Model Cost vs Performance anthropic/claude-sonnet-4 0.030 x-ai/grok-3-beta Cost per Attempt 0.025 openai/o4-mini 0.020 0.015 Average 0.010 google/gemini-2.5-flash-preview-05-20:thinking x-ai/grok-3-mini-beta 0.005 google/gemini-2.5-flash-preview-05-20 openai/gpt-4.1 45 50 55 60 65 Success Rate (%)