

# HW 5

Pranav Kallem

11/07/2024

The COMPAS algorithm, developed by Northpointe, Inc., aims to predict the likelihood of recidivism for criminal defendants. Despite its growing use in the judicial system, concerns regarding its fairness and predictive accuracy have surfaced. As a statistical consultant advising a judge, I recommend against the inclusion of COMPAS results in parole decisions due to its inconsistent performance across racial groups and its inability to achieve an acceptable level of predictive accuracy. Both statistical evidence and ethical considerations make it unwise to rely on this tool for decisions that profoundly impact individuals' lives.

The use of algorithms like COMPAS in the criminal justice system is intended to enhance objectivity and reduce human biases. However, these tools often introduce new biases or aggravate existing ones, particularly when the data used to train the algorithm reflects societal inequalities. The COMPAS algorithm, for instance, has faced significant investigation for its flawed predictions which disproportionately affect minority groups. This not only questions the reliability of the tool but also raises concerns about the ethical implications of using such technology in life-altering decisions.

The ProPublica analysis of COMPAS found significant racial disparities in its predictions. While the overall accuracy of predicting recidivism was approximately 61%, the predictive value significantly differed between racial groups. Black defendants were more likely to be incorrectly classified as high risk, with nearly twice the false positive rate compared to white defendants who were mistakenly classified as low risk (45% for Black defendants versus 23% for White defendants). This suggests a systematic bias that disproportionately affects Black individuals, leading to harsher outcomes.

Moreover, the COMPAS score correctly predicted violent recidivism only 20% of the time. Even when controlling for age, gender, and prior criminal records, racial disparities persisted. Black defendants were 77% more likely to receive higher scores for violent recidivism than White defendants, highlighting the algorithm's inability to treat similar cases equally. The inherent biases in the predictions indicate that the COMPAS tool is not only unreliable but also perpetuates racial inequalities that are prevalent in the justice system.

In addition to racial biases, the COMPAS algorithm's overall predictive accuracy has been questioned. A concordance score of 63.6% was reported, indicating that the tool's ability to rank defendants' recidivism risk is only marginally better than chance. According to Northpointe's guidelines, a predictive accuracy of 70% or higher is typically considered acceptable. The failure of COMPAS to reach this threshold further supports the argument that its use in judicial decisions may be inappropriate. Additionally, the discrepancies in the performance of the tool across different demographic groups suggest that it does not provide a consistent measure of risk.

The statistical shortcomings of the COMPAS algorithm are not limited to racial bias. The predictive power of the algorithm varies significantly based on age and gender as well. For instance, younger defendants are more likely to be assigned higher risk scores, even when controlling for prior offenses and other factors. This is particularly concerning, as it implies that young individuals are being disproportionately penalized by the system, which may limit their opportunities for rehabilitation and reintegration into society. Female defendants, despite having lower overall rates of criminality, were also found to be more likely to receive higher risk scores compared to their male counterparts, suggesting that the algorithm may contain inherent gender biases as well.

From a deontological perspective, the use of the COMPAS algorithm in parole decisions is ethically problematic because it violates the principle of justice and fairness. The observed racial bias in the risk assessments indicates that similar individuals are not treated equally under the algorithm. Judges are ethically bound to ensure that decisions are made equitably, without bias based on race or any other immutable characteristic. Relying on an algorithm that demonstrably discriminates against Black defendants would undermine the justice system's commitment to impartiality.

Furthermore, the principle of respect for persons is compromised when a flawed algorithm is used to make determinations that affect people's freedom. Defendants are entitled to individualized assessments, and using a generalized, biased tool like COMPAS deprives them of the right to be evaluated on their unique circumstances. The lack of transparency in how the COMPAS algorithm calculates risk scores also makes it difficult for defendants to contest or understand the basis of decisions made about their future, which further undermines their autonomy and right to a fair hearing.

Additionally, under consequentialist reasoning, the negative consequences of using COMPAS outweigh the potential benefits. The disproportionate number of false positives for Black defendants could lead to unnecessarily extended incarcerations, exacerbating the systemic inequities in the criminal justice system. This unjust treatment has profound impacts on defendants' lives, affecting their families and communities, and could result in a mistrust of the judicial process by marginalized communities. By potentially causing more harm than good, the use of COMPAS in parole decisions fails the utilitarian test of promoting the greatest good for the greatest number of people.

The ethical issues with using COMPAS also extend to its impact on public trust in the judicial system. When communities perceive that decisions are being made based on biased algorithms, confidence in the fairness and legitimacy of the judicial system is eroded. This can lead to broader societal consequences, such as reduced cooperation with law enforcement and a general sense of disenfranchisement among affected populations. The damage to public trust is not easily reversible, and the long-term consequences of relying on flawed technology in such a sensitive context must be carefully considered.

The COMPAS algorithm, while intended to assist in making data-driven decisions, presents significant risks due to its racial bias, limited predictive accuracy, and potential ethical violations. The judicial system has an ethical duty to uphold fairness, impartiality, and respect for individuals' rights. These principles cannot be met by incorporating a tool that disproportionately harms marginalized groups and fails to provide reliable predictions. Therefore, I recommend that COMPAS should not be used in the parole decision-making process, and alternative methods that are transparent, free from racial bias, and capable of individualized assessments should be sought to support judicial decision-making.

In conclusion, while the use of algorithms in the criminal justice system has the potential to improve efficiency and objectivity, the shortcomings of the COMPAS tool make it unsuitable for such critical decisions. The stakes in parole decisions are too high to rely on a system that lacks transparency and fairness. Moving forward, it is crucial to invest in the development of risk assessment tools that are not only statistically robust but also ethically sound, ensuring that all defendants are given a fair chance at justice.