

Table for model performance for each group on original testing set. "F" represents female, "M" represents male, "Diff" represents the difference between two groups								
Org			SP	TPR	FPR	Acc	F1	AUC
	Prompt1	F	0.6606	0.7451	0.5215	0.6443	0.7227	0.6118
		M	0.6446	0.7978	0.4762	0.6673	0.7152	0.6608
		Diff	0.0160	0.0527	0.0453	0.0230	0.0075	0.0499
	Prompt2	F	0.6504	0.7418	0.5000	0.6504	0.7252	0.6209
		M	0.6163	0.7690	0.4484	0.6654	0.7065	0.6603
		Diff	0.0341	0.0272	0.0516	0.0150	0.0187	0.0394
	Prompt3	F	0.7988	0.8627	0.6935	0.6524	0.7553	0.5846
		M	0.7996	0.8953	0.6944	0.6144	0.7086	0.6004
		Diff	0.0008	0.0326	0.0009	0.0380	0.0467	0.0158
	Prompt4	F	0.6728	0.7680	0.5161	0.6606	0.7378	0.6259
		M	0.6541	0.8087	0.4841	0.6692	0.7191	0.6623
		Diff	0.0187	0.0407	0.0320	0.0086	0.0187	0.0364
	Prompt5	F	0.5752	0.6471	0.4570	0.6077	0.6723	0.5950
		M	0.6011	0.7437	0.4444	0.6541	0.6924	0.6496
		Diff	0.0259	0.0966	0.0126	0.0464	0.0201	0.0546
	Prompt6	F	0.5874	0.6765	0.4409	0.6321	0.6958	0.6178
		M	0.6673	0.8051	0.5159	0.6522	0.7079	0.6446
		Diff	0.0799	0.1286	0.0750	0.0201	0.0121	0.0268
	Prompt7	F	0.5427	0.6307	0.3978	0.6199	0.6736	0.6164
		M	0.6163	0.7509	0.4683	0.6465	0.6899	0.6413
		Diff	0.0736	0.1202	0.0705	0.0266	0.0163	0.0249
	Prompt8	F	0.5102	0.5980	0.3656	0.6118	0.6571	0.6162
		M	0.4650	0.6317	0.2817	0.6730	0.6692	0.6750
		Diff	0.0452	0.0337	0.0839	0.0612	0.0121	0.0588
	LR	F	0.8211	0.8889	0.7097	0.6626	0.7662	0.5896
		M	0.4858	0.6426	0.3135	0.6635	0.6667	0.6646
		Diff	0.3353	0.2463	0.3962	0.0009	0.0995	0.0750
	MLP	F	0.6199	0.7092	0.4631	0.6402	0.7103	0.6180
		M	0.4405	0.5451	0.3254	0.6068	0.5922	0.6099
		Diff	0.1794	0.1641	0.1377	0.0334	0.1181	0.0081