

HLE-Raw Leaderboard

Gemini 3 Pro  40.4%

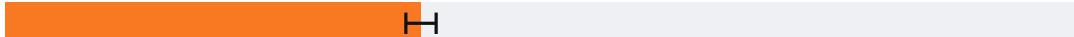
Claude Opus 4.6  39.0%

GPT-5.2  33.3%

Claude Opus 4.5  30.0%

Qwen3-Max-Thinking  30.0%

DeepSeek-V3.2  24.9%

Grok 4.1 (Fast)  19.9%

