

Bias Detection Report

Language Bias Analysis Report

Dataset Analyzed: COMPAS dataset, focusing on the 'Language' column.

Features Examined: Language

Bias Type: Distribution

Detection and Analysis Methods Used:

- Shannon Entropy and Balance Metric
- Max/Min Ratio
- Gini Index
- Relative Risk

Values Obtained and Extent of Bias:

- ****Balance****: 0.0387
- ****Max/Min Ratio****: 240.44

- **Corrected Gini Index**: 0.00828
- **Relative Risk (English)**: 1.99, **(Spanish)**: 0.00828
- **Normalized Bias Score**: 240.44

Bias Level: Extreme Bias

Detailed Summary:

The normalized bias score and max/min ratio both indicate a significant dominance of English speakers, with extreme disparities compared to Spanish speakers. The Shannon Entropy and Balance Metric also reflect a highly unequal distribution. The Gini Index, while low, supports the presence of inequality in language distribution.

Recommendations:

Due to the extreme bias observed, it is advisable to consider the implications of this bias when using this dataset for any analysis or decision-making. Efforts to mitigate this bias may be necessary to ensure fair representation.

