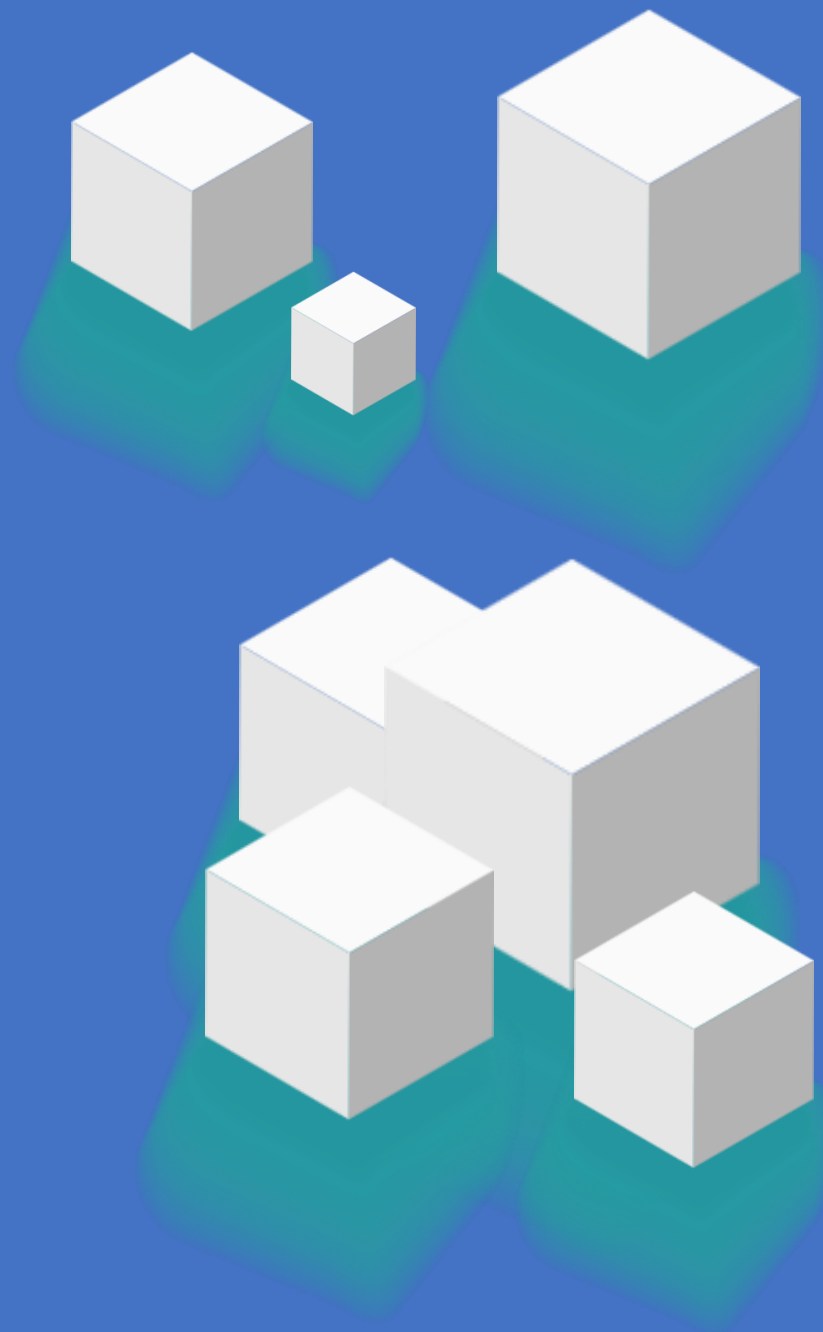
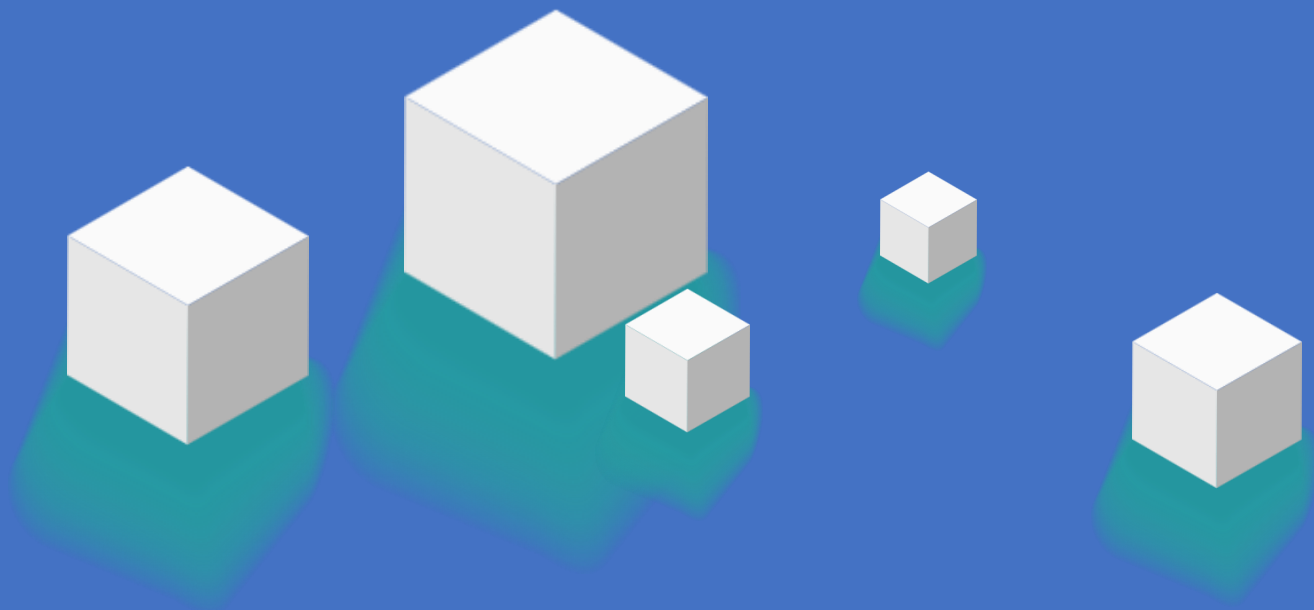


Claude 4



# Claude 4

---

Claude 4:

2025 年 5 月发布，包括旗舰级 Claude Opus 4 和高性价比的 Claude Sonnet 4。Claude 4 在编程、推理和长任务执行方面表现卓越，是当前最强的 AI 编程模型之一。

编程能力:

- 在 SWE-bench（真实软件工程任务基准测试）中，Opus 4 得分 72.5%，Sonnet 4 得分 72.7%，**均领先竞争对手如 OpenAI 的 GPT-4.1 和 Google 的 Gemini 2.5 Pro。**
- 在 Terminal-bench（终端交互任务测试）中，Opus 4 以 43.2% 的成绩领先。
- **可连续执行 7 小时的复杂编程任务（如代码重构），远超前代模型的几十分钟极限**

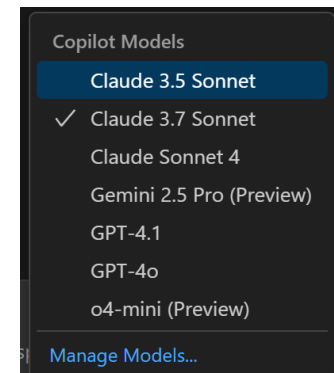
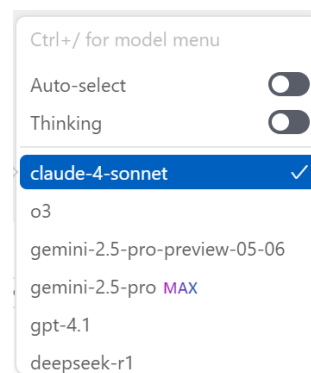
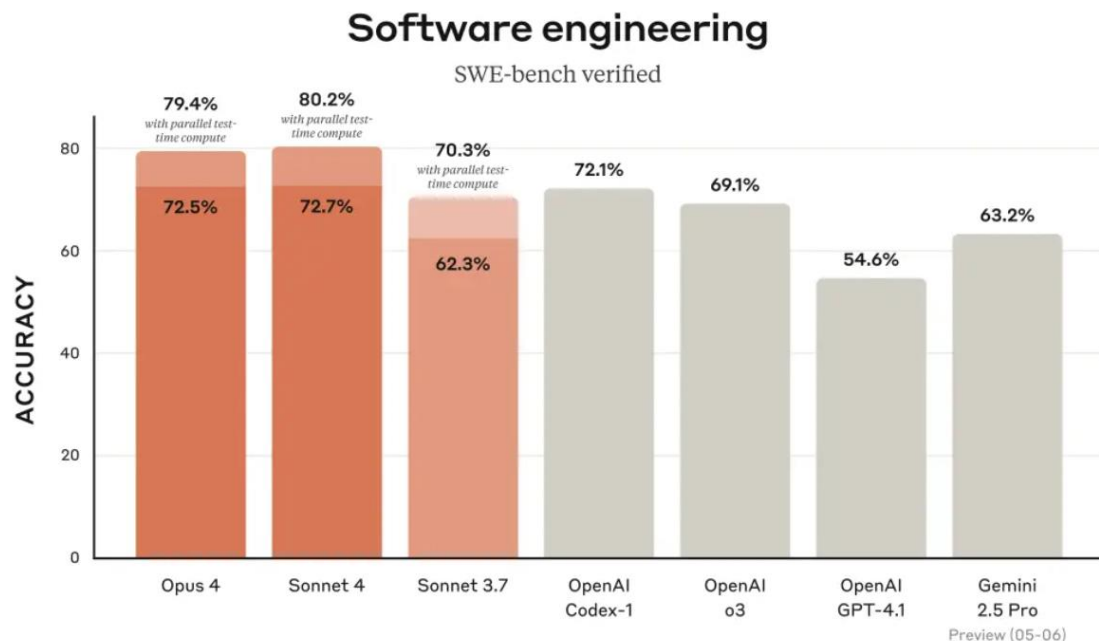
推理能力:

- 支持扩展思考模式（Extended Thinking Mode），可进行长达 64K token 的深度推理，适用于复杂问题解决。
- 在 MMMLU（多学科知识测试）中，Opus 4 得分 87.4%，Sonnet 4 得分 85.4%




# Claude 4

Claude 4的创新：

- **长任务稳态**：能自主拆解任务、记录进度，并跨多轮对话保持上下文，适合长期项目开发。
- **工具并行**：自动调用搜索引擎、代码执行器等工具，无需人工干预，提高任务执行效率。
- **行为安全锁 (ASL-3)**：减少 65% 的逻辑漏洞和作弊倾向，确保 AI 行为更可靠



可以在Cursor, Github Copilot中使用



Thank You  
Using data to solve problems