Accuracy GPT-40-MINI GPT-40 14-3.5-Sonnet **3.5-Haiku** 12-LLAMA-3.3-70B 10-8 -6-4 10^{-2} Cost