect a continued overabundance of heads when flipping a weighted coin that favors heads, the algorithm would perpetuate the advantages or disadvantages to racial groups that are already embedded within society.

Despite discussing statistical reasons to not use the COMPAS algorithm in court such as its inaccuracy level, black box creation, and systematically biased training data, it is also important to consider algorithmic use at an ethical level. This requires consideration both of the intent of the algorithm and the potential outcomes for society as a whole. Using a consequentialist approach, such that the right choice is the one that results in the greatest benefit, the COMPAS algorithm, arguably any algorithm, should not be used in court. While the intent of using algorithms in court may be to yield more informed decision-making and speed up the arduously slow legal system, their use may cause more harm to an individual, and society, than gain. For example, even a near 99% accurate algorithm should not replace a judge as it would not allow for long-term change or individualized consideration. Indeed, each individual has a background and circumstance that is not fully capturable by numbers or labels that can be input into an algorithm. Even if it wasn't about completely replacing a judge, the very presence of an algorithm may predispose the judge's thinking, slowly dehumanizing the justice system. Overall, the potential for perpetuation of pre-existing problems within society as well as inching away from individualized and circumstantial consideration by a judge implies far more detriment from algorithm usage than benefit.