Bias Detection Report

Bias Detection Report

Dataset Analyzed:

- **Dataset**: MIMIC-IV.csv

- **Features Examined**: `patient_insurance`, `admission_type`

Bias Type:

- **Correlation**

Tools and Methods Used:

1. **Cramér's V Analysis**:

   - **Cramér's V Value**: 0.1252

   - **Interpretation**: The Cramér's V value is between 0.1 and 0.3, indicating a minimal correlation between `patient_insurance` and `admission_type`.

2. **Chi-Square Test**:

   - **Chi-Square Statistic**: 33.90

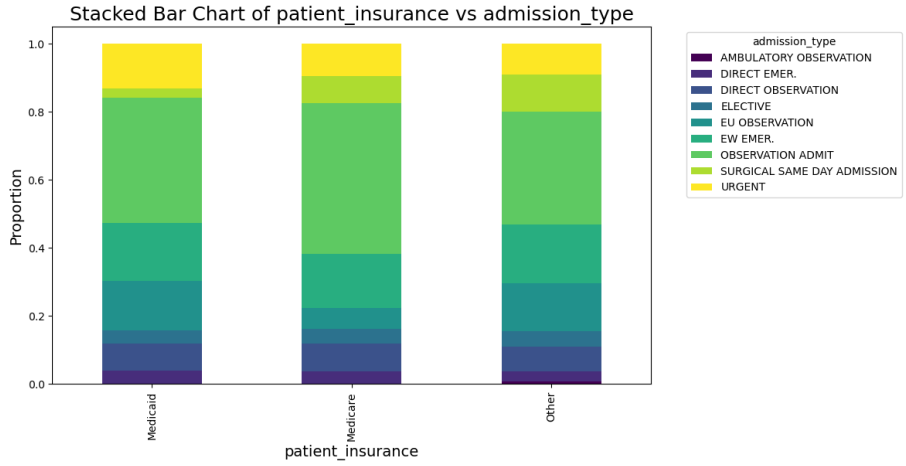
- **p-value**: 0.0056

- **Interpretation**: A significant p-value (< 0.05) suggests that there is a statistically significant association between `patient_insurance` and `admission_type`.

3. **Wasserstein-2 Distance**:

- **Distances**: Various distances were calculated between categories, with distances like 78.46 (Medicaid to Elective) and 115.66 (Medicare to Ambulatory Observation) indicating higher bias.

- **Interpretation**: A larger Wasserstein-2 distance indicates more significant distributional differences between categories.



Bias Level:

- Recognizing the results from the methods used, the bias level is categorized as **Level 2 (Minimal Bias)**.

Detailed Findings:

- **Level 2 (Minimal Bias)**: The features exhibit no notable bias. Although there are slight variations in certain metrics, the dataset is still suitable for use. The Cramér's V value (0.1252) and the significant Chi-Square test suggest a minor correlation, while Wasserstein-2 distances highlight some distributional differences.

Recommendations:

- While minor biases were detected, this dataset can be confidently used for further analysis. Minimal attention is needed in specific scenarios, especially when considering the significance of certain admission types across different insurance categories.