

Q&A Accuracy vs AUROC: no\_abst\_raw\_logits\_second\_prompt, hellaswag (r = 0.94)

