

&A Accuracy vs AUROC: no_abst_norm_logits_first_prompt, hellaswag ($r = 0.$

