

Q&A Accuracy vs AUROC: Max Logit, hellaswag ( $r = 0.85$ )

