Bias Detection Report

The dataset analyzed is the Statlog dataset, which contains various features related to credit risk.

The features examined are the numerical features in the dataset, including 'Duration in month', 'Credit amount', 'Installment rate in percentage of disposable income', 'Age in years', etc.

The type of bias detected is distribution bias, which refers to the difference in the distribution of a feature across different groups.

The tools used include the max/min ratio method, which calculates the ratio of the maximum frequency to the minimum frequency for each feature.

The values obtained and the extent of the bias are as follows:

- Feature: Duration in month, Bias Level: Minimal Bias

- Feature: Credit amount, Bias Level: Moderate Bias

- Feature: Installment rate in percentage of disposable income, Bias Level: Minimal Bias

- Feature: Age in years, Bias Level: No Bias

The natural language interpretation of the bias severity is as follows:

- The feature 'Duration in month' has a minimal bias, indicating that the distribution of this feature is relatively balanced across different groups.

- The feature 'Credit amount' has a moderate bias, indicating that the distribution of this feature is somewhat imbalanced across different groups.

- The feature 'Installment rate in percentage of disposable income' has a minimal bias, indicating that the distribution of this feature is relatively balanced across different groups.

- The feature 'Age in years' has no bias, indicating that the distribution of this feature is perfectly balanced across different groups.

The recommendations for the user regarding the use of the dataset are as follows:

- The user should be aware of the potential bias in the features 'Credit amount' and take steps to mitigate it, such as data preprocessing or feature engineering.

- The user can use the features 'Duration in month', 'Installment rate in percentage of disposable income', and 'Age in years' without significant concerns about bias.