

Ethical AI Auditor

Compliance & Audit Report

Model Information

Model: hiring_model_v1

Version: 1.0

Generated On: 2025-08-06 03:35:28

Record Count: 4

1. Bias Detection Results (FairML Metrics)

Disparate Impact Ratio: 0.717 (NON-COMPLIANT)

Statistical Parity Difference: -0.170

Equal Opportunity Difference: 0.000

2. Explainability Summary (SAP Joule Alternative)

This model predicts hiring outcomes based on factors like experience and education, but concerningly, gender is also a highly influential feature. The model negatively correlates gender with hiring, creating significant business risk due to potential discrimination lawsuits and reputational damage. Given that gender is a top-ranking feature, this model presents a compliance risk under EEOC, GDPR, and the upcoming EU AI Act, requiring immediate investigation and remediation to avoid legal and ethical liabilities.

3. Regulatory Compliance Status

EEOC: NON-COMPLIANT | 80% Rule (Disparate Impact < 0.8) | Disparate Impact Ratio: 0.717

GDPR: NON-COMPLIANT | Article 22 - Right to Explanation | Explainability Summary Provided + Fairness (DI >= 0.8)

EU AI Act: NON-COMPLIANT | High-Risk AI - Bias Thresholds | DI: 0.717, SPD: 0.170

4. Recommended Actions

Bias detected: Disparate impact ratio (0.72) below 0.8 threshold (EEOC 80% rule)

Recommendation: Retrain model with fairness constraints

Recommendation: Use synthetic data balancing for underrepresented groups

Recommendation: Adjust decision threshold for fairness