

Red-highlighted cells denote corrected values that were misreported in the original publication. All remaining entries are unchanged from the original papers.

Table 2: AUC comparison with only the LA fine-tuned (ViT-B frozen) on an out-of-domain LLM and dataset. Last row shows ACT-ViT’s gain over the best baseline from prior work.

	Mis-7B					LlaMa-8B					Qwen-7B				
	HotpotQA	Movies	HQA-Wc	IMDB	TriviaQA	HotpotQA	Movies	HQA-Wc	IMDB	TriviaQA	HotpotQA	Movies	HQA-Wc	IMDB	TriviaQA
Best-Probas	63.00	66.00	56.00	57.00	63.00	67.00	75.00	56.00	73.00	74.00	67.50	74.20	67.40	74.80	69.70
Probe[*]	82.69	74.87	66.80	97.61	79.95	79.71	83.61	70.56	93.58	80.13	80.57	93.22	75.01	93.76	87.08
ACT-MLP	80.22	72.83	60.68	95.05	79.64	76.83	75.51	64.50	88.30	76.65	83.40	86.03	73.64	89.14	86.34
ACT-ViT	83.73	79.07	68.87	97.58	83.04	80.83	82.31	70.66	94.71	82.63	87.27	95.00	77.56	96.25	91.14
Improvement	+1.04	+4.20	+2.07	−0.03	+3.09	+1.12	−1.30	+0.10	+1.13	+2.50	+6.70	+1.78	+2.55	+2.49	+4.06