Bias Audit Report: Ethics Statement

This audit of a synthetic employment prediction dataset (income > R50,000/year) reveals how AI can deepen South Africa's inequalities if left unchecked. With a 33.2% unemployment rate [5] and a Gini coefficient of 0.63 the world's highest [0] biased AI risks entrenching apartheid's legacy. Our analysis shows females face a 16% lower employment prediction rate (0.557 vs. 0.715 for males) and non-Whites a 25% gap (0.594 vs. 0.845 for Whites) [12]. These disparities, validated by chi-squared tests ($p<0.05$), reflect real-world inequities, like unequal access to education, that AI models can amplify if data isn't carefully curated [6].

Ethically, deploying biased AI in hiring is a non-starter. It's not just about numbers; it's about people's livelihoods. Denying jobs to qualified women or non-Whites because of skewed algorithms violates the Employment Equity Act and undermines BEE goals, which aim to level the playing field [6]. Such discrimination fuels poverty and social unrest, a dangerous prospect in a country already battling a 0.63 Gini [0]. Drawing from global frameworks like NIST's AI risk management, fairness isn't optional, it's a moral and legal must [11]. Our mitigations removing protected attributes and reweighing data pushed Disparate Impact closer to 1 (from 0.696-0.788 to 0.989-1.043), though accuracy dipped slightly (67.75% to 63.50-65.00%) [9]. This trade-off prioritizes equity, aligning with utilitarian principles: the greatest good for SA's diverse workforce [10].

Transparency is non-negotiable. Hiding how models make decisions erodes trust, as seen when global firms like Amazon scrapped biased hiring tools [12]. Our open notebook, shared on GitHub, lays bare the code, data, and visuals, inviting scrutiny from SA stakeholders like unions or policymakers [13]. Inclusivity matters too developing AI without diverse voices risks blind spots, especially in a nation shaped by historical divides [17]. Per SA's AI Policy Framework, we need teams reflecting our rainbow nation to catch these issues early [13].

Accountability seals the deal. Developers must own biased outcomes, not shrug and blame the data. Regular audits, as mandated by emerging SA regulations, ensure models don't drift back to unfairness [14]. Our recommendations, balancing datasets, using skills-based features, and conducting annual fairness checks embody non-maleficence (do no harm) and beneficence (promote good) [16]. For SA, where unemployment and inequality bite hard, ethical AI is not just a buzzword, it's a lifeline to break poverty cycles and honor BEE's spirit [6]. Policymakers should lean on global models like the EU AI Act, tailoring

them to SA's unique challenges [11]. This audit urges a commitment to fairness, ensuring AI uplifts rather than divides Mzansi's workforce.