

□ Model Consistency (Lower Variance = More Consistent)

Score Variance

0.035
0.030
0.025
0.020
0.015
0.010
0.005
0.000

Claude-Sonnet

DeepSeek-V3

Gemini-Flash

gemini-2.5-pro

gpt-4o

gpt-5

gpt-5-mini

AI Model

0.0005

0.0215

0.0083

0.0009

0.0341

0.0141

0.0007

