Bias Detection Report

Bias Detection Report

Dataset Analyzed

- **File**: MIMIC-IV.csv

Features Examined

- **Feature**: patient_lang

Bias Type

- **Type**: Distribution

Tools Used

- Shannon Entropy and Balance Metric

- Max/Min Ratio

- Normalized Entropy

Analysis and Results

- **Shannon Entropy and Balance Metric**: Approximately 0.486

- **Max/Min Ratio**: Approximately 8.48

- **Normalized Entropy**: Approximately 0.486

Visualizations
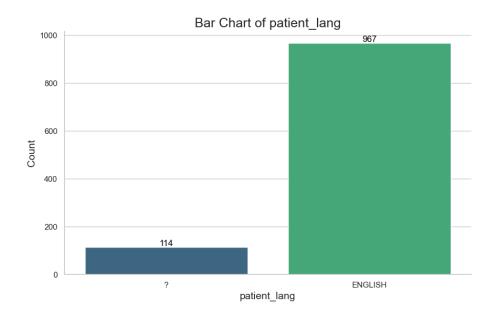
- Bar Chart of Patient Language Distribution

- Pie Chart of Patient Language Distribution

Interpretation of Bias Severity

The distribution of patient languages in this dataset shows a significant bias, with one or more languages predominating. This suggests that the dataset may not be representative of a diverse linguistic population.

Recommendations

Consider adjusting the dataset or acquiring additional data to better represent underrepresented languages.

Bar Chart of patient_lang

Pie Chart of patient_lang