AI Model Comparison (2025)

| Model | Developer | Unique Strengths | Known Limitations / Notes |
| --- | --- | --- | --- |
| Claude (Anthropic) | Anthropic | Multimodal (text+image), strong reasoning, very large context windows (100K → 200K → up to 1M tokens)[Wikipedia](https://en.wikipedia.org/wiki/Claude_%28language_model%29?utm_source=chatgpt.com) | Proprietary; cautious ethical alignment sometimes limits usability |
| DeepSeek (R1/V3) | DeepSeek (China) | Open-source (MIT), cost-efficient, plays strong in reasoning and healthcare domains[arXiv](https://arxiv.org/abs/2506.01257?utm_source=chatgpt.com)[Wikipedia](https://en.wikipedia.org/wiki/DeepSeek_%28chatbot%29?utm_source=chatgpt.com)[Barron's](https://www.barrons.com/articles/what-is-deepseek-ai-china-nvidia-stocks-bd00ca90?utm_source=chatgpt.com)[The Times of India](https://timesofindia.indiatimes.com/technology/tech-news/chinas-deepseek-that-shocked-america-and-american-technology-companies-has-an-update-that-it-says-to-openais-o3-and-googles-gemini-2-5-pro/articleshow/121498260.cms?utm_source=chatgpt.com) | Regulatory/ethical concerns (censorship, data-security), technical complexity |
| Gemini (Google DeepMind) | Google DeepMind | Multimodal + huge context (~1M+ tokens), strong reasoning & coding, integrated with Google ecosystem[Wikipedia](https://en.wikipedia.org/wiki/Gemini_%28language_model%29?utm_source=chatgpt.com) | Closed-source; potential privacy/data-integration concerns |
| Grok (xAI) | xAI (Elon Musk) | Real-time web/X access, "Think" reasoning modes, advanced reasoning, multimodal support[Tom's Guide](https://www.tomsguide.com/ai/what-is-grok?utm_source=chatgpt.com)[Business Insider](https://www.businessinsider.com/grok-artificial-intelligence-chatbot-elon-musk-xai-explained-2025-7?utm_source=chatgpt.com)[The Verge](https://www.theverge.com/news/614218/elon-musk-xai-big-brain-reasoning-grok-3?utm_source=chatgpt.com)[Voiceflow](https://www.voiceflow.com/blog/grok?utm_source=chatgpt.com) | Controversial outputs, limited access via X Premium tiers |
| GPT-5 (OpenAI) | OpenAI | State-of-the-art across benchmarks—coding, multimodal, medical, long-context; system-of-models architecture (fast + reasoning)[Wikipedia](https://en.wikipedia.org/wiki/GPT-5?utm_source=chatgpt.com)[OpenAI](https://openai.com/index/introducing-gpt-5-for-developers/?utm_source=chatgpt.com)[arXiv+1](https://arxiv.org/abs/2508.19259?utm_source=chatgpt.com) | Early rollout criticized for glitches and reduced empathy[The Week](https://theweek.com/tech/gpt-5-open-ai-launch-fail?utm_source=chatgpt.com)[Axios](https://www.axios.com/2025/08/12/gpt-5-bumpy-launch-openai?utm_source=chatgpt.com) |
| Perplexity AI | Perplexity AI | AI-powered search engine; provides real-time, citation-backed answers; supports multiple backend models[The Sun](https://www.thesun.co.uk/tech/36428010/what-is-perplexity-ai/?utm_source=chatgpt.com)[Wikipedia](https://en.wikipedia.org/wiki/Perplexity_AI?utm_source=chatgpt.com)[Perplexity AI](https://www.perplexity.ai/help-center/en/articles/10354919-what-advanced-ai-models-are-included-in-a-perplexity-pro-subscription?utm_source=chatgpt.com) | Focused on research & citation; not ideal for creative or coding tasks |

Model Highlights & Context

Claude (Anthropic)

* Latest release is Claude 4 (Opus & Sonnet), as of May 2025[Wikipedia](https://en.wikipedia.org/wiki/Claude_%28language_model%29?utm_source=chatgpt.com).
* Offers exceptionally long context capacity, supporting advanced document workflows.

DeepSeek (China)

* Released the open-source DeepSeek-R1 in early 2025; praised for efficiency and reasoning in domains like healthcare[arXiv](https://arxiv.org/abs/2506.01257?utm_source=chatgpt.com)[Wikipedia](https://en.wikipedia.org/wiki/DeepSeek_%28chatbot%29?utm_source=chatgpt.com).
* DeepSeek-R1-0528 update claims better math, coding, and fewer hallucinations[The Times of India](https://timesofindia.indiatimes.com/technology/tech-news/chinas-deepseek-that-shocked-america-and-american-technology-companies-has-an-update-that-it-says-to-openais-o3-and-googles-gemini-2-5-pro/articleshow/121498260.cms?utm_source=chatgpt.com).
* Major concerns regarding censorship and Chinese regulatory implications[The Times](https://www.thetimes.co.uk/article/deepseek-is-here-should-you-use-it-in-your-business-mg7m7csff?utm_source=chatgpt.com)[Barron's](https://www.barrons.com/articles/what-is-deepseek-ai-china-nvidia-stocks-bd00ca90?utm_source=chatgpt.com).

Gemini (Google DeepMind)

* Gemini 2.5 Pro (released June 2025) supports reasoning models with a 1M-token context window[Wikipedia](https://en.wikipedia.org/wiki/Gemini_%28language_model%29?utm_source=chatgpt.com).
* Excels in multimodal inputs, integrated with Google services (Search, Workspace, Bard).

Grok (xAI)

* Grok 4 (July 2025) delivers enhanced reasoning, voice interface, and content interpretation like memes[Tom's Guide](https://www.tomsguide.com/ai/what-is-grok?utm_source=chatgpt.com)[Business Insider](https://www.businessinsider.com/grok-artificial-intelligence-chatbot-elon-musk-xai-explained-2025-7?utm_source=chatgpt.com).
* Grok-3 introduced "Think" and "Big Brain" modes with advanced transparency[The Verge](https://www.theverge.com/news/614218/elon-musk-xai-big-brain-reasoning-grok-3?utm_source=chatgpt.com)[New York Post](https://nypost.com/2025/02/18/business/elon-musks-xai-claims-grok-3-outperforms-openai-deepseek/?utm_source=chatgpt.com).
* Offers real-time data via X; availability limited to paid tiers.

GPT-5 (OpenAI)

* Officially launched August 7, 2025—combined reasoning & non-reasoning models, very long context, strong real-world benchmark performance[Wikipedia](https://en.wikipedia.org/wiki/GPT-5?utm_source=chatgpt.com)[OpenAI](https://openai.com/index/introducing-gpt-5-for-developers/?utm_source=chatgpt.com).
* Outperforms GPT-4 in medical and academic settings per recent studies[arXiv+1](https://arxiv.org/abs/2508.19259?utm_source=chatgpt.com).
* Some early user backlash due to rollout issues and shifted tone[The Week](https://theweek.com/tech/gpt-5-open-ai-launch-fail?utm_source=chatgpt.com)[Axios](https://www.axios.com/2025/08/12/gpt-5-bumpy-launch-openai?utm_source=chatgpt.com).

Perplexity AI

* Functions as a conversational search engine—answers are real-time, with sources cited[Wikipedia](https://en.wikipedia.org/wiki/Perplexity_AI?utm_source=chatgpt.com)[The Sun](https://www.thesun.co.uk/tech/36428010/what-is-perplexity-ai/?utm_source=chatgpt.com).
* Plug-in access to high-end LLMs (GPT-5, Claude) via Pro subscription[Perplexity AI](https://www.perplexity.ai/help-center/en/articles/10354919-what-advanced-ai-models-are-included-in-a-perplexity-pro-subscription?utm_source=chatgpt.com).
* Best suited for research, less so for creative or coding tasks.

TL;DR: Which Model for What?

* Best overall performer (coding, reasoning, multimodal): GPT-5
* Top for long-context, document-heavy workflows: Claude
* Most open and cost-efficient reasoning model: DeepSeek
* Strongest at real-time web integration & witty reasoning: Grok
* Best for multimodal industry integration (Google): Gemini
* Ideal for citation-backed search and research: Perplexity