Bias Detection Report

Bias Detection Report

Dataset Analyzed: COMPAS

Feature Examined: Ethnic_Code_Text

Bias Type Detected: Distribution

Tools and Methods Used

1. Shannon Entropy and Balance Metric

2. Max/Min Ratio

3. Gini Index

4. Chi-Square Test

Analysis and Results

- **Shannon Entropy and Balance Metric**: Balance = 0.550, Shannon Entropy = 1.743 (Significant Bias)

- **Max/Min Ratio**: 692.77 (Extreme Bias)

- **Gini Index**: Corrected Gini Index = 0.652, Adjusted Gini Index = 0.734 (Significant Bias)

- **Chi-Square Test**: Chi-Square Statistic = 7980.72, p-value = 0.0 (Extreme Bias)

Conclusion

The `Ethnic_Code_Text` feature exhibits a high level of distribution bias. The bias levels range from significant to extreme across different methods, indicating a pronounced imbalance in the distribution of ethnic categories.

Recommendations

The dataset requires careful consideration for applications sensitive to ethnic distribution bias. Mitigation strategies or adjustments may be necessary to address these biases.

Pie Chart of Ethnic_Code_Text

African-American 44.4%

Caucasian 35.8%

Hispanic 14.4%