

Best Model Performance by Prompt

Best: Mistral 7B

Best: OpenChat 3.5

Best: LLaMA 3.1

Best: Mistral 7B DP

Best: Zephyr 7B

