

A Accuracy vs AUROC: no\_abst\_norm\_logits\_second\_prompt, hellaswag (r = 0.92)

