



ReasonEval-34B

Skywork-1.5B

o1-mini

MATHMinos-7B

Gemini-thinking

GPT-4o

MathShepherd-7B