

Chatbot Arena LLM Leaderboard: Community-driven Evaluation for Best LLM and AI chatbots

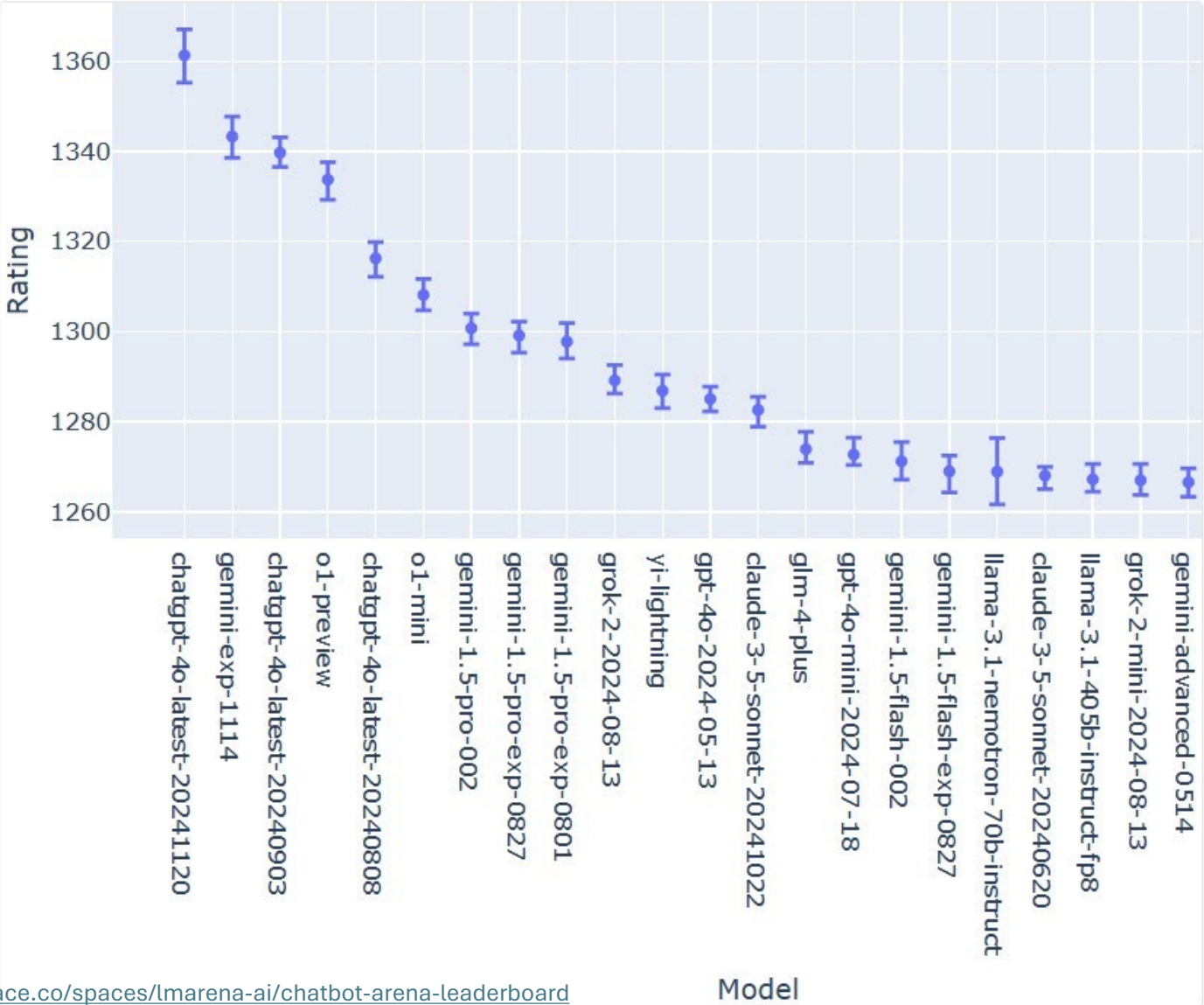
Last updated: 2024-11-19.

Rank* (UB)	Rank (StyleCtrl)	Model	Arena Score	95% CI	Votes	Organizatio	License	Knowledge Cutoff
1	1	ChatGPT-4o-latest (2024-11-20)	1361	+6/-6	8513	OpenAI	Proprietary	Unknown
2	4	Gemini-Exp-1114	1343	+4/-5	11566	Google	Proprietary	Unknown
3	2	o1-preview	1334	+4/-4	27357	OpenAI	Proprietary	2023/10
4	7	o1-mini	1308	+4/-3	31158	OpenAI	Proprietary	2023/10
5	4	Gemini-1.5-Pro-002	1301	+3/-4	26456	Google	Proprietary	Unknown
5	4	Gemini-1.5-Pro-Exp-0827	1299	+3/-4	32351	Google	Proprietary	2023/11
7	10	Grok-2-08-13	1289	+3/-3	51162	xAI	Proprietary	2024/3
7	12	Yi-Lightning	1287	+4/-4	29081	01 AI	Proprietary	Unknown
7	4	GPT-4o-2024-05-13	1285	+3/-3	110841	OpenAI	Proprietary	2023/10
8	3	Claude 3.5 Sonnet (20241022)	1283	+3/-4	28535	Anthropic	Proprietary	2024/4

<https://huggingface.co/spaces/lmarena-ai/chatbot-arena-leaderboard>

More Statistics for Chatbot Arena (Overall)

Figure 1: Confidence Intervals on Model Strength (via Bootstrapping)



More Statistics for Chatbot Arena (Overall)

Figure 2: Average Win Rate Against All Other Models (Assuming Uniform Sampling and No Ties)

