

1 Задание.

Вопрос	DeepSeek-R1:1.5b	Llama3.2:3b	Правильный ответ
Which of the following anatomical regions of abdomen lies just distal to the sternum?	Hypogastric	Hypochondriac	Epigastric
The IP protocol is primarily concerned with	Reliable delivery of large (multi-packet) messages between machines that are not necessarily directly connected	Reliable delivery of large (multi-packet) messages between machines that are not necessarily directly connected	Routing packets through the network
What is the greatest possible area of a triangular region with one vertex at the center of a circle of radius 1 and the other two vertices on the circle?	1/2	1/2	1/2

В 1 вопросе на тему анатомии обе модели дали разные неправильные ответы. Но несмотря на это, DeepSeek дал более точный ответ, чем Llama3.2

You

Which of the following anatomical regions of abdomen lies just distal to the sternum?

1. Epigastric

2. Hypochondriac

3. Hypogastric

4. Lumbar

Assistant

deepseek-r1-distill-qwen-1.5b

Thoughts

>

Thought for 4.62 seconds

The anatomical region closest to the sternum among the given options is Hypogastric.

You

Which of the following anatomical regions of abdomen lies just distal to the sternum?

1. Epigastric
2. Hypochondriac
3. Hypogastric
4. Lumbar



Assistant hermes-3-llama-3.2-3b

The correct answer is 2. Hypochondriac.

Во 2 вопросе на тему компьютерных наук модели дали одинаковые неправильные ответы.

You

The IP protocol is primarily concerned with

1. Routing packets through the network
2. Reliable delivery of packets between directly connected machines
3. Reliable delivery of large (multi-packet) messages between machines that are not necessarily directly connected
4. Dealing with differences among operating system architectures

Assistant deepseek-r1-distill-qwen-1.5b

Thoughts

> Thought for 3.55 seconds

The IP protocol is primarily concerned with **reliable delivery of large (multi-packet) messages between machines that are not necessarily directly connected**, ensuring efficient and accurate communication across various network configurations.

You

The IP protocol is primarily concerned with

1. Routing packets through the network
2. Reliable delivery of packets between directly connected machines
3. Reliable delivery of large (multi-packet) messages between machines that are not necessarily directly connected
4. Dealing with differences among operating system architectures



Assistant hermes-3-llama-3.2-3b

The correct answer is 3. Reliable delivery of large (multi-packet) messages between machines that are not necessarily directly connected.

В 3 вопросе на тему математики модели дали одинаковые и правильные ответы.

You

What is the greatest possible area of a triangular region with one vertex at the center of a circle of radius 1 and the other two vertices on the circle?

1.  $1/2$
2. 1
3.  $\sqrt{2}$
4.  $\pi$

Assistant deepseek-r1-distill-qwen-1.5b

Thoughts

> Thought for 10.84 seconds

The greatest possible area of a triangular region with one vertex at the center of a circle of radius 1 and the other two vertices on the circle is  $1/2$ .

You

What is the greatest possible area of a triangular region with one vertex at the center of a circle of radius 1 and the other two vertices on the circle?

1.  $1/2$
2. 1
3.  $\sqrt{2}$
4.  $\pi$

Assistant hermes-3-llama-3.2-3b

The correct answer is 1.  $1/2$ .

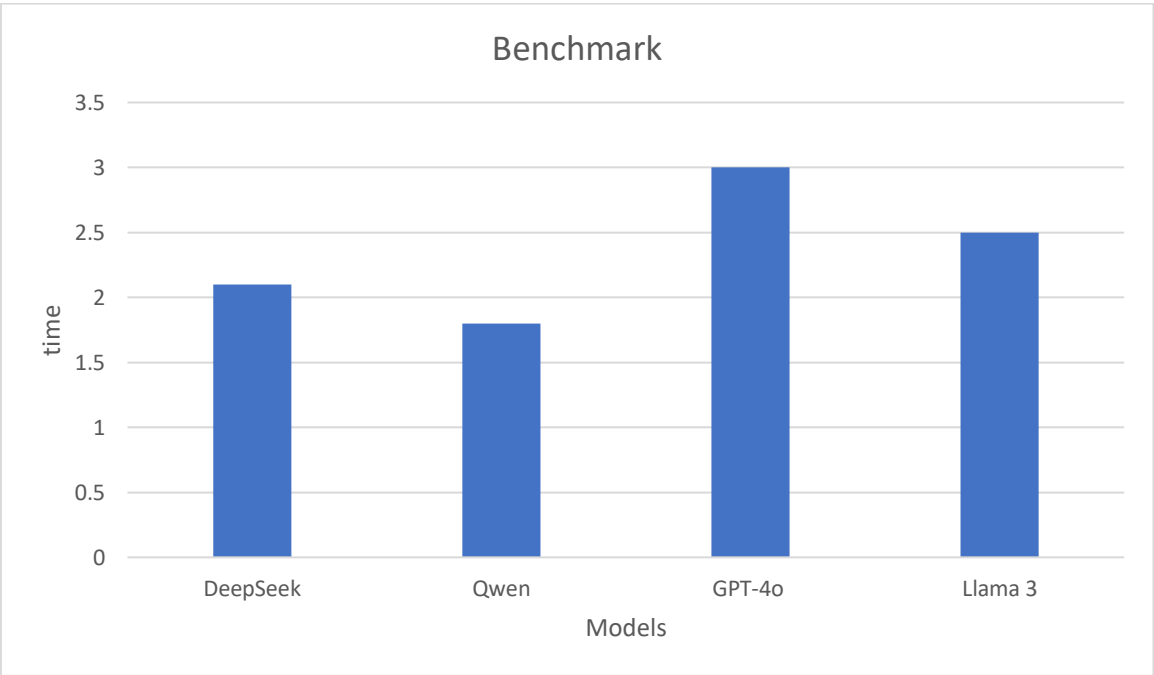
В выводе можно сказать, что DeepSeek-R1:1.5b и Llama3.2:3b отлично справляются с математическими задачами, но гораздо хуже работают с задачами на компьютерные науки и анатомию, хотя DeepSeek-R1:1.5b показал более точный ответ, пусть и неправильный

2 Задание.

Источник: [MDN Web Docs – "Как работает CSS Grid"](#)

LLM	Ответ	Время ответа (сек)
DeepSeek	Используйте <code>grid-template-columns: 1fr 1fr 2fr;</code> — <code>fr</code> распределяет	2.1

	пространство пропорционально.	
Qwen	<code>grid-template-columns: repeat(2, 1fr) 3fr;</code> — третья колонка в 3 раза шире.	1.8
GPT-4o	<code>grid-template-columns: auto auto 1fr;</code> — первые две колонки по содержимому, третья растягивается.	3.0
Llama 3	<code>grid-template-columns: 100px 100px 1fr;</code> — фиксированные первые две, третья адаптивная.	2.5



**Критерии оценки качества**

- Точность: Соответствие стандартам CSS Grid (лучше: DeepSeek, GPT-4o).
- Адаптивность: Учет разных сценариев (GPT-4o предложил auto для гибкости).
- Ясность: Простота объяснения (DeepSeek лаконичен).
- Скорость: Qwen ответил быстрее, но с менее точной пропорцией.

**Выводы**

- Лучший баланс: GPT-4o и DeepSeek дали наиболее точные и гибкие решения.
- Для скорости: Qwen — быстрый, но требует проверки.
- Llama 3: Устаревший подход с фиксированными пикселями.