Q&A Accuracy vs AUROC: yes abst norm logits, piqa (r = 0.46) SOLAR 10.7B Mixtral 8x7B Llama2YB6B 70 Llama2 13B Mistral 7B Llama2 7B Falcon 40B 65 60 55 Falcon 7B 50 45 40 35 30 52.5 62.5 65.0 67.5 50.0 55.0 57.5 60.0 70.0

AUROC

Q&A Accuracy