

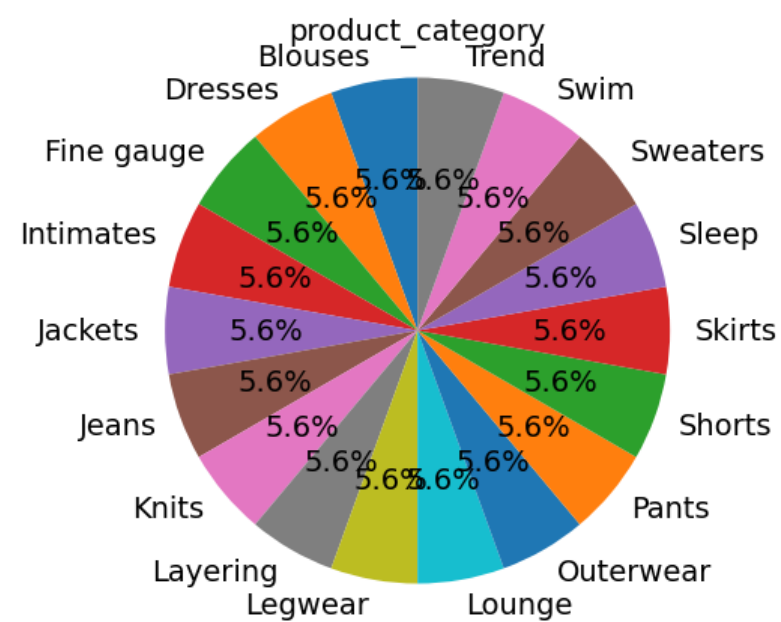
Analysis Report

We report the following SageMaker analysis.

Pre-training Bias Metrics

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

- 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

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"
    }
  }
}
```