

Performance on simple inferences vs. LMSYS Elo score

Temperature: 1, Condition: few-shot

