## COMPAS Dataset Bias Audit Report

This report summarizes a practical audit of the COMPAS dataset, focusing on racial bias in recidivism prediction. The analysis specifically targeted the false positive rates (FPR) among African-American and Caucasian individuals. False positives occur when a person is incorrectly predicted to reoffend within two years but does not. A high FPR in a specific group can suggest algorithmic bias.

We loaded the dataset using Python and filtered it to remove incomplete or irrelevant records. The analysis retained only African-American and Caucasian individuals to ensure statistical comparison. We then converted risk scores into binary predictions assigning '1' to 'High' risk scores and '0' otherwise.

Our results showed that African-American defendants had a significantly higher false positive rate compared to Caucasians. This indicates the model was more likely to incorrectly flag African-Americans as high risk even when they did not reoffend. The disparity was visualized using a bar chart, reinforcing the concern of racial imbalance.

Several factors may contribute to this issue, including biased historical data and a lack of fairness constraints in the model design. Although race is not explicitly used as an input, proxies for race can indirectly influence predictions.

To address these issues, we recommend the following steps: applying fairness-aware modeling techniques, conducting routine audits, and involving community stakeholders in the development process. Transparency in how risk scores are computed and used in decision-making should also be a requirement.

This audit highlights the ethical implications of deploying AI systems in sensitive domains such as criminal justice. Biases in algorithmic outcomes can have life-altering consequences. Developers and decision-makers must commit to ongoing fairness evaluations to build trustworthy and equitable AI.