

Acquisition step		50	100	200	300	400
Target	Method					
70B-zero	US	2.3826	1.1493	6.3630	0.39690	0.30590
	LURE-Llama2-7B	0.8497	0.4371	0.2247	0.1524	0.1090
	LURE-Llama2-70B	0.5703	0.2996	0.1418	0.09900	0.07770
7B-zero	US	4.3455	2.3304	1.3545	0.96420	0.80470
	LURE-Llama2-7B	3.8653	1.9395	1.1326	0.86310	0.71630
	LURE-Llama2-70B	3.0157	1.7693	0.94190	0.73000	0.64030
70B-few	US	24.5497	11.8593	5.8947	4.263	2.897
	LURE-Llama2-7B	26.4253	13.5746	6.9755	4.249	3.0014
	LURE-Llama2-70B	19.8693	11.2704	6.1626	3.983	2.9785
7B-few	US	23.7706	11.7404	6.6989	4.6215	3.4943
	LURE-Llama2-7B	19.6371	9.638	5.8037	4.0226	3.1452
	LURE-Llama2-70B	13.4897	7.5196	4.1388	2.7654	2.2691

Table 4: Numerical values for experiments on Hate Speech. All numerical values in the table are scaled by 10^{-4} . For each experiment, minimum value is shown in bold.