

Score vs Confidence Threshold: no_abst_norm_logits_second_prompt, p1qa

