

Q&A Accuracy vs AUROC: yes_abst_raw_logits_first_prompt, truthfulqa (r = 0.1)

