

Q&A Accuracy vs AUROC: no_abst_norm_logits_second_prompt, mmlu ($r = 0.6$)

