

Bias Detection Report

The dataset analyzed is the Statlog dataset, and the features examined are 'Telephone' and 'Credit risk (Good/Bad)' for correlation bias and 'Age in years' for distribution bias.

Image not found: generated_files/dependency_plot.png

The type of bias detected is correlation bias between 'Telephone' and 'Credit risk (Good/Bad)' and distribution bias in 'Age in years'.

Image not found: generated_files/histogram.png

The tools used are the SHAP method for correlation bias and the skewness method for distribution bias.

The values obtained are a Cramers V of 0.034242641998258334 for correlation bias and a skewness of 1.0207392686768317 for distribution bias.

The extent of the bias is minimal for correlation bias and moderate for distribution bias.

My natural language interpretation of the bias severity is that there is minimal correlation bias between 'Telephone' and 'Credit risk (Good/Bad)' and moderate distribution bias in 'Age in years'.

My recommendations for the user regarding the use of the dataset are to be cautious when using the 'Telephone' feature in models that predict 'Credit risk (Good/Bad)' and to consider transforming the 'Age in years' feature to reduce skewness.