Bias Detection Report

Bias Detection Report for Patient Insurance in MIMIC-IV Dataset

Analysis Overview

This report analyzes the distribution bias of the 'patient_insurance' feature in the MIMIC-IV dataset.

The goal is to determine if all insurance types are well-represented or if there's an overrepresentation of a particular type.

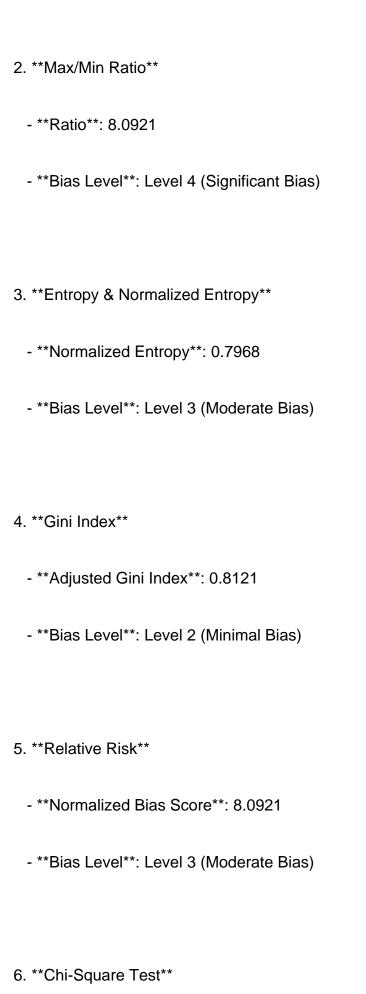
Bias Type: Distribution

- **Relevant Feature**: patient_insurance

Detection and Analysis Methods

Multiple methods were employed to analyze the distribution bias:

- 1. **Shannon Entropy & Balance Metric**
 - **Balance**: 0.7968
 - **Shannon Entropy**: 1.2629
 - **Bias Level**: Level 3 (Moderate Bias)



- **Chi-Square Statistic**: 252.9520

- **p-value**: 1.18e-55

Conclusion

The analysis indicates a moderate to significant level of bias in the 'patient_insurance' feature.

Several insurance types are overrepresented, as evidenced by the Max/Min Ratio and Chi-Square

Test results.

Recommendations

- Consider addressing insurance type representation in the dataset.

- Use visualizations like bar charts or pie charts to present the distribution clearly.

Image not found: generated_files/bar_chart_insurance.png

Image not found: generated_files/pie_chart_insurance.png