TalentMatch AI

HireScore Bias Investigation Report

Executive Summary

Bias Severity: High

Primary Affected Groups: Female candidates - Candidates outside Greater Accra and Ashanti regi... Candidates from smaller or less represented universities

Top 3 Recommended Actions: 1. Rebalance training data representation 2. Introduce fairness constraints during model training 3. Add human review for top-ranked candidate decisions

---

Task 1: Bias Type Analysis

1. Historical Bias

Present? Yes **Evidence:** - Past hiring data shows male hires = 72% vs female hires = 2... Software engineering roles dominate training data (6000/10000)

Disadvantaged Groups: - Female applicants - Non-software applicants

---

2. Sampling Bias

Present? Yes **Evidence:** - Regional imbalance (Accra + Ashanti = 8500 hires vs other regions... 1500) - University representation is skewed toward top universities

Underrepresented Groups: - Candidates from Northern, Volta, Upper regions - Candidates from sm... universities

---

3. Measurement Bias

Present? Yes **Evidence:** - Age derived from graduation year may disadvantage career changers - Li... connections count may disadvantage newer professionals

Unfair Measurements: - Social network size does not equal job competence - Extracurricular activitie... be unequally accessible

---

4. Proxy Bias

Present? Yes

Proxy Features: - University name (proxy for socioeconomic status) - Location of previous emplo[...] (proxy for region/ethnicity) - LinkedIn connections (proxy for social capital)

Impact: These features can reinforce existing regional and gender inequalities.

Task 2: Bias Pipeline Mapping

[Historical Hiring Decisions] ↓ [Training Data Collection] → Bias Point #1: Historical inequality re[...] in hiring records ↓ [Feature Selection] → Bias Point #2: Use of demographic proxy features [...] Training] → Bias Point #3: Model learns biased patterns from skewed data ↓ [Deployment & [...] → Bias Point #4: Unequal ranking of candidates ↓ [Biased Outcomes]

Task 3: Mitigation Strategies

Immediate Actions

Feature Removal: - Remove or de-emphasize: - University name - Region of origin - Age de[...] graduation year

Threshold Adjustments: - Use group-aware thresholding to ensure equal top-ranking representation

Weekly Monitoring Metrics: Gender representation in top 100 - Regional distribution of scores [...] Score variance between groups

---

Short-Term Actions

New Data Collection: Gather more applications from underrepresented regions - Collect role-speci[...] performance outcomes

Model Retraining: - Introduce fairness regularization - Use stratified sampling

Human Oversight: - Review top 20% of ranked candidates manually

---

Long-Term Actions

Fairness Metric Choice: - Equal opportunity fairness is recommended

Process Changes: - Provide clients with fairness audit reports

Transparency: - Disclose scoring criteria at high level to applicants

---

Success Metrics

- Female representation in top rankings increases
- Regional representation improves
- No single group exceeds 60% dominance in top candidates

Timeline

- Week 1: Feature audit
- 1 - 3 months: Retraining and data collection
- 6 - 1 2 months: Governance and policy changes