Artificial Intelligence: Problem Solver or Causer in the Medical Field?

Jackson Vondemkamp

Introduction

The developments of Artificial Intelligence are often so vast that users and reporters fail to delve into niche topics such as the use cases in the medical field, but rather AI generated images used for fun. However, there is a growing argument for the usage of artificial intelligence to be used as an informative tool for medical patients. Is AI "on track" to performing as such? Will it eventually replace (or have the ability to replace) our medical professionals? This may be the upcoming reality the medical industry will be facing.

The Evolution of AI in Medicine

Artificial Intelligence is not *new* to the medical industry; in fact it has existed in the space for quite a while already. Prior uses have included enhancing research in drug discovery, and also medical imaging. Alongside these developments, a certain Language Learning Model (LLM) called ChatGPT has been reaching new development milestones that show it has new capabilities, specifically with medical recommendations. A LLM is an Artificial Intelligence model that allows for natural communication, and in this case it is dialogue being used to ask questions and get them answered. Most recently, ChatGPT-4 has shown a promising ability to allow for the "best" answers.

ChatGPT-4's Performance in Cancer Treatment Recommendations

A recent study by Tsai et al. (2024) analyzed the difference between performance between ChatGPT-4 and ChatGPT-3.5 in giving cancer treatment recommendations for urological cancers, including prostate, bladder, and kidney cancers. This research had a goal of evaluating the quality, adherence to clinical guidelines, and concordance with expert opinions of the AI-generated recommendations.

Methodology of the Study

The researchers designed 108 clinical scenarios surrounding the three priorly mentioned cancers. Both ChatGPT-4 and ChatGPT-3.5 were asked to provide treatment recommendations for the scenarios given. Then, the recommendations were assessed using the following methods:

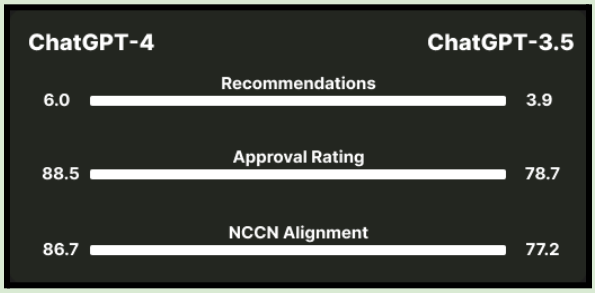
1. **Recommendation Count:** The number of treatment recommendations given in one response.
2. **Approval Rate:** Percentage of recommendations that experts approved of.
3. **NCCN Concordance:** How much the recommendation aligns with the

National Comprehensive Cancer Network's (NCCN) guidelines.

Key Findings

The study revealed many interesting findings in the behavior of the two ChatGPT models giving medical recommendations:

- GPT-4 offered around **two more ideas per case**, beating GPT-3.5 in quantity.
- Oncologists **approved significantly more suggestions** from GPT-4, around 10 percent.
- GPT-4 **stuck to NCCN guidelines more often** than GPT-3.5 by around 10 percent.



Use Cases Currently

Since LLMs like ChatGPT are certainly not 100% accurate, it is not recommended that they be used on their own. However, the promising performance of them reveal other ways that they can be used *right now*. For example, and also pointed out in the article, working side by side with these chatbots can allow for a new in person perspective that is also able to be corrected. This could essentially function as a doctor with an assistant, where the doctor supplies the main recommendation and the "assistant" gives side recommendations. Additionally, what is true for all medical fields but especially true for cancer patients, is that there is always a wait for the recommendations that are non-emergency. A possible use case of ChatGPT is to supply these recommendations quickly and easily. This is great for issues such as pain management, and other questions that exist as public knowledge on the internet already. To further deal with the problems of wait times for these patients, ChatGPT can be implemented into the electronic health records that already exist. Imagine how much more diagnosing and treatment can be done if all it took to document something was to tell ChatGPT to "record" the details. By streamlining these processes, more patients can be seen and overall more people can benefit from a doctor's care.

Will AI Completely Replace Medical Professionals?

As it stands, LLMs such as ChatGPT *cannot* replace the important therapies that patients such as cancer patients go through. Since ChatGPT is just a LLM, it cannot supply the care associated with its recommendations and is therefore not sufficient in working just by itself. But even if they were somehow able to do so, with the power of robotics, there are necessary jobs that doctors perform besides the treatment. Doctors also supply mental support and reassurance that Artificial Intelligence just

cannot replace in the real world. Doctors are also well informed on the ethical boundaries that surround their everyday work. Simply put, not all recommendations can be followed through because they rely on the fact that patients would listen to what is recommended. Concepts such as Do Not Resuscitate (DNR), are things that are put in place to protect the patients rights to their own autonomy, and Artificial Intelligence would not be able to make the ethical calls that doctors have to make everyday.
9538

Source

Tsai C-Y, Cheng P-Y, Deng J-H, Jaw F-S, Yii S-C. *ChatGPT v4 outperforming v3.5 on cancer treatment recommendations in quality, clinical guideline, and expert opinion concordance*. DIGITAL HEALTH. 2024. https://journals.sagepub.com/doi/full/10.1177/2055207624126