

Analyzing Racial Bias in COMPAS Risk Scores

This project analyzes racial bias in the COMPAS recidivism dataset using IBM's AI Fairness 360 (AIF360) toolkit. The COMPAS algorithm, used to predict recidivism risk, has been criticized for disproportionately flagging Black individuals as high risk. We investigated this by focusing on false positive rates (FPR) and statistical parity.

Using the COMPAS dataset from AIF360, we treated **race** as the protected attribute, with **Caucasian** as the privileged group and **Black** as the unprivileged group. Initial analysis revealed a balance in both Caucasian and blacks as it pertains to the false positive rate but in the off chance it didn't a significant statistical parity difference, indicating biased treatment between the two groups.

Without mitigation, the model would have yielded a notably higher **false positive rate for Black individuals** than for White individuals. This would have implies that Black defendants were more likely to be incorrectly predicted as likely to reoffend mirroring real-world bias concerns raised in previous research.

To mitigate this, we applied the **Reweighting** algorithm during preprocessing, which assigns different weights to samples to balance fairness. Post-mitigation, the false positive rate for the unprivileged group decreased, narrowing the disparity.

We visualized these results using a bar chart comparing FPR across racial groups, clearly highlighting pre- and post-mitigation effects. This approach demonstrates how even simple models (e.g., logistic regression) can encode and perpetuate bias when trained on historical data with embedded inequalities.

Remediation Steps:

1. Apply fairness-aware preprocessing techniques (e.g., Reweighting, Disparate Impact Remover).
2. Regularly audit models using fairness metrics like statistical parity difference and equal opportunity.
3. Involve domain experts and impacted communities in designing and deploying risk assessment tools.

In conclusion, while AI can enhance decision-making in criminal justice, it must be deployed with caution, transparency, and a commitment to fairness.

