Bias Detection Report

Bias Detection Report for Age Feature

This report provides a detailed analysis of the distribution bias in the 'Age' feature from the COMPAS dataset.

Dataset Analyzed

- **Source File**: COMPAS.csv

- **Derived Feature**: Age

The 'Age' feature was derived from 'DateOfBirth' and 'Screening_Date' columns.

Bias Type: Distribution

Relevant Feature(s): Age

Bias Analysis Methods and Results:

1. **Skewness Analysis**:

   - Skewness value: -0.902

   - Indicates moderate skewness.

2. **Kurtosis Analysis**:

   - Kurtosis value: -1.039

   - Suggests a flatter distribution.

3. **Z-score Outlier Detection**:

   - Outlier percentage: 0.0%

   - No significant outliers.

4. **Cohen's d Analysis**:

   - Cohen's d value: -2.287

   - Significant deviation.

5. **Quantile Deviation**:

   - Quantile deviation value: 0.085

   - Minimal deviation.

6. **Max/Min Ratio**:

- Ratio: Infinite

  - Indicates extreme bias due to zero frequency bins.

7. **Shannon Entropy**:

  - Entropy: 1.046

  - Max Entropy: 3.332

  - Significant deviation from maximum entropy.

8. **Gini Coefficient**:

  - Value: -292573.236

  - Indicates calculation error or data issue.

Overall Bias Level:

Severe distribution bias indicated by extreme Max/Min Ratio and significant entropy deviation.

Recommendations:

- Re-evaluate binning strategy.

- Verify data integrity.

- Consider data transformations to balance distribution.

Bar Chart of Age

Box Plot of Age grouped by Age