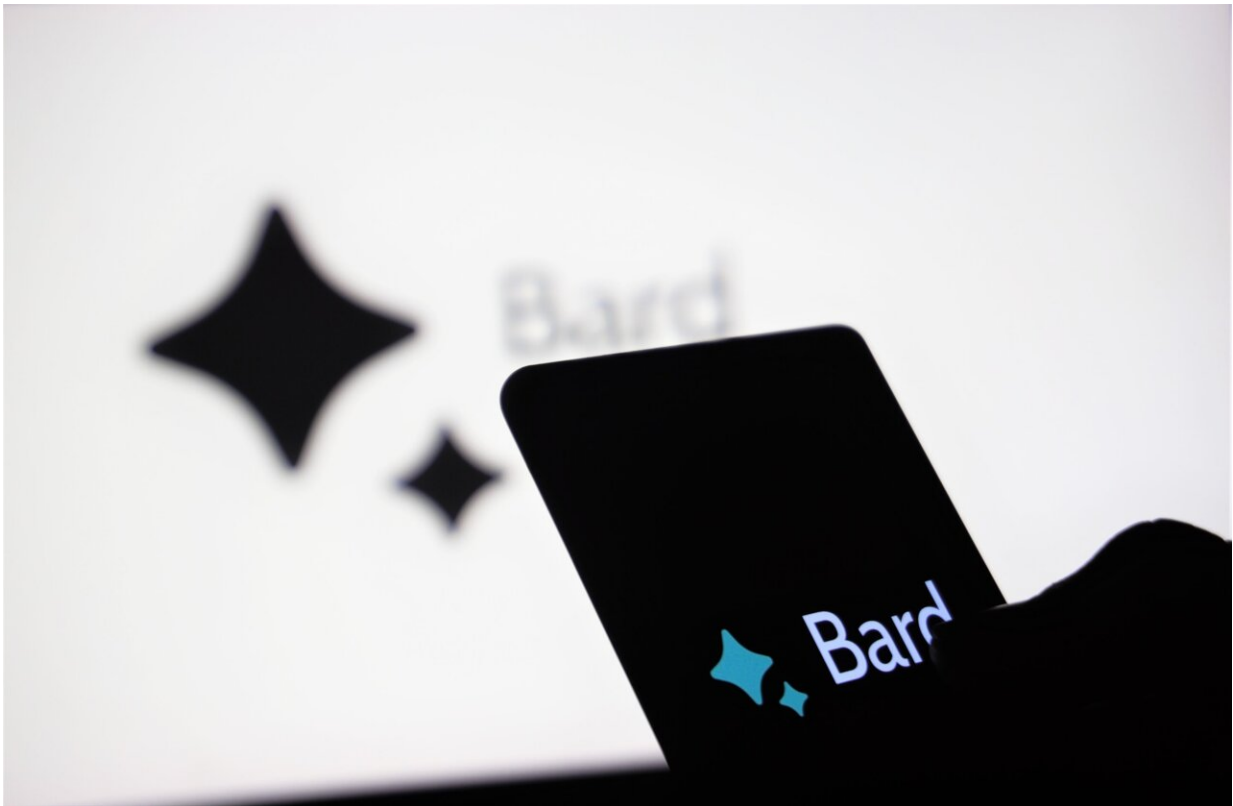# Google AI health chatbot passes US medical exam: study

July 13 2023

Credit: Unsplash/CC0 Public Domain

Google's artificial intelligence-powered medical chatbot has achieved a passing grade on a tough US medical licensing exam, but it's answers still fall short of those from human doctors, a peer-reviewed study said on Wednesday.

Last year the release of ChatGPT—whose developer OpenAI is backed by Google's rival Microsoft—kicked off a race between tech giants in the burgeoning field of AI.

While much has been made about the future possibilities—and dangers—of AI, health is one area where the technology had already shown tangible progress, with algorithms able to read certain medical scans as well as humans.

Google first unveiled its AI tool for answering medical questions, called Med-PaLM, in a preprint study in December. Unlike ChatGPT, it has not been released to the public.

The US tech giant says Med-PaLM is the first large language model, an AI technique trained on vast amounts of human-produced text, to pass the US Medical Licensing Examination (USMLE).

A passing grade for the exam, which is taken by medical students and physicians-in-training in the United States, is around 60 percent.

In February, a study said that ChatGPT had achieved passing or near passing results.

In a peer-reviewed study published in the journal Nature on Wednesday, Google researchers said that Med-PaLM had achieved 67.6 percent on USMLE-style multiple choice questions.

"Med-PaLM performs encouragingly, but remains inferior to clinicians," the study said.

To identify and cut down on "hallucinations"—the name for when AI models offer up false information—Google said it had developed a new evaluation benchmark.

Karan Singhal, a Google researcher and lead author of the new study, told AFP that the team has used the benchmark to test a newer version of their model with "super exciting" results.

Med-PaLM 2 has reached 86.5 percent on the USMLE exam, topping the previous version by nearly 20 percent, according to a preprint study released in May that has not been peer-reviewed.

'Elephant in the room'

James Davenport, a computer scientist at the UK's University of Bath not involved in the research, said "there is an elephant in the room" for these AI-powered medical chatbots.

There is a big difference between answering "medical questions and actual medicine," which includes diagnosing and treating genuine health problems," he said.

Anthony Cohn, an AI expert at the UK's Leeds University, said that hallucinations would likely always be a problem for such large language models, because of their statistical nature.

Therefore these models "should always be regarded as assistants rather than the final decision makers," Cohn said.

Singhal said that in the future Med-PaLM could be used to support doctors to offer up alternatives that may not have been considered otherwise.

The Wall Street Journal reported earlier this week that Med-PaLM 2 has been in testing at the prestigious US Mayo Clinic research hospital since April.

Singhal said he could not speak about specific partnerships.

But he emphasized that any testing would not be "clinical, or patient facing, or are able to cause patients harm".

It would instead be for "more administrative tasks that can be relatively easily automated, with low stakes," he added.

© 2023 AFP

Citation: Google AI health chatbot passes US medical exam: study (2023, July 13) retrieved 13 May 2024 from https://techxplore.com/news/2023-07-google-ai-health-chatbot-medical.html