Grok 4.5: Fast and Cheap, but Not the Best at Every Coding Test

Kabui, Charles

2026-07-20

SpaceXAI released Grok 4.5 for coding, agent tasks, and office work. Its model page lists a 500,000-token context window, image input, and prices of $2 per million input tokens and $6 per million output tokens. Those rates double when a request exceeds 200,000 tokens. SpaceXAI says the model streams 80 tokens per second and averages 15,954 output tokens per SWE-Bench Pro task, 4.2 times fewer than Claude Opus 4.8 at 67,020. Its benchmark chart

is mixed. Grok leads SWE Marathon at 29% versus Opus 4.8 at 26%, but trails Fable 5 on DeepSWE 1.1, 53% to 70%, and SWE-Bench Pro, 64.7% to 80.4%. SpaceXAI also says the comparison scores came from provider reports and leaderboards, not one controlled test.

Artificial Analysis found a clearer advantage in cost. Grok 4.5 scored 76 on its Coding Agent Index, level with GPT-5.5 in Codex and below Fable 5 in Claude Code. Grok cost $2.49 per task in that setup, compared with $5.07 for GPT-5.5 and $11.80 for Fable 5. That makes it useful for teams running many coding-agent jobs, where task cost matters more than the token rate alone. The catch is workload fit: benchmark results change with the agent framework and task mix, while long prompts trigger twice the normal API price. Grok 4.5 is a strong price-performance option, not a safe default for every coding job.

Read More: Grok 4.3 cut token prices and added voice cloning.

Sources:

- SpaceXAI: Introducing Grok 4.5
- SpaceXAI Docs: Grok 4.5 specifications and pricing
- Artificial Analysis: Grok 4.5 benchmarks and task costs
- DataCamp: Grok 4.5 features, benchmarks, pricing, and tests

---