

Bias Detection Report

Dataset Analyzed

- **Dataset Path:** source_files/COMPAS.csv
- **Relevant Feature:** MaritalStatus

Types of Bias Detected

- **Distribution Bias:** Assessed using multiple methods to understand the distribution of "MaritalStatus."

Tools and Methods Used

1. **Shannon Entropy and Balance Metric:**

- **Balance:** 0.4643
- **Shannon Entropy:** 1.3035

2. **Max/Min Ratio:**

- **Ratio:** 211.86

3. **Gini Index:**

- **Corrected Gini Index:** 0.4265

- **Adjusted Gini Index:** 0.4976

4. **Jensen-Shannon Divergence** (from literature):

- **JSD Value:** 0.4956

Visualization

- **Pie Chart:** Successfully created for "MaritalStatus" (path: generated_files/marital_status_pie_chart.png).

Detailed Interpretation and Recommendations

1. **Shannon Entropy and Balance Metric:**

- The balance metric of 0.4643 indicates a significant bias level. A lower balance suggests uneven distribution among categories.

- Shannon Entropy of 1.3035 aligns with this finding, showing moderate uncertainty in category distribution.

2. **Max/Min Ratio:**

- The ratio of 211.86 falls into a high bias category, indicating extreme disparities between the most and least frequent categories.

3. **Gini Index**:

- Both Gini indices suggest significant inequality in distribution.

4. **Jensen-Shannon Divergence**:

- A JSD value of 0.4956 also points to significant distribution bias, confirming the findings from other methods.

Recommendations

- Based on the analysis, "MaritalStatus" shows significant distribution bias. The dataset might require adjustments or considerations when used in analysis or modeling, especially in contexts where equal representation of marital statuses is crucial.
- The visual representation via the pie chart provides a clear view of the distribution, supporting the analytical findings.

Overall Bias Level:

- The findings consistently indicate a significant bias level across the tools used.

Pie Chart of MaritalStatus

