

BiasGuard

AI Fairness Compliance Report

Report ID:	RPT-model_2d5eff13694f-20260224
Generated:	February 24, 2026 at 02:51 PM
Model ID:	model_2d5eff13694f
Model Type:	sklearn
Compliance Status:	WARNING - Intersectional bias detected

Executive Summary

The AI fairness compliance report for the sklearn model, which achieved an accuracy of 74.98%, indicates a compliance status of "WARNING - Intersectional bias detected". The analysis, based on a sample size of 10,000, highlights potential disparities in prediction outcomes across sensitive attributes, including race, gender, age.

Key metrics reveal that while the statistical parity difference for race is 0.2154, indicating potential bias, the equal opportunity metric shows a significant disparity in true positive rates between groups. The disparate impact ratio of 0.4683 does not meet the CFPB 80% rule threshold.

Current assessment indicates high risk of discriminatory outcomes. Action recommended: immediate attention required.

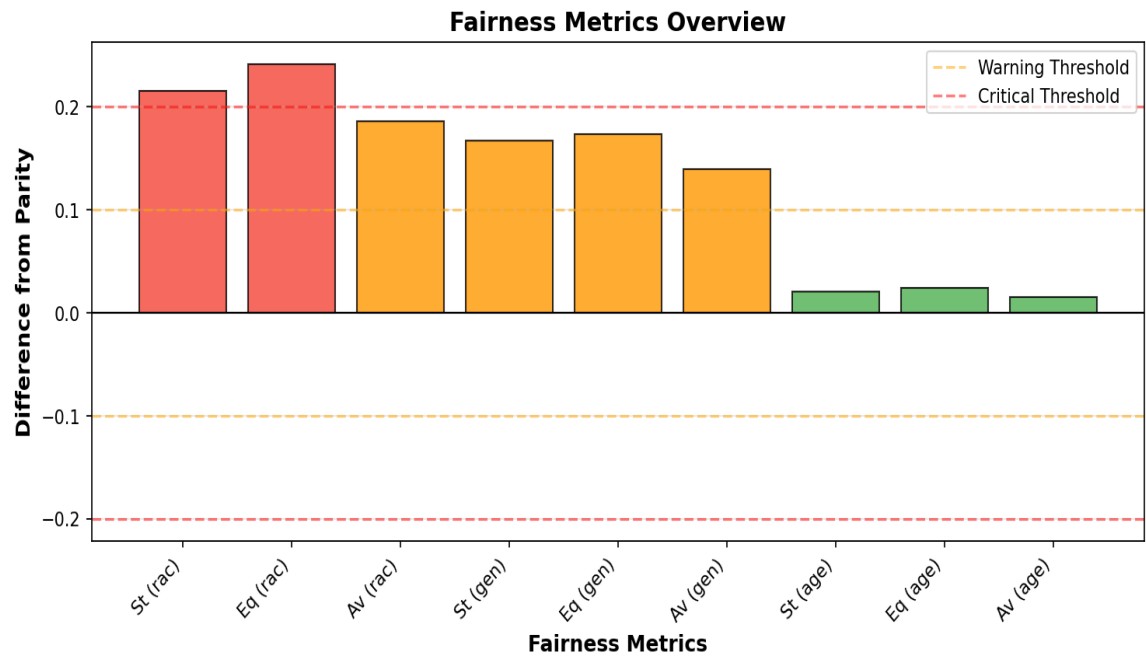
To address these concerns, it is recommended that the model undergoes further refinement to reduce bias. This may include implementing bias mitigation strategies, such as re-weighting training samples or incorporating fairness constraints during model training. Additionally, ongoing monitoring of fairness metrics is essential to ensure compliance with regulatory expectations and to foster equitable outcomes across all demographic groups.

Compliance Status: NON-COMPLIANT

Analyzed: 2026-02-24 19:49:07

Fairness Metrics

Metric	Sensitive Attr	Value	Threshold	Status
Statistical Parity	race	0.2154	0.1	HIGH
Equal Opportunity	race	0.2407	0.1	HIGH
Disparate Impact	race	0.4683	[0.8, 1.2]	HIGH
Average Odds	race	0.1860	0.1	MODERATE
Fpr Parity	race	0.1313	0.1	MODERATE
Predictive Parity	race	0.0521	0.1	LOW
Treatment Equality	race	3.6375	1.0	HIGH
Statistical Parity	gender	0.1672	0.1	MODERATE
Equal Opportunity	gender	0.1732	0.1	MODERATE
Disparate Impact	gender	1.7063	[0.8, 1.2]	HIGH
Average Odds	gender	0.1388	0.1	MODERATE
Fpr Parity	gender	0.1043	0.1	MODERATE
Predictive Parity	gender	0.0286	0.1	LOW
Treatment Equality	gender	1.9907	1.0	MODERATE
Statistical Parity	age	0.0207	0.1	LOW
Equal Opportunity	age	0.0236	0.1	LOW
Disparate Impact	age	0.9371	[0.8, 1.2]	LOW
Average Odds	age	0.0148	0.1	LOW
Fpr Parity	age	0.0060	0.1	LOW
Predictive Parity	age	0.0123	0.1	LOW
Treatment Equality	age	0.0676	1.0	LOW



Intersectionality Analysis

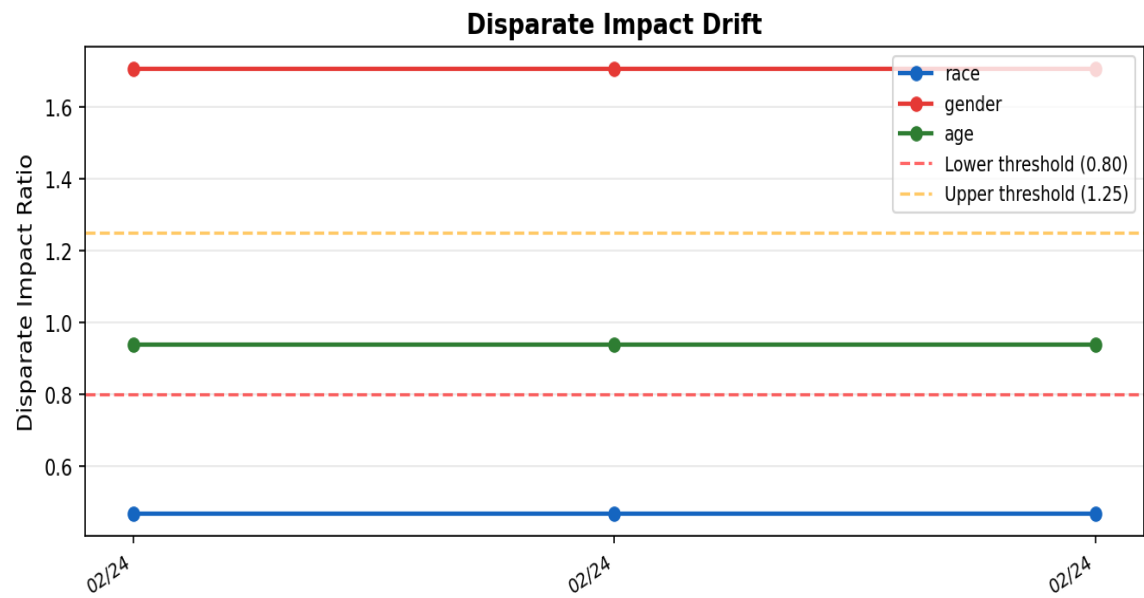
Metric	Value
Total intersectional groups analyzed	126
Groups below DI threshold (< 0.8)	33
Groups above threshold	43
Minimum disparate impact	0.134
Maximum disparate impact	2.081

Highest Risk Intersectional Groups (Top 10)

Group	Disparate Impact	Sample Size	Risk Level
race=Hispanic, age=18	0.134	70	CRITICAL
race=Black, gender=Female	0.142	397	CRITICAL
race=Black, age=18	0.209	60	CRITICAL
race=Hispanic, gender=Female	0.267	458	CRITICAL
gender=Female, age=21	0.299	42	CRITICAL

Bias Drift Over Time

Disparate Impact ratio across the last 3 analysis runs, per sensitive attribute.



Run Date	race	gender	age
02/24	0.468	1.706	0.937
02/24	0.468	1.706	0.937
02/24	0.468	1.706	0.937

Recommendations

1. **[CRITICAL] Severe bias for race: DI=0.468. Immediate model review required. Consider retraining with fairness constraints.**
2. [HIGH] Large statistical parity gap for race: 0.215. Review model features for proxy discrimination.
3. [HIGH] Large statistical parity gap for gender: 0.167. Review model features for proxy discrimination.
4. [HIGH] 33 intersectional groups below threshold. Review compounding discrimination effects.
5. **[CRITICAL] Worst group: race=Hispanic, age=18 (DI=0.13). Conduct targeted bias audit.**

Regulatory Compliance Checklist

Regulation	Requirement	Status	Notes
CFPB AI/ML Guidelines	Regular bias audits	FAIL	Disparate impact ratio 0.468 outside 0.8-1.25 range
ECOA (80% Rule)	No disparate impact	FAIL	80% rule violation
Title VII	No discrimination	FAIL	Statistical parity difference 0.215 exceeds 0.1
Explainability	Adverse action notices	FAIL	No MLflow tracking

Technical Details

Model Architecture:	sklearn
Task Type:	binary
Training Samples:	5,000
Test Samples:	5,000
Accuracy:	74.98%
Target Variable:	prediction
Protected Attributes:	race, gender, age
MLflow Run ID:	N/A

This report was automatically generated by BiasGuard AI Fairness Platform
For questions or support, contact: compliance@biasguard.ai
© 2025 BiasGuard. All rights reserved.