

Q&A Accuracy vs AUROC: no\_abst\_norm\_logits\_first\_prompt, piqa ( $r = 0.24$ )

