

Task Length

$R^2 = 0.82$

2022-01

2022-07

2023-01

2023-07

2024-01

2024-07

2025-01

2025-07

2026-01

Model Release Date

GPT-3.5

GPT-4

Claude 3.5 Sonnet

GPT-4o

Claude 3.7 Sonnet

Claude 4 Sonnet

GPT-5  
Claude 4.5 Opus  
Gemini 3 Pro

