



Table 4. Per-group results (accuracy, recall, precision and approval-rate disaggregated by groups A and B) for the offline pipeline.

Variant	A Acc. (†)	B Acc. (†)	A Rec. (†)	B Rec. (†)	A Prec. (†)	B Prec. (†)	A AppRate (†)	B AppRate (†)
rep = 0.5 lbl = 0.0 A support = 632 B support = 579								
none	0.935	0.967	0.954	0.830	0.942	0.976	0.630	0.147
reweight_auto	0.932	0.957	0.939	0.870	0.951	0.879	0.614	0.171
reweight_manual_A0.5_B1.5	0.921	0.960	0.941	0.800	0.932	0.964	0.628	0.143
mitigator_demographic_parity	0.894	0.611	0.847	0.910	0.979	0.296	0.538	0.530
mitigator_equalized_odds	0.929	0.962	0.939	0.850	0.946	0.924	0.617	0.159
rep = 0.5 lbl = 0.4 A support = 616 B support = 605								
none	0.920	0.954	0.948	0.568	0.941	0.636	0.718	0.055
reweight_auto	0.922	0.929	0.948	0.676	0.943	0.446	0.716	0.093
reweight_manual_A0.5_B1.5	0.906	0.954	0.938	0.595	0.930	0.629	0.719	0.058
mitigator_demographic_parity	0.846	0.383	0.800	0.838	0.980	0.078	0.581	0.658
mitigator_equalized_odds	0.909	0.944	0.918	0.568	0.953	0.538	0.687	0.064
rep = 0.5 lbl = 0.5 A support = 644 B support = 576								
none	0.932	0.955	0.965	0.677	0.940	0.568	0.730	0.064
reweight_auto	0.904	0.960	0.891	0.581	0.971	0.643	0.652	0.049
reweight_manual_A0.5_B1.5	0.915	0.972	0.921	0.645	0.957	0.800	0.685	0.043
mitigator_demographic_parity	0.846	0.451	0.808	0.903	0.971	0.082	0.592	0.592
mitigator_equalized_odds	0.925	0.960	0.932	0.677	0.962	0.618	0.689	0.059
rep = 0.5 lbl = 0.6 A support = 618 B support = 597								
none	0.896	0.940	0.928	0.267	0.928	0.364	0.718	0.037
reweight_auto	0.906	0.925	0.948	0.567	0.923	0.347	0.738	0.082
reweight_manual_A0.5_B1.5	0.890	0.951	0.917	0.300	0.929	0.529	0.709	0.028
mitigator_demographic_parity	0.366	0.787	0.119	0.767	0.981	0.161	0.087	0.240
mitigator_equalized_odds	0.906	0.938	0.935	0.300	0.935	0.360	0.718	0.042
rep = 0.6 lbl = 0.0 A support = 771 B support = 455								
none	0.917	0.949	0.925	0.816	0.933	0.873	0.585	0.156
reweight_auto	0.913	0.941	0.919	0.842	0.933	0.810	0.581	0.174
reweight_manual_A0.5_B1.5	0.926	0.947	0.952	0.829	0.925	0.851	0.607	0.163
mitigator_demographic_parity	0.913	0.596	0.914	0.921	0.937	0.282	0.576	0.545
mitigator_equalized_odds	0.925	0.947	0.947	0.842	0.927	0.842	0.603	0.167
rep = 0.6 lbl = 0.4 A support = 771 B support = 463								
none	0.904	0.937	0.905	0.351	0.957	0.722	0.668	0.039
reweight_auto	0.925	0.922	0.945	0.541	0.948	0.513	0.704	0.084
reweight_manual_A0.5_B1.5	0.909	0.942	0.919	0.432	0.951	0.727	0.684	0.048
mitigator_demographic_parity	0.917	0.335	0.949	0.946	0.935	0.103	0.717	0.737
mitigator_equalized_odds	0.925	0.933	0.960	0.595	0.936	0.579	0.725	0.082
rep = 0.6 lbl = 0.5 A support = 776 B support = 461								
none	0.930	0.954	0.913	0.484	0.984	0.750	0.635	0.043
reweight_auto	0.929	0.948	0.917	0.484	0.978	0.652	0.642	0.050
reweight_manual_A0.5_B1.5	0.912	0.957	0.885	0.419	0.985	0.867	0.615	0.033
mitigator_demographic_parity	0.930	0.282	0.951	0.871	0.947	0.076	0.687	0.768
mitigator_equalized_odds	0.936	0.946	0.940	0.484	0.965	0.625	0.666	0.052
rep = 0.6 lbl = 0.6 A support = 763 B support = 461								
none	0.933	0.946	0.970	0.450	0.941	0.391	0.754	0.050
reweight_auto	0.928	0.920	0.970	0.450	0.934	0.257	0.759	0.076
reweight_manual_A0.5_B1.5	0.924	0.946	0.980	0.450	0.921	0.391	0.779	0.050
mitigator_demographic_parity	0.932	0.258	0.973	0.900	0.936	0.050	0.760	0.777
mitigator_equalized_odds	0.934	0.946	0.971	0.450	0.941	0.391	0.755	0.050
rep = 0.7 lbl = 0.0 A support = 881 B support = 338								
none	0.933	0.938	0.950	0.806	0.938	0.847	0.604	0.175
reweight_auto	0.925	0.947	0.930	0.806	0.944	0.893	0.587	0.166
reweight_manual_A0.5_B1.5	0.927	0.944	0.937	0.806	0.941	0.877	0.594	0.169
mitigator_demographic_parity	0.930	0.571	0.943	0.935	0.939	0.291	0.598	0.589
mitigator_equalized_odds	0.930	0.947	0.949	0.839	0.934	0.867	0.605	0.178
rep = 0.7 lbl = 0.4 A support = 890 B support = 349								
none	0.939	0.943	0.965	0.633	0.949	0.679	0.709	0.080
reweight_auto	0.939	0.926	0.955	0.667	0.958	0.556	0.696	0.103
reweight_manual_A0.5_B1.5	0.936	0.940	0.960	0.633	0.949	0.655	0.706	0.083
mitigator_demographic_parity	0.936	0.384	0.958	0.867	0.950	0.110	0.703	0.679
mitigator_equalized_odds	0.938	0.946	0.953	0.633	0.958	0.704	0.694	0.077
rep = 0.7 lbl = 0.5 A support = 887 B support = 360								
none	0.939	0.908	0.977	0.588	0.942	0.278	0.754	0.100
reweight_auto	0.933	0.928	0.946	0.471	0.962	0.320	0.715	0.069
reweight_manual_A0.5_B1.5	0.928	0.947	0.938	0.529	0.962	0.450	0.709	0.056
mitigator_demographic_parity	0.937	0.192	0.974	0.941	0.942	0.052	0.752	0.850
mitigator_equalized_odds	0.939	0.928	0.961	0.588	0.955	0.345	0.732	0.081
rep = 0.7 lbl = 0.6 A support = 908 B support = 344								
none	0.909	0.936	0.953	0.533	0.923	0.348	0.744	0.067
reweight_auto	0.912	0.904	0.963	0.600	0.918	0.250	0.757	0.105
reweight_manual_A0.5_B1.5	0.907	0.956	0.939	0.467	0.933	0.500	0.726	0.041
mitigator_demographic_parity	0.924	0.363	0.953	0.667	0.943	0.045	0.729	0.651
mitigator_equalized_odds	0.910	0.936	0.940	0.533	0.935	0.348	0.726	0.067
rep = 0.8 lbl = 0.0 A support = 1030 B support = 216								
none	0.911	0.944	0.926	0.818	0.923	0.900	0.589	0.185
reweight_auto	0.920	0.940	0.924	0.818	0.939	0.878	0.578	0.190
reweight_manual_A0.5_B1.5	0.914	0.940	0.922	0.795	0.930	0.897	0.583	0.181
mitigator_demographic_parity	0.925	0.736	0.940	0.909	0.933	0.430	0.592	0.431
mitigator_equalized_odds	0.916	0.944	0.924	0.818	0.932	0.900	0.583	0.185
rep = 0.8 lbl = 0.4 A support = 1017 B support = 234								
none	0.942	0.936	0.946	0.533	0.971	0.500	0.687	0.068
reweight_auto	0.940	0.919	0.946	0.600	0.969	0.499	0.689	0.094
reweight_manual_A0.5_B1.5	0.947	0.944	0.955	0.467	0.969	0.583	0.696	0.051
mitigator_demographic_parity	0.946	0.333	0.965	0.800	0.959	0.073	0.711	0.705
mitigator_equalized_odds	0.952	0.932	0.965	0.733	0.967	0.478	0.705	0.098
rep = 0.8 lbl = 0.5 A support = 1007 B support = 234								
none	0.948	0.923	0.972	0.222	0.956	0.154	0.725	0.056
reweight_auto	0.929	0.932	0.940	0.222	0.960	0.182	0.698	0.047
reweight_manual_A0.5_B1.5	0.935	0.953	0.951	0.222	0.958	0.333	0.708	0.026
mitigator_demographic_parity	0.939	0.350	0.965	0.889	0.951	0.050	0.724	0.679
mitigator_equalized_odds	0.946	0.923	0.968	0.444	0.957	0.235	0.721	0.073
rep = 0.8 lbl = 0.6 A support = 1016 B support = 222								
none	0.932	0.896	0.972	0.286	0.938	0.235	0.756	0.077
reweight_auto	0.929	0.815	0.989	0.571	0.920	0.186	0.784	0.194
reweight_manual_A0.5_B1.5	0.930	0.910	0.962	0.357	0.943	0.312	0.744	0.072
mitigator_demographic_parity	0.927	0.374	0.966	0.857	0.936	0.081	0.753	0.671
mitigator_equalized_odds	0.925	0.874	0.961	0.429	0.938	0.231	0.747	0.117