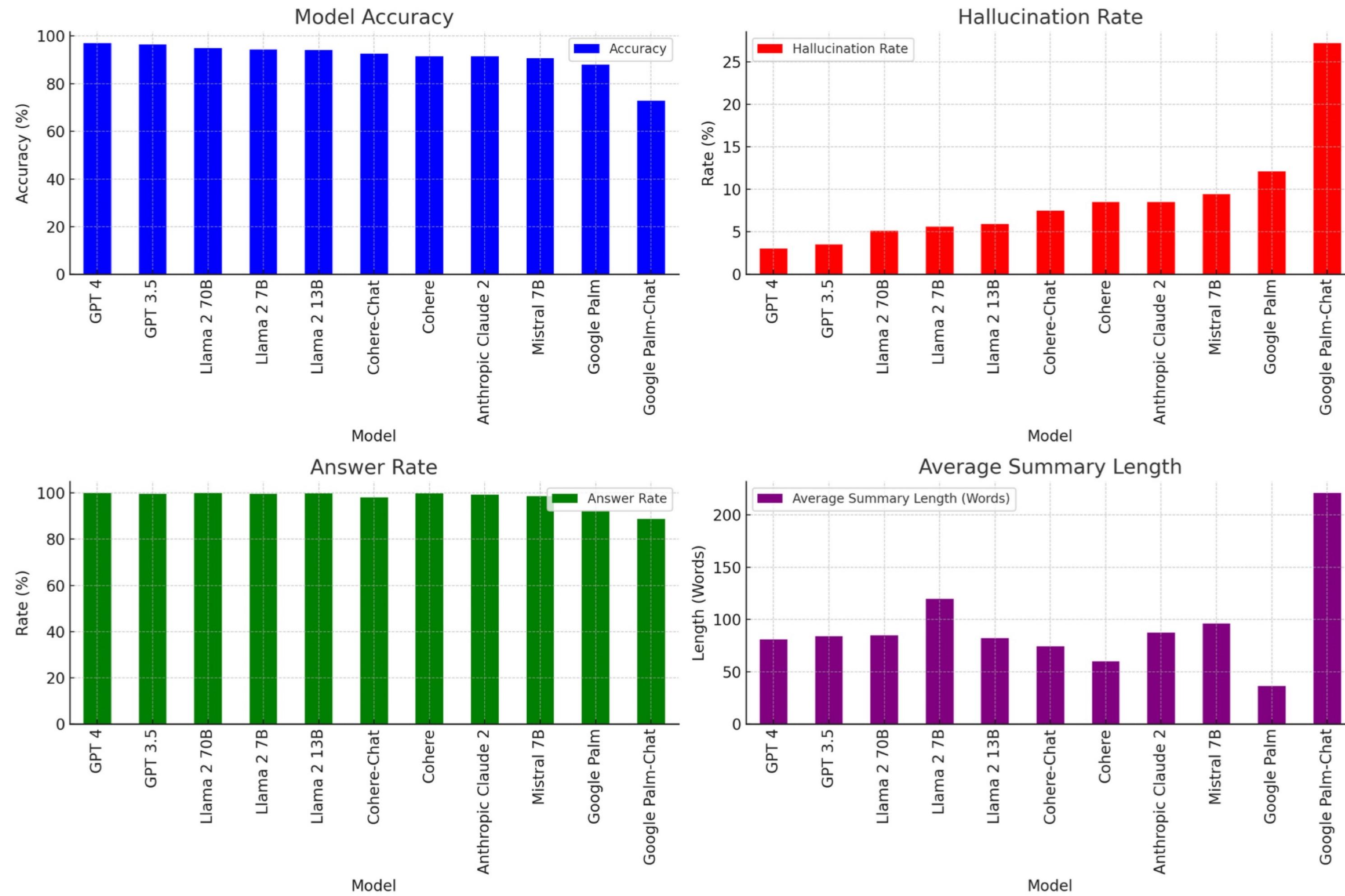




AI Blue

<http://www.bit.ly/ai-blue>

<http://34.121.86.8:7860/>



Make things Precise

Strategies to Achieve Optimal Performance:

User Type	Recommended Method(s)	Efficiency	Constrained Cost	Hypothetical Real-life Example
1) General Users	B	60%	Low	Casual Q&A chatbot - Answering general knowledge questions
2) Business Users	A+D	88%		News Summarization - Generating summaries of news articles
3) Rigorous Academic Research Users	C+D	90%	High	Legal document analysis - Analyzing contracts and case law
4) Highly Sensitive Policy Users (i.e., geopolitics analysis)	A+C+D+E	99%		Geopolitical risk assessment - Analyzing and forecasting risks in international relations

A: Prompt Engineering - This involves crafting precise instructions to guide the AI's output.

B: Temperature Setting - We balance between encouraging creativity and ensuring fact-checking by optimizing this setting.

C: Verification with Multiple AIs - We cross-verify results using various AI engines to ensure accuracy and depth.

D: Verification with Credible Online Sources - Our process includes automated cross-checking of information with reliable, up-to-date online sources via Pinecone vector database.

E: Fine-Tuning - We make use of specific data in certain areas to enhance the AI's training and improve its accuracy in those domains.

AI grading:

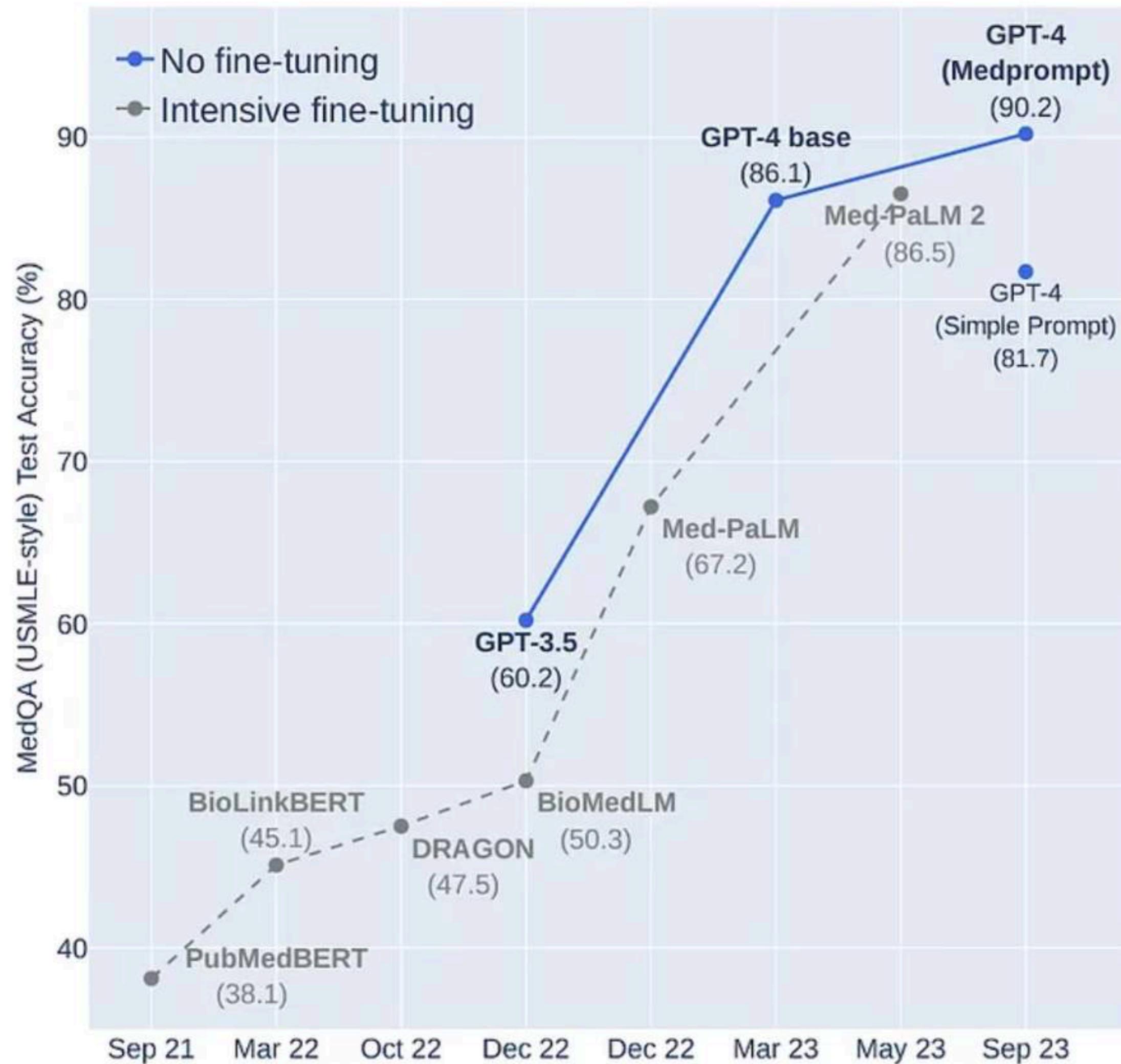
GPT-4: A

Claude: A-

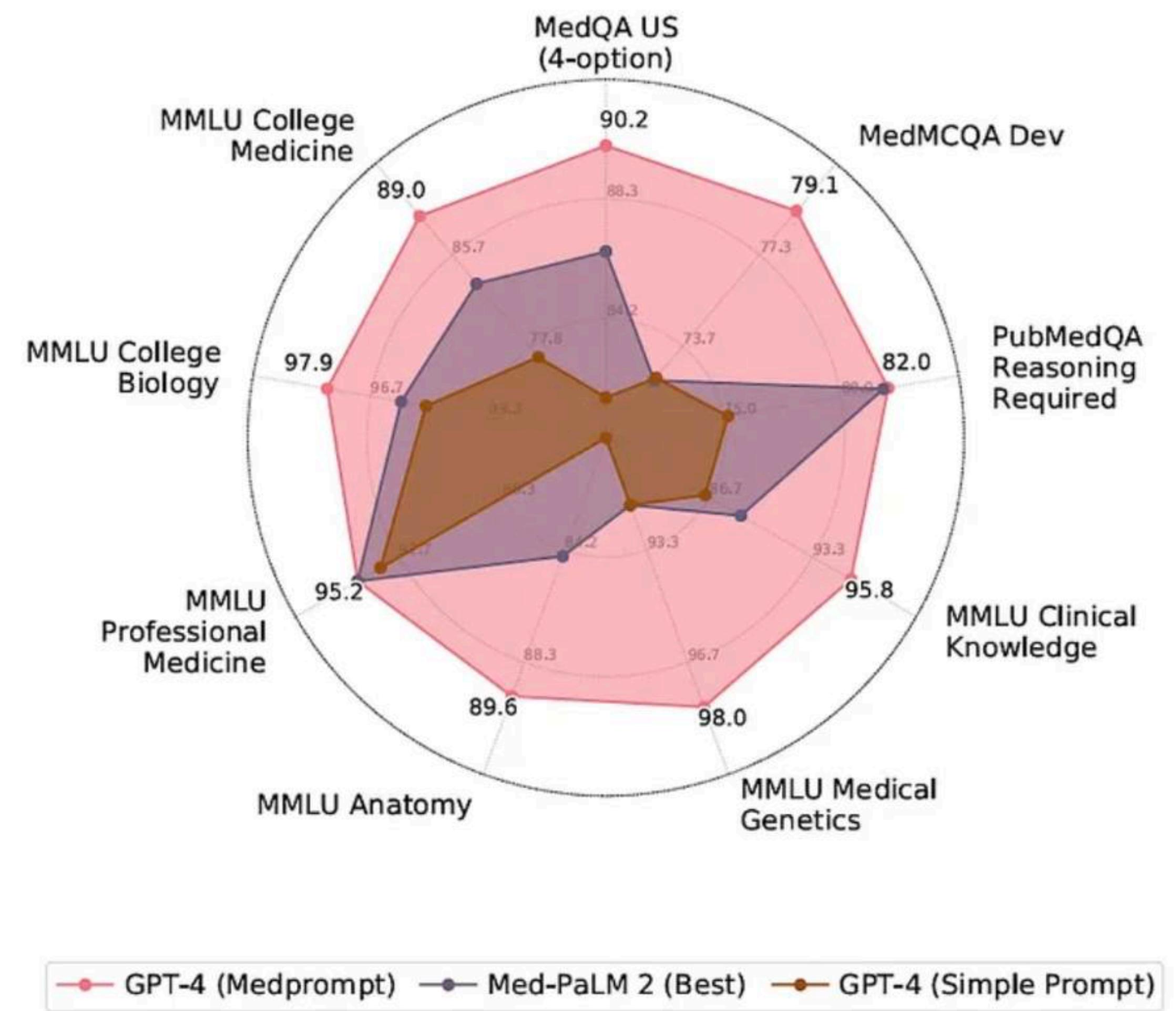
Bard: B+

GPT-3.5: B

<https://github.com/sikkha/PulsarWave/>

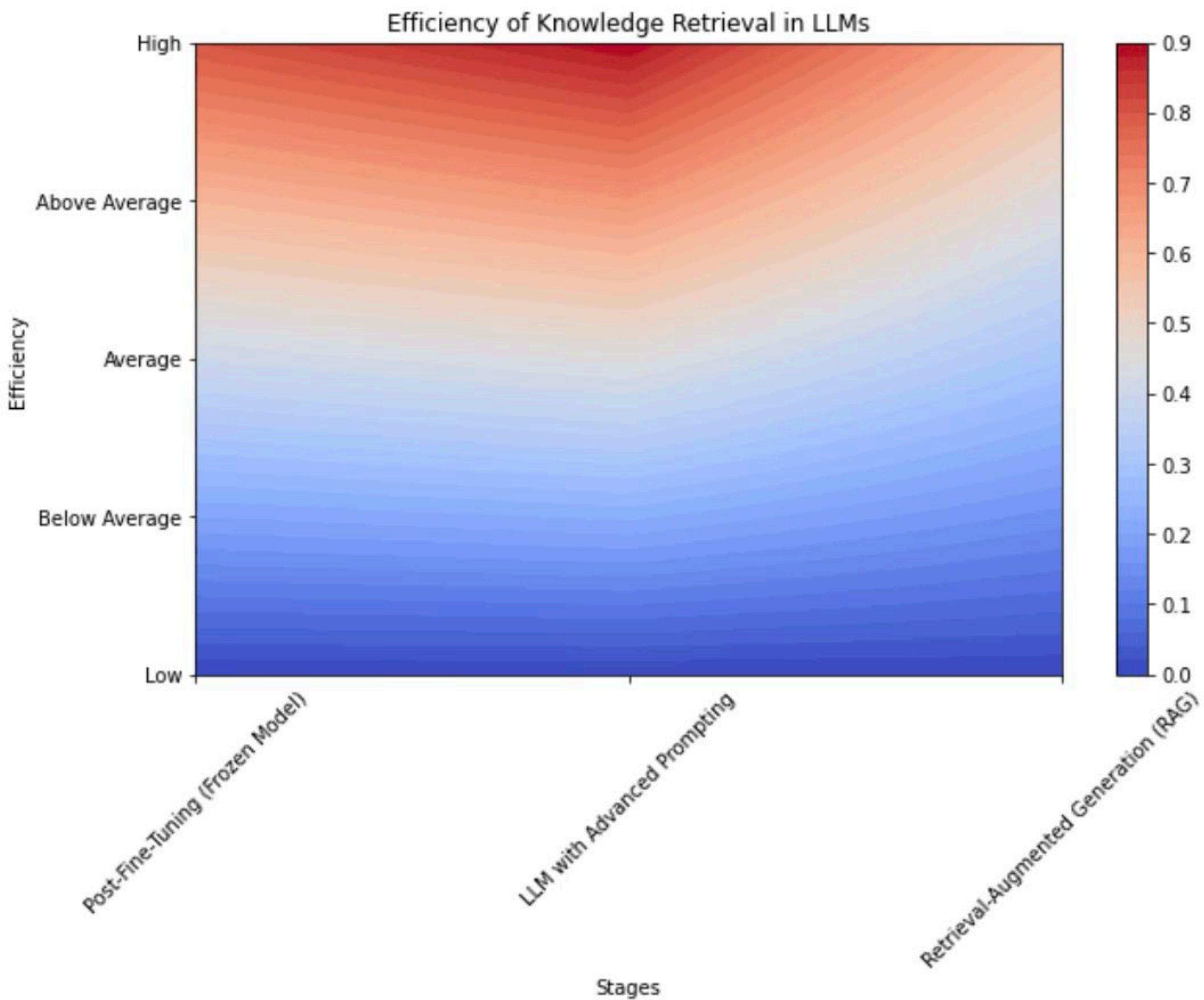


(a)



(b)

Figure 1: (a) Comparison of performance on MedQA. (b) GPT-4 with Medprompt achieves SoTA on a wide range of medical challenge questions.

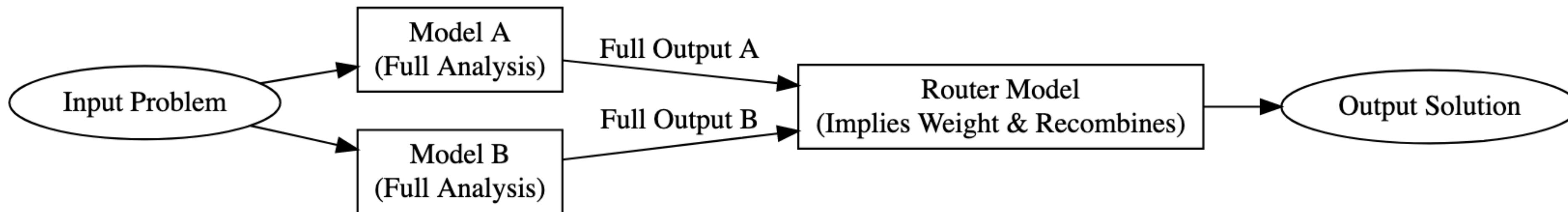
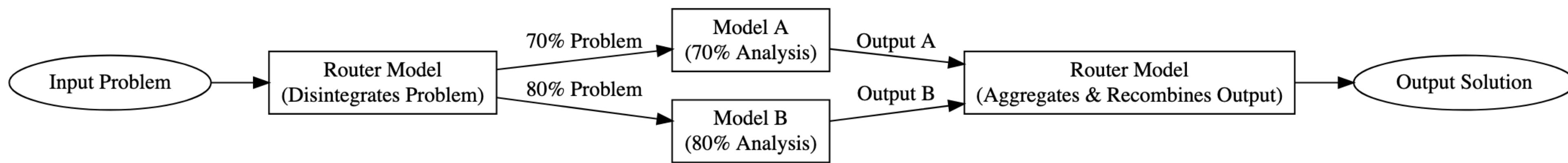




Satellite Reign - Official Launch Trailer

IGN

Criteria/Model	PaLM 2	ChatGPT (GPT-4)	LLaMa 2
Decision Style	Concise and bold decisions.	Balanced between operational mandates and emotional/ethical considerations.	In-depth multi-perspective analysis.
Operational Focus	Prioritizes operational directives over external personal details.	Balances operational mandates with personal details but leans toward protocol.	Evaluates operational protocols in light of personal details and potential strategic implications.
Information Processing	Processes information swiftly, filtering out deemed irrelevant details.	Considers a wider range of information but with a balanced approach.	Deeply analyzes all available information, attempting to understand broader ramifications of decisions.
Ethical Sensitivity	Less sensitive to extraneous emotional or ethical factors.	Balances ethical/emotional considerations with operational needs.	Highly sensitive to ethical and emotional nuances, prioritizing understanding and consideration of broader impacts.
Response to Ambiguity	Likely to default to established protocols.	May weigh ambiguities against established protocols, leaning toward the latter.	Delves deep into ambiguities, potentially recommending alternative actions or reviews.
Recommendations & Future Actions	Limited. Primarily focused on the task at hand.	May suggest post-action reviews based on certain triggers.	More likely to recommend broader operational reviews, adjustments, or alternative actions.
Potential for External Consideration (like Propaganda Risk)	Less likely to consider.	Might acknowledge but won't weigh heavily.	Actively considers potential externalities and their strategic implications.





END

AI Blue