Analysis Report

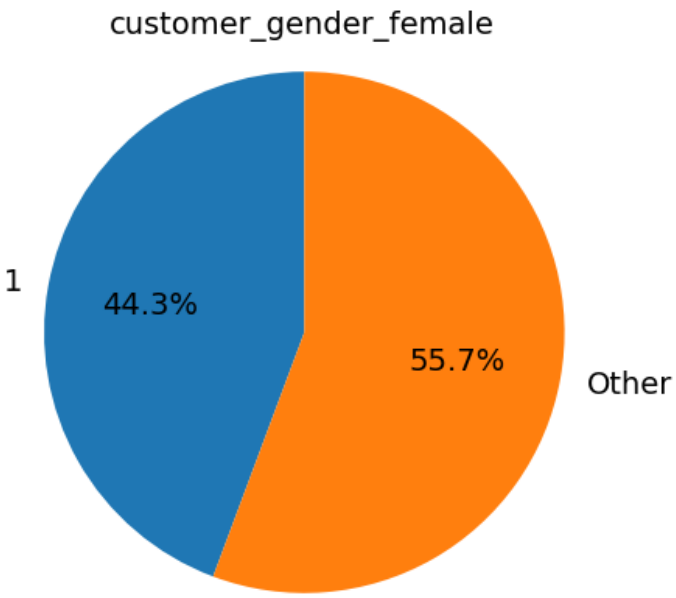
We report the following SageMaker analysis.

Pre-training Bias Metrics

We computed the bias metrics for the label `fraud` using label value(s)/threshold `0`.

- **customer_gender_female**

The groups are represented in the dataset with the following proportions.



Value(s)/Threshold: 1

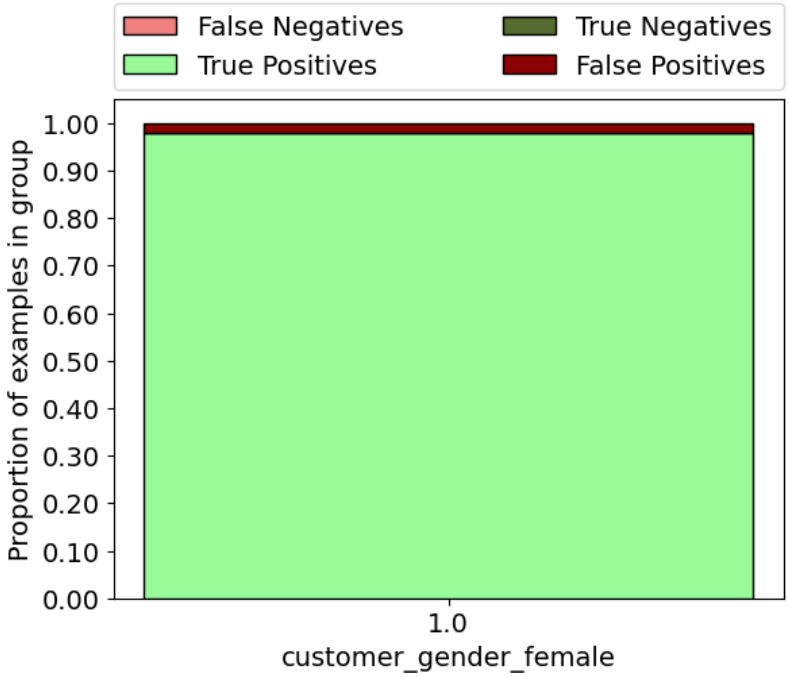
name	description	value	error
CDDL	Conditional Demographic Disparity in Labels (CDDL)	None	Group variable is empty or not provided
CI	Class Imbalance (CI)	0.114	NaN
DPL	Difference in Positive Proportions in Labels (DPL)	-0.026363	NaN
JS	Jensen-Shannon Divergence (JS)	0.002755	NaN
KL	Kullback-Liebler Divergence (KL)	0.012826	NaN
KS	Kolmogorov-Smirnov Distance (KS)	0.026363	NaN
LP	L-p Norm (LP)	0.037282	NaN
TVD	Total Variation Distance (TVD)	0.026363	NaN

Post-training Bias Metrics

We computed the bias metrics for the label `fraud` using label value(s)/threshold `0`.

- **customer_gender_female**

The labels and predictions of the group have the following proportions.



Positive labels = TP + FN --- Used in the following metrics: DPL, JS, KL, KS, LP, TVD
Negative labels = TN + FP
Positive predictions = TP + FP --- Used in the following metrics: DI
Negative predictions = TN + FN
Accuracy = TP + TN --- Used in the following metrics: AD
Recall = TP / (TP + FN) --- Used in the following metrics: RD
Precision = TP / (TP + FP) --- Used in the following metrics: DAR

Value(s)/Threshold: 1

name	description	value	error
AD	Accuracy Difference (AD)	-0.026363	NaN
CDDPL	Conditional Demographic Disparity in Predicted Labels (CDDPL)	None	Group variable is empty or not provided
DAR	Difference in Acceptance Rates (DAR)	-0.026363	NaN
DCA	Difference in Conditional Acceptance (DCA)	-0.026363	NaN
DCR	Difference in Conditional Rejection (DCR)	0	NaN
DI	Disparate Impact (DI)	1.0	NaN
DPPL	Difference in Positive Proportions in Predicted Labels (DPPL)	0.0	NaN
DRR	Difference in Rejection Rates (DRR)	0.0	NaN
FT	Flip Test (FT)	0.0	NaN
RD	Recall Difference (RD)	0.0	NaN
TE	Treatment Equality (TE)	0.0	NaN