Analysis Report

Global dataset report

This report is the output of the Amazon SageMaker Clarify analysis. The report is split into following parts:

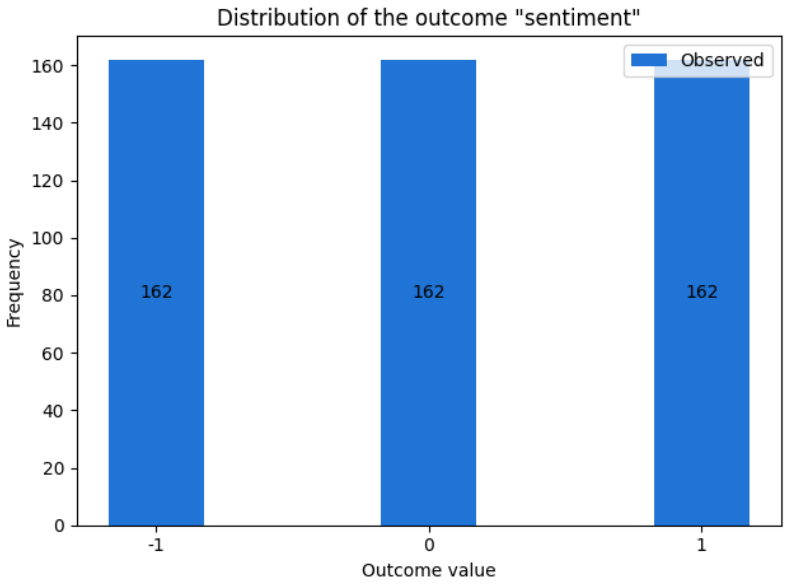
1. Analysis configuration
2. Pretraining bias metrics

Analysis Configuration

Bias analysis requires you to configure the outcome label column, the facet and optionally a group variable. Generating explanations requires you to configure the outcome label. You configured the analysis with the following variables. The complete analysis configuration is appended at the end.

Outcome label: You chose the column `sentiment` in the input data as the outcome label. Bias metric computation requires designating the positive outcome. You chose `sentiment=1` as the positive outcome. `sentiment` consisted of values `[-1, 0, 1]`.

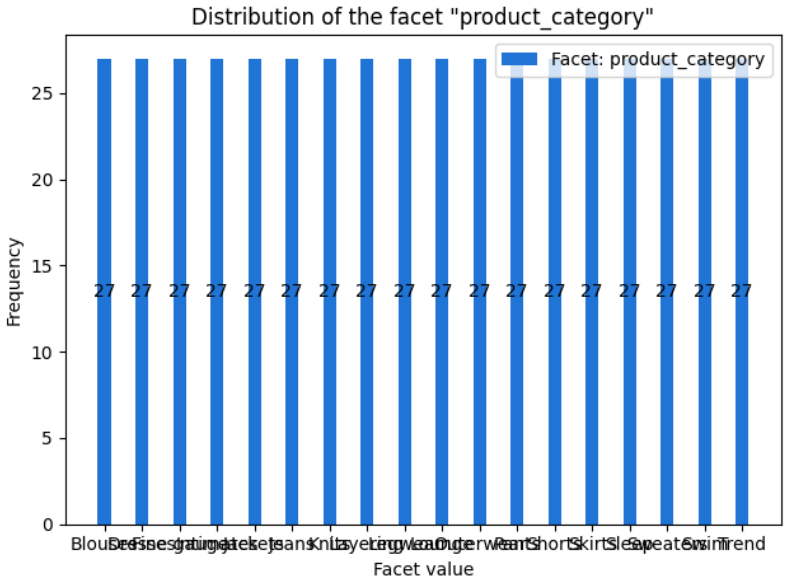
The figure below shows the distribution of values of `sentiment`.



Facet: You chose the column `product_category` in the input data as the facet. `product_category` consisted of values `['Blouses', 'Dresses', 'Fine gauge', 'Intimates', 'Jackets', 'Jeans', 'Knits', 'Layering', 'Legwear', 'Lounge', 'Outerwear', 'Pants', 'Shorts', 'Skirts', 'Sleep', 'Sweaters', 'Swim', 'Trend']`. Bias metrics were computed by comparing the inputs `product_category=Blouses` with all other inputs, then by comparing inputs `product_category=Dresses` with all other inputs, then by comparing inputs `product_category=Fine gauge` with all other inputs, then by comparing inputs `product_category=Intimates` with all other inputs, then by comparing inputs `product_category=Jackets` with all other inputs, then by comparing inputs `product_category=Jeans` with all other inputs, then by comparing inputs `product_category=Knits` with all other inputs, then by comparing inputs `product_category=Layering` with all other inputs, then by comparing inputs `product_category=Legwear` with all other inputs, then by comparing inputs `product_category=Lounge` with all other inputs, then by comparing inputs `product_category=Outerwear` with all other inputs, then by comparing inputs `product_category=Pants` with all other inputs, then by comparing inputs `product_category=Shorts` with all other inputs, then by comparing inputs `product_category=Skirts` with all other inputs, then by comparing inputs `product_category=Sleep` with all other inputs, then by comparing inputs `product_category=Sweaters` with all other inputs, then by comparing inputs `product_category=Swim` with all other inputs, then by comparing inputs `product_category=Trend` with all other

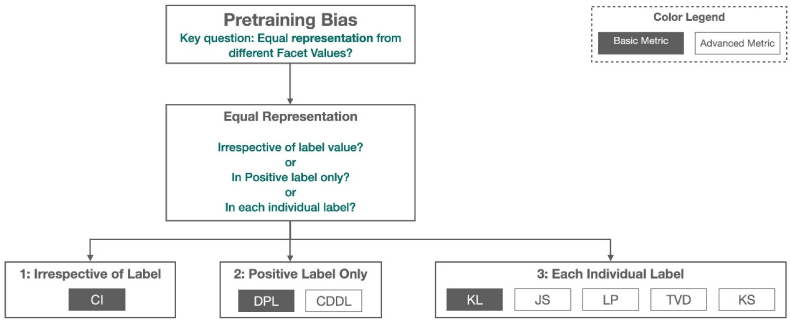
inputs.

The figure below shows the distribution of values of `product_category`.



Pre-training Bias Metrics

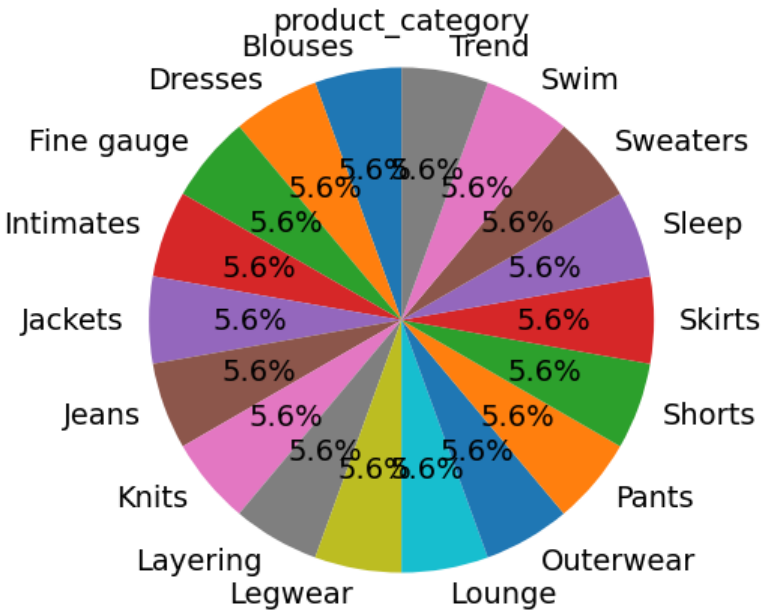
Pretraining bias metrics measure imbalances in facet value representation in the training data. Imbalances can be measured across different dimensions. For instance, you could focus imbalances within the inputs with positive observed label only. The figure below shows how different pretraining bias metrics focus on different dimensions. For a detailed description of these dimensions, see Learn How Amazon SageMaker Clarify Helps Detect Bias.



The metric values along with an informal description of what they mean are shown below. For mathematical formulas and examples, see the Measure Pretraining Bias section of the AWS documentation.

We computed the bias metrics for the label `sentiment` using label value(s)/threshold `1`

- **product_category**
  The groups are represented in the dataset with the following proportions.



Value(s)/Threshold: Blouses

name	description	value
CI	Class Imbalance (CI)	0.888889
DPL	Difference in Positive Proportions in Labels (DPL)	0.0
JS	Jensen-Shannon Divergence (JS)	0.0
KL	Kullback-Liebler Divergence (KL)	0.0
KS	Kolmogorov-Smirnov Distance (KS)	0.0
LP	L-p Norm (LP)	0.0
TVD	Total Variation Distance (TVD)	0.0

Value(s)/Threshold: Dresses

name	description	value
CI	Class Imbalance (CI)	0.888889
DPL	Difference in Positive Proportions in Labels (DPL)	0.0
JS	Jensen-Shannon Divergence (JS)	0.0
KL	Kullback-Liebler Divergence (KL)	0.0
KS	Kolmogorov-Smirnov Distance (KS)	0.0
LP	L-p Norm (LP)	0.0
TVD	Total Variation Distance (TVD)	0.0

Value(s)/Threshold: Fine gauge

name	description	value
CI	Class Imbalance (CI)	0.888889
DPL	Difference in Positive Proportions in Labels (DPL)	0.0
JS	Jensen-Shannon Divergence (JS)	0.0
KL	Kullback-Liebler Divergence (KL)	0.0
KS	Kolmogorov-Smirnov Distance (KS)	0.0
LP	L-p Norm (LP)	0.0
TVD	Total Variation Distance (TVD)	0.0

Value(s)/Threshold: Intimates

name	description	value
CI	Class Imbalance (CI)	0.888889
DPL	Difference in Positive Proportions in Labels (DPL)	0.0
JS	Jensen-Shannon Divergence (JS)	0.0
KL	Kullback-Liebler Divergence (KL)	0.0
KS	Kolmogorov-Smirnov Distance (KS)	0.0
LP	L-p Norm (LP)	0.0
TVD	Total Variation Distance (TVD)	0.0

Value(s)/Threshold: Jackets

name	description	value
CI	Class Imbalance (CI)	0.888889
DPL	Difference in Positive Proportions in Labels (DPL)	0.0
JS	Jensen-Shannon Divergence (JS)	0.0
KL	Kullback-Liebler Divergence (KL)	0.0
KS	Kolmogorov-Smirnov Distance (KS)	0.0
LP	L-p Norm (LP)	0.0
TVD	Total Variation Distance (TVD)	0.0

Value(s)/Threshold: Jeans

name	description	value
CI	Class Imbalance (CI)	0.888889
DPL	Difference in Positive Proportions in Labels (DPL)	0.0
JS	Jensen-Shannon Divergence (JS)	0.0
KL	Kullback-Liebler Divergence (KL)	0.0
KS	Kolmogorov-Smirnov Distance (KS)	0.0
LP	L-p Norm (LP)	0.0
TVD	Total Variation Distance (TVD)	0.0

Value(s)/Threshold: Knits

name	description	value
CI	Class Imbalance (CI)	0.888889
DPL	Difference in Positive Proportions in Labels (DPL)	0.0
JS	Jensen-Shannon Divergence (JS)	0.0
KL	Kullback-Liebler Divergence (KL)	0.0
KS	Kolmogorov-Smirnov Distance (KS)	0.0
LP	L-p Norm (LP)	0.0
TVD	Total Variation Distance (TVD)	0.0

Value(s)/Threshold: Layering

name	description	value
CI	Class Imbalance (CI)	0.888889
DPL	Difference in Positive Proportions in Labels (DPL)	0.0
JS	Jensen-Shannon Divergence (JS)	0.0
KL	Kullback-Liebler Divergence (KL)	0.0
KS	Kolmogorov-Smirnov Distance (KS)	0.0
LP	L-p Norm (LP)	0.0
TVD	Total Variation Distance (TVD)	0.0

Value(s)/Threshold: Legwear

name	description	value
CI	Class Imbalance (CI)	0.888889
DPL	Difference in Positive Proportions in Labels (DPL)	0.0
JS	Jensen-Shannon Divergence (JS)	0.0
KL	Kullback-Liebler Divergence (KL)	0.0
KS	Kolmogorov-Smirnov Distance (KS)	0.0
LP	L-p Norm (LP)	0.0
TVD	Total Variation Distance (TVD)	0.0

Value(s)/Threshold: Lounge

name	description	value
CI	Class Imbalance (CI)	0.888889
DPL	Difference in Positive Proportions in Labels (DPL)	0.0
JS	Jensen-Shannon Divergence (JS)	0.0
KL	Kullback-Liebler Divergence (KL)	0.0
KS	Kolmogorov-Smirnov Distance (KS)	0.0
LP	L-p Norm (LP)	0.0
TVD	Total Variation Distance (TVD)	0.0

Value(s)/Threshold: Outerwear

name	description	value
CI	Class Imbalance (CI)	0.888889
DPL	Difference in Positive Proportions in Labels (DPL)	0.0
JS	Jensen-Shannon Divergence (JS)	0.0
KL	Kullback-Liebler Divergence (KL)	0.0
KS	Kolmogorov-Smirnov Distance (KS)	0.0
LP	L-p Norm (LP)	0.0
TVD	Total Variation Distance (TVD)	0.0

Value(s)/Threshold: Pants

name	description	value
CI	Class Imbalance (CI)	0.888889
DPL	Difference in Positive Proportions in Labels (DPL)	0.0
JS	Jensen-Shannon Divergence (JS)	0.0
KL	Kullback-Liebler Divergence (KL)	0.0
KS	Kolmogorov-Smirnov Distance (KS)	0.0
LP	L-p Norm (LP)	0.0
TVD	Total Variation Distance (TVD)	0.0

Value(s)/Threshold: Shorts

name	description	value
CI	Class Imbalance (CI)	0.888889
DPL	Difference in Positive Proportions in Labels (DPL)	0.0
JS	Jensen-Shannon Divergence (JS)	0.0
KL	Kullback-Liebler Divergence (KL)	0.0
KS	Kolmogorov-Smirnov Distance (KS)	0.0
LP	L-p Norm (LP)	0.0
TVD	Total Variation Distance (TVD)	0.0

Value(s)/Threshold: Skirts

name	description	value
CI	Class Imbalance (CI)	0.888889
DPL	Difference in Positive Proportions in Labels (DPL)	0.0
JS	Jensen-Shannon Divergence (JS)	0.0
KL	Kullback-Liebler Divergence (KL)	0.0
KS	Kolmogorov-Smirnov Distance (KS)	0.0
LP	L-p Norm (LP)	0.0
TVD	Total Variation Distance (TVD)	0.0

Value(s)/Threshold: Sleep

name	description	value
CI	Class Imbalance (CI)	0.888889
DPL	Difference in Positive Proportions in Labels (DPL)	0.0
JS	Jensen-Shannon Divergence (JS)	0.0
KL	Kullback-Liebler Divergence (KL)	0.0
KS	Kolmogorov-Smirnov Distance (KS)	0.0
LP	L-p Norm (LP)	0.0
TVD	Total Variation Distance (TVD)	0.0

Value(s)/Threshold: Sweaters

name	description	value
CI	Class Imbalance (CI)	0.888889
DPL	Difference in Positive Proportions in Labels (DPL)	0.0
JS	Jensen-Shannon Divergence (JS)	0.0
KL	Kullback-Liebler Divergence (KL)	0.0
KS	Kolmogorov-Smirnov Distance (KS)	0.0
LP	L-p Norm (LP)	0.0
TVD	Total Variation Distance (TVD)	0.0

Value(s)/Threshold: Swim

name	description	value
CI	Class Imbalance (CI)	0.888889
DPL	Difference in Positive Proportions in Labels (DPL)	0.0
JS	Jensen-Shannon Divergence (JS)	0.0
KL	Kullback-Liebler Divergence (KL)	0.0
KS	Kolmogorov-Smirnov Distance (KS)	0.0
LP	L-p Norm (LP)	0.0
TVD	Total Variation Distance (TVD)	0.0

Value(s)/Threshold: Trend

name	description	value
CI	Class Imbalance (CI)	0.888889
DPL	Difference in Positive Proportions in Labels (DPL)	0.0
JS	Jensen-Shannon Divergence (JS)	0.0
KL	Kullback-Liebler Divergence (KL)	0.0
KS	Kolmogorov-Smirnov Distance (KS)	0.0
LP	L-p Norm (LP)	0.0
TVD	Total Variation Distance (TVD)	0.0

Appendix: Analysis Configuration Parameters

```
{
   "dataset_type": "text/csv",
   "headers": [
      "sentiment",
      "review_body",
      "product_category"
   ],
   "label": "sentiment",
   "label_values_or_threshold": [
      1
   ],
   "facet": [
      {
         "name_or_index": "product_category"
      }
   ],
   "methods": {
      "pre_training_bias": {
         "methods": [
            "CI",
            "DPL",
            "KL",
            "JS",
            "LP",
            "TVD",
            "KS"
         ]
      },
      "report": {
         "name": "report",
         "title": "Analysis Report"
      }
   }
}
```