

Q&A Accuracy vs AUROC: no\_abst\_norm\_logits\_second\_prompt, mmlu ( $r = 0.6$ )

