

Q&A Accuracy vs AUROC: yes_abst_norm_logits, hellaswag ($r = 0.34$)

