

Ethical AI Auditor

Compliance & Audit Report

Model Information

Model: hiring_model_v1

Version: 1.0

Generated On: 2025-08-06 19:14:46

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 primarily on experience, gender, and education. The significant influence of 'gender' (negative correlation) raises immediate compliance concerns under EEOC/GDPR/EU AI Act, posing a high risk to brand reputation and potential regulatory penalties. We must immediately investigate and mitigate this bias to avoid legal action and reputational damage.

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