

AI Cost Calculator

Compare Model Pricing & ROI Before You Build

Model Pricing Comparison

GPT-4o (OpenAI)	\$\$\$
Input: \$2.50/1M tokens Output: \$10/1M Context: 128K	
GPT-4o mini (OpenAI)	\$
Input: \$0.15/1M tokens Output: \$0.60/1M Context: 128K	
Claude 3.5 Sonnet (Anthropic)	\$\$\$
Input: \$3/1M tokens Output: \$15/1M Context: 200K	
Gemini 1.5 Pro (Google)	\$\$
Input: \$1.25/1M tokens Output: \$5/1M Context: 2M	
Gemini 1.5 Flash (Google)	\$
Input: \$0.075/1M tokens Output: \$0.30/1M Context: 1M	
Llama 3.1 70B (Meta/self-hosted)	\$
Input: ~\$0.10/1M* Output: ~\$0.10/1M* Context: 128K	
Mistral Large (Mistral)	\$\$
Input: \$2/1M tokens Output: \$6/1M Context: 128K	

ROI Examples

- Customer Support Bot**
100K req/mo | GPT-4o mini | \$45/mo AI cost | Replaces 2 FTE | Net: ~\$7,955/mo saved
- Content Generation**
50K req/mo | Gemini Flash | \$15/mo AI cost | Replaces 1 FTE | Net: ~\$2,985/mo saved
- Code Review Assistant**
10K req/mo | Claude Sonnet | \$150/mo AI cost | Replaces 0.5 FTE | Net: ~\$2,350/mo saved

Quick Cost Formula

Monthly cost = (Monthly requests × Avg input tokens × Input price) + (Monthly requests × Avg output tokens × Output price)

Example: 10K requests/mo, 500 input tokens, 200 output tokens, GPT-4o mini:
= (10,000 × 500 × \$0.00000015) + (10,000 × 200 × \$0.00000006)
= \$0.75 + \$1.20 = \$1.95/month

Cost-Saving Tips

- Use Flash/Mini models for simple tasks**
80% of tasks don't need GPT-4 or Claude Sonnet
- Cache frequent prompts**
OpenAI and Anthropic offer prompt caching (50% discount)
- Self-host open models at scale**
Llama 3.1 on RunPod/AWS can be 10x cheaper at high volume

🔗 **Batch API for non-realtime tasks**

OpenAI Batch API offers 50% discount with 24h turnaround

🔗 **Optimize prompt length**

Shorter prompts = fewer tokens = lower cost. Be concise.

Recommended Tools

- 🔗 **OpenAI Platform** GPT-4o, GPT-4o mini, DALL-E — platform.openai.com
- 🔗 **Google AI Studio** Gemini Pro & Flash — aistudio.google.com
- 🔗 **RunPod** Self-host Llama & more — runpod.io

Disclosure: Some links are affiliate links. We earn a commission at no extra cost to you.

