
Munjal Shah: Hippocratic AI Partnering With Healthcare Providers to Build its Large Language Model



CEO Munjal Shah and Hippocratic AI are working to ensure healthcare LLMs have expert feedback and evidence-based training sets.

Hippocratic AI is a start-up designing a large language model for non-diagnostic healthcare applications. Despite the hype surrounding the recent boom in artificial intelligence, which can now convincingly engage in conversations and answer questions on a dizzying array of topics, there remain concerns about the risks of applying LLMs in some spaces. After all, these models aren't infallible. Despite high user adoption rates, most nonexperts remain baffled as to how AI can come up with such accurate answers on topics that seem to require human experience and knowledge. As a result, many remain sceptical of trusting LLMs.

For Munjal Shah, the co-founder and CEO of Hippocratic AI, the response to this problem is twofold. First, demonstrate the effectiveness of LLMs in low-risk applications so providers and patients are willing to engage and eventually understand that these models can help solve important problems. Second, highlight the role of human expertise in building LLMs in the healthcare space so that the public can better understand where the model's knowledge comes from and trust in its safety.

In a recent interview on "The Mad Podcast with Matt Turck", Shah emphasised the importance of a "bottom-up" approach to establishing safe LLMs. This strategy involves experts providing feedback on AI responses and ultimately deciding when the model is equivalent to — or better than — a human providing medical services.

"You have two choices: You can use LLM A or LLM B," said Shah. "LLM A follows the standards put out by a governmental body that says it'll make it safe, and LLM B had a thousand chronic care nurses [perform quality assurance on] the chronic care nurse functionality, and they gave it feedback, and it was only released when the majority of them said it was safe in a blind test".

Shah says the best outcome is to harness both LLM A and LLM B.

"Would you want to use an LLM that only follows top-down safety standards or one that does that and has 1,000 nurses or clinical professionals

reviewing it? Why aren't we using the experts who do the job today, who know the job best, as our way of determining safety?"

Reinforcement Learning From Human Feedback

The foundation of this bottom-up approach is what's known as reinforcement learning from human feedback. RLHF involves refining an LLM's outputs based on human feedback. Essentially, humans review and rate the responses generated by the model, and these ratings are used to teach the model what constitutes a good or a desirable response. The model learns to predict responses that would likely receive positive feedback, thereby improving its performance and accuracy over time.

In the case of Hippocratic AI, the goal is to use experienced healthcare professionals to provide the RLHF for its model. In chronic care nursing, for instance, an LLM could learn to provide more empathetic, informed responses to patients by drawing on the real-world experience of chronic care nurses' evaluations of its responses.

Feedback is designed to guide the model to respond in ways that are both supportive and informative, enhancing patient engagement and adherence to treatment plans. This approach improves the model's ability to provide personalised and contextually relevant assistance to a variety of clients, as it draws on a wide range of real-world healthcare worker experiences and expertise.

"We're partnering with the health systems first. We did not just release this. We did not just build [the LLM], then take it to them," Munjal Shah told Turck. "We are literally having them work with us in joint development sessions every minute, and we believe they understand better than anybody how to deliver health care safely. We'll release it when they tell us it's safe, really, and by they, I mean their managerial teams, but really, I mean their bottom-up teams — when all the nurses at that health system say this is safe".

Evidence-Based Training

While RLHF can provide invaluable feedback on how to empathically communicate accurate information to patients, it doesn't operate in a vacuum. Feedback from thousands of humans can help AI refine its knowledge and communicative skills, but a system first needs to develop a knowledge base — and in the case of large language models, this process requires learning from unstructured data such as text, images, or audio.

LLMs rely on neural networks to predict appropriate linguistic responses. Similar to the human brain, these networks strengthen connections when prediction leads to accurate pattern recognition. However, unlike a human brain, a generative AI neural network's method of pattern recognition isn't hardwired into an evolved structure but rather is built from the ground up based on simple inputs and outputs. For this process to work, they must be fed vast volumes of training data.

When Munjal Shah began to think about how an LLM could best be trained to perform non-diagnostic healthcare services, he turned to evidence-based research as the ideal source. If LLMs learn to generate responses based on predicting a fit with a training set, why not train a healthcare LLM to try to fit a set of research and solutions developed and peer-reviewed by medical professionals?

Unlike popular LLM chatbots such as ChatGPT, the Hippocratic AI LLM doesn't crawl the entire internet — full of facts and misinformation — for its training data. It's instead fed evidence-based medical data that's not necessarily readily available on the internet.

"Health care data is very different from other data". said Shah. "Imagine an iceberg floating in the water. Everything below the line is behind a firewall, and everything above the water line is on the internet. Health care looks like an iceberg. Most of the data is below the water line".

He explained that Hippocratic AI has had to build its LLM using its own data and, as a result, partner with healthcare systems to provide it with the data it needs to work better for health-specific functions.

"We've instruction-tuned it better, and we've actually gone in and acquired 1.5 trillion health care tokens", he said, referring to the units of text LLMs use to understand and generate language. "It's pre-trained differently on health care content. We just announced ten health systems that partnered with us".

Hippocratic AI hopes this combination of expert-based human feedback and health-specific data will result in an effective healthcare LLM that could have a significant impact on patient outcomes and staffing shortages. But attention to safety remains paramount, as evidenced in the company's tagline of "Do No Harm". Safety measures will ultimately dictate when the company decides to release its LLM, said Shah, but [he believes](#) that by focusing on non-diagnostic applications, safety isn't an insurmountable barrier.

"I think most people who build LLMs just did not have enough experience in the health care world to understand there are so many other problems to be solved that LLMs can solve".

This article is part of the HealthManagement.org Point-of-View Programme.

Published on : Tue, 19 Dec 2023