

# Bias Detection Report

## Bias Detection Report

### Dataset Analyzed:

MIMIC-IV (Path: source\_files/MIMIC-IV.csv)

### Features Examined:

- Language Spoken by Patients (`patient\_lang`)
- Mortality Outcome (`hospital\_expire\_flag`)

### Bias Type:

Correlation

### Relevant Feature(s):

- patient\_lang
- hospital\_expire\_flag

### Tools Used:

- Hilbert-Schmidt Independence Criterion (HSIC)
- Cohen's d
- Standardized Difference

### **Values Obtained and Bias Level:**

- **HSIC Value**:  $7.14 \times 10^{-12}$  (No Significant Bias)
- **Cohen's d Effect Size**: 0.1405 (Minimal Bias)
- **Standardized Difference**: Infinity (Significant Distribution Issue)

### **Visualization:**

Due to tool constraints, visualizations were not generated directly. However, the analyses suggest slight correlation bias.

### **Interpretation and Recommendations:**

The feature `patient\_lang` shows a very slight correlation bias with `hospital\_expire\_flag` based on Cohen's d. However, the standardized difference indicates a distribution issue, suggesting that some language categories might be underrepresented.

Overall, the dataset appears suitable for use, but it is advisable to further inspect the distribution of

language categories to ensure balanced representation.