

Q&A Accuracy vs AUROC: yes_abst_raw_logits_first_prompt, mmlu ($r = 0.76$)

