

Q&A Accuracy vs AUROC: no_abst_norm_logits_first_prompt, winogrande (r = 0.70)

