

Model / Source	Recent Advances in Generative AI and Large Language Models: Current Status, Challenges, and Perspectives	How Efficient is LLM-Generated Code? A Rigorous & High Standard Benchmark	Top Leaderboard Ranking = Top Coding Proficiency, Always? EVOEVAL: Evolving Coding Benchmarks via LLM	Generative AI in Academic Writing: A Comparison of DeepSeek, Qwen, ChatGPT, Gemini, Llama, Mistral, and Gemma	Shakudo	Thisisdefinition	Unite.ai	Zapier	OpenRouter	Vellum	BotPress	Livebench	Total
GPT	1	1	1	1	1	1	1	1	1	1	1	1	12
Llama	1	1	1	1	1	1	0	1	1	1	1	1	11
Claude	0	1	1	0	1	1	1	1	1	1	1	1	10
Gemini	1	0	1	1	1	1	1	1	1	0	1	1	10
Deepseek	0	0	1	1	1	0	1	1	1	1	1	1	9
Mixtral	0	1	1	1	1	1	0	1	1	0	1	1	9
Qwen	0	0	1	1	1	0	0	1	1	1	1	1	8
Grok	0	0	0	0	1	0	1	1	0	0	0	0	3
Command	0	0	0	0	1	0	0	1	0	0	0	0	2
StarCoder	0	1	1	0	0	0	0	0	0	0	0	0	2
Palm	1	0	1	0	0	0	0	0	0	0	0	0	2
Others models	appears at most 1 time (phi, nova, incoder, copilot etc)												0