Q&A Accuracy vs AUROC: MSP and Max Logit Mixtral-8x7B Yi-6B 0.70 Mistral-7B SOLAR-10.7B Llama2-70B Falcon-40B Yi-6Bı Yi-34B Llama2-13B 0.65 Llama2-13B Falcon-40B Llama2-7B Llama2-7B 0.60 0.55 Q&A Accuracy 0.50 0.45 0.40 0.35 **MSP** Max Logit 0.30 0.650 0.500 0.525 0.550 0.575 0.600 0.625 0.675 0.700

AUROC