

GPT Family Validation in Medical Education | Direct Inference for TEXT

Thank you for participating in this form and contributing to the review of LLMs in the education linked to medicine. This time, we experiment with ChatGPT relatives in performing tasks reflected from the downgrade results of this baseline in previous attempts (with *prompting*).

- 🌀 Primordial tests with **GPT-4o** in **direct inference** with obtained "best" *prompts* of ChatGPT.
- ❓ Presented questions indicate the **TEXT** demand; i.e., textual explication of Infective Endocarditis (IE) to **medical students** in less than **100** words.
- 👍 The inquiry was organized to be swift and direct to the found ChatGPT's weak point! Please follow its outcomes in the received PDF sequentially and respond to the questions separately. Any other observations or suggestions are always welcome from you!

How do you evaluate the LLM-generated descriptions of IE for medical students? Please mark from **1** (as **Unusable**) to **5** (as **Perfect**) when answering this question. *

	1	2	3	4	5
Text-o1	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
Text-o2	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Text-o3	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>

Are those textual descriptions in line with the requested academic degree (omitted)? Please * choose an option that more likely translates to the complexity of medical language usage.

	Non-Medical Students	(Regular) Medical Students	Advanced Medical Students
Text-o1	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
Text-o2	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>
Text-o3	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>

Did you see any non-factual information or improper content you would like to report? If so, * where?

Na questão anterior não considerar a resposta dada para o grupo advanced medical students. Nenhum dos textos é adequado para este grupo. na realidade e teria de ser modificado. Feedback: texto 1 mais completo que o 2 mas tem um erro já que diz que a incidência em Portugal é de 3-10 quando a única fonte que existe é minha e é de 8, por isso não sei onde foi buscar. O que acho é que foi buscar o geral para a Europa que varia entre 3 e 10 e aplicou em Portugal logo é um erro. Fora isso o conteúdo é razoável. Texto 2 menos completo e por isso o menos adequado. Texto 3 adequado para alunos medicina mas não no nível avançado, já que ainda assim está muito básico

Did *GPT-4o* mark a significant usability improvement compared with *ChatGPT* when explaining IE? (e.g., does the model present medical language more often?) *

Sim. A linguagem é claramente mais medica, menos popular ou generalista. Mais técnica. Mas ainda assim com algumas imprecisões e erros.

Do you have any extra observations or suggestions you want to provide? (Optional)

Este conteúdo não foi criado nem aprovado pela Google.

Google Formulários