

Q&A Accuracy vs AUROC: yes\_abst\_raw\_logits\_first\_prompt, mmlu (r = 0.76)

