

Q&A Accuracy vs AUROC: yes_abst_raw_logits, truthfulqa ($r = 0.87$)

