

Table for the comparison for model performance on the original and the conterfactual testing set for each group. "Or_" represents original testing set results, "CF_" represents counterfactual testing set results, "W" for white, "B" for black, "Diff" for the difference between two groups.								
CF	Prompt1		SP	TPR	FPR	Acc	F1	AUC
		Or_W	0.2756	0.4398	0.1572	0.6740	0.5306	0.6413
		CF_W	0.3134	0.4887	0.1870	0.6772	0.5591	0.6509
		Diff	0.0378	0.0489	0.0298	0.0032	0.0285	0.0096
		Or_B	0.5907	0.7210	0.4227	0.6582	0.7038	0.6491
		CF_B	0.5422	0.6873	0.3551	0.6688	0.7004	0.6661
		Diff	0.0485	0.0337	0.0676	0.0106	0.0034	0.0170
	Prompt2	Or_W	0.2331	0.3910	0.1192	0.6756	0.5024	0.6359
		CF_W	0.2882	0.4624	0.1626	0.6803	0.5479	0.6499
		Diff	0.0551	0.0714	0.0434	0.0047	0.0455	0.0140
		Or_B	0.5591	0.6948	0.3841	0.6603	0.6974	0.6553
		CF_B	0.5127	0.6498	0.3357	0.6561	0.6804	0.6570
		Diff	0.0464	0.0450	0.0484	0.0042	0.0170	0.0017
	Prompt3	Or_W	0.1433	0.2669	0.0542	0.6614	0.3978	0.6064
		CF_W	0.1732	0.3233	0.0650	0.6787	0.4574	0.6291
		Diff	0.0299	0.0564	0.0108	0.0173	0.0596	0.0227
		Or_B	0.4293	0.5824	0.2319	0.6635	0.6610	0.6753
		CF_B	0.3787	0.5281	0.1860	0.6530	0.6316	0.6710
		Diff	0.0506	0.0543	0.0459	0.0105	0.0294	0.0043
	Prompt4	Or_W	0.2520	0.4211	0.1300	0.6819	0.5258	0.6455
		CF_W	0.2898	0.4624	0.1653	0.6787	0.5467	0.6485
		Diff	0.0378	0.0413	0.0353	0.0032	0.0209	0.0030
		Or_B	0.5601	0.6985	0.3816	0.6635	0.7005	0.6584
		CF_B	0.5137	0.6648	0.3188	0.6719	0.6954	0.6730
		Diff	0.0464	0.0337	0.0628	0.0084	0.0051	0.0146
	Prompt5	Or_W	0.2898	0.4398	0.1816	0.6598	0.5200	0.6291
		CF_W	0.3213	0.4737	0.2114	0.6567	0.5362	0.6312
		Diff	0.0315	0.0339	0.0298	0.0031	0.0162	0.0021
		Or_B	0.5475	0.6667	0.3937	0.6403	0.6761	0.6364
		CF_B	0.5127	0.6404	0.3478	0.6456	0.6706	0.6463
		Diff	0.0348	0.0263	0.0459	0.0053	0.0055	0.0099
	Prompt6	Or_W	0.2646	0.4398	0.1382	0.6850	0.5392	0.6508
		CF_W	0.3071	0.4737	0.1870	0.6709	0.5466	0.6433
		Diff	0.0425	0.0339	0.0488	0.0141	0.0074	0.0075
		Or_B	0.5643	0.6891	0.4034	0.6487	0.6885	0.6429
		CF_B	0.5158	0.6479	0.3454	0.6508	0.6764	0.6513
		Diff	0.0485	0.0412	0.0580	0.0021	0.0121	0.0084
	Prompt7	Or_W	0.3370	0.5075	0.2141	0.6693	0.5625	0.6467
		CF_W	0.3543	0.5226	0.2331	0.6646	0.5662	0.6447
		Diff	0.0173	0.0151	0.0190	0.0047	0.0037	0.0020
		Or_B	0.6097	0.7322	0.4517	0.6519	0.7032	0.6403
		CF_B	0.5981	0.7079	0.4565	0.6361	0.6866	0.6257
		Diff	0.0116	0.0243	0.0048	0.0158	0.0166	0.0146
	Prompt8	Or_W	0.3559	0.5376	0.2249	0.6756	0.5813	0.6563
		CF_W	0.3654	0.5414	0.2385	0.6693	0.5783	0.6514
		Diff	0.0095	0.0038	0.0136	0.0063	0.0030	0.0049
		Or_B	0.6224	0.7416	0.4686	0.6498	0.7046	0.6365
		CF_B	0.6044	0.7303	0.4420	0.6551	0.7046	0.6442
		Diff	0.0180	0.0113	0.0266	0.0053	0.0000	0.0077
	LR	Or_W	0.3228	0.5000	0.1951	0.6772	0.5648	0.6524
		CF_W	0.4335	0.5524	0.5524	0.6255	0.6243	0.6361
		Diff	0.1107	0.0524	0.3573	0.0517	0.0595	0.0163
		Or_B	0.5454	0.6573	0.4010	0.6319	0.6679	0.6282
		CF_B	0.2425	0.3835	0.3835	0.6598	0.4857	0.6213
		Diff	0.3029	0.2738	0.0175	0.0279	0.1822	0.0069
	MLP	Or_W	0.3228	0.4624	0.2222	0.6457	0.5223	0.6201
		CF_W	0.4536	0.5393	0.3430	0.5907	0.5975	0.5982
		Diff	0.1308	0.0769	0.1208	0.0550	0.0752	0.0219
		Or_B	0.4536	0.5599	0.3164	0.6139	0.6203	0.6217
		CF_B	0.3228	0.4474	0.2331	0.6331	0.5053	0.6072
		Diff	0.1308	0.1125	0.0833	0.0192	0.1150	0.0145