

Q&A Accuracy vs AUROC: Max Logit, prompt comparison

