

A Accuracy vs AUROC: no_abst_norm_logits_second_prompt, hellaswag (r = 0.92)

