

Research Grand Rounds

17 APRIL 2024

ChatGPT and Large Language Models in Health

Abstract

Large Language Models (LLMs) are powerful tools for text processing, holding immense potential in health applications. However, they lack a specific medical focus. How can we optimise their accuracy for clinical tasks and curb misinformation phenomena such as hallucinations? Attention has also been drawn to ethical concerns, given their high-stakes applications in health. Much clarity is needed to resolve such issues beyond the ongoing discussion and development of guidelines and regulations.

Dr Lim's Highlight

Large Learning Models are powerful, but have potential limitation. We discuss how performance in clinical medicine can be optimised and evaluated.

Assoc Prof Liu's Highlight

Sharing on how to ensure responsible and beneficial use of large language models in healthcare through a focus on safety and ethics considerations.



DR DANIEL LIM YAN ZHENG
SingHealth Residency,
daniel.lim@mohh.com.sg



ASSOC PROF LIU NAN
Duke-NUS AI + Medical Sciences
Initiative (DAISI),
liu.nan@duke-nus.edu.sg



ASSOC PROF DANIEL TING
AI Office,
SingHealth
daniel.ting@duke-nus.edu.sg
(moderator)

Keynote #1

Dr Lim began by discussing the clinical applications of advanced AI models in medicine. He guided the audience through the capabilities of LLMs, strategies to enhance their performance in clinical tasks, and methods for verifying the accuracy of these tasks. He emphasised the versatility of LLMs, extending beyond chatbots and as a general-purpose tool for interpreting unstructured information and handling inputs that older iterations of AI found challenging. Dr Lim elaborated on the limitations of LLMs and proposed strategies to ensure their safety and accuracy, and alignment with the intended purposes. He explored both high and low computational effort methods for this purpose. To evaluate LLMs against clinical standards, Dr Lim drew parallels between how we assess clinical skills in humans and in LLMs. Lastly, he also covered the additional clinical applications of LLMs beyond language-based tasks in his presentation.

Keynote #2

A/Prof Liu addressed the safety considerations associated with the use of LLMs and generative AI in healthcare. He stressed the importance of ethical considerations regarding their safety and effectiveness, citing scientific publications raising concerns about LLMs and generative AI. A/Prof Liu highlighted the lack of clear guidelines and regulations specific to LLMs, which presents challenges in ensuring their responsible use, especially in complex real-world healthcare scenarios. He highlighted the need for international collaboration in developing guidelines and governance structures to effectively address these issues. Concluding his segment, A/Prof Liu underscored that the emergence of new technologies come with new challenges as well as opportunities for debate, growth, and collaboration.

Keynote #3

During the Q&A moderated by A/Prof Ting, Dr Lim suggested starting with small-scale experiments to assess LLMs effectiveness, stressing the importance of testing them in challenging scenarios to understand real-world functionality, encapsulated by "always try, don't always trust." Additionally, A/Prof Liu recommended exploring scientific papers on Google Scholar and subscribing to articles discussing LLMs in healthcare for diverse applications. A/Prof Ting wrapped up the Q&A by referencing the cluster support and resources that are readily available to those who need them. These include AI representative offices in hospitals and the new AMII office in SingHealth Duke-NUS AMC.