

Score vs Confidence Threshold: yes\_abst\_norm\_logits\_second\_prompt, winogrande

