

Q&A Accuracy vs AUROC: yes\_abst\_norm\_logits\_first\_prompt, hellaswag ( $r = 0$ )

