BiasGuard - AI Fairness Report

Model Overview

This report analyzes bias in the LogisticRegression with respect to sensitive attribute: Gender.

Fairness Metrics

Accuracy: 0.8000

Demographic Parity Diff: 0.0500

Equal Opportunity Diff: 0.6100

Disparate Impact: 0.9400

Selection Rates by Gender

Gender Group	Selection Rate
0	0.86
1	0.81

Interpretation

The model achieved an accuracy of 0.80. However, high accuracy does not imply fairness, as the model may be systematically favoring one group over another.

Disparate Impact is 0.94, which lies within the fairness threshold. This suggests that the model is treating groups more equitably in terms of overall outcomes.

Equal Opportunity Difference is 0.61, which exceeds the fairness tolerance. This means the model provides different true positive rates across groups, disadvantaging some.

Demographic Parity Difference is 0.05. A higher value indicates unequal positive prediction rates between groups, a key sign of potential bias.

Fairness Summary

The highest selection rate among groups is 0.86, while the lowest is 0.81. This results in a gap of 0.05, indicating minor bias in favor of the higher-rated group.

Selection Rates Chart

Selection Rates by Group