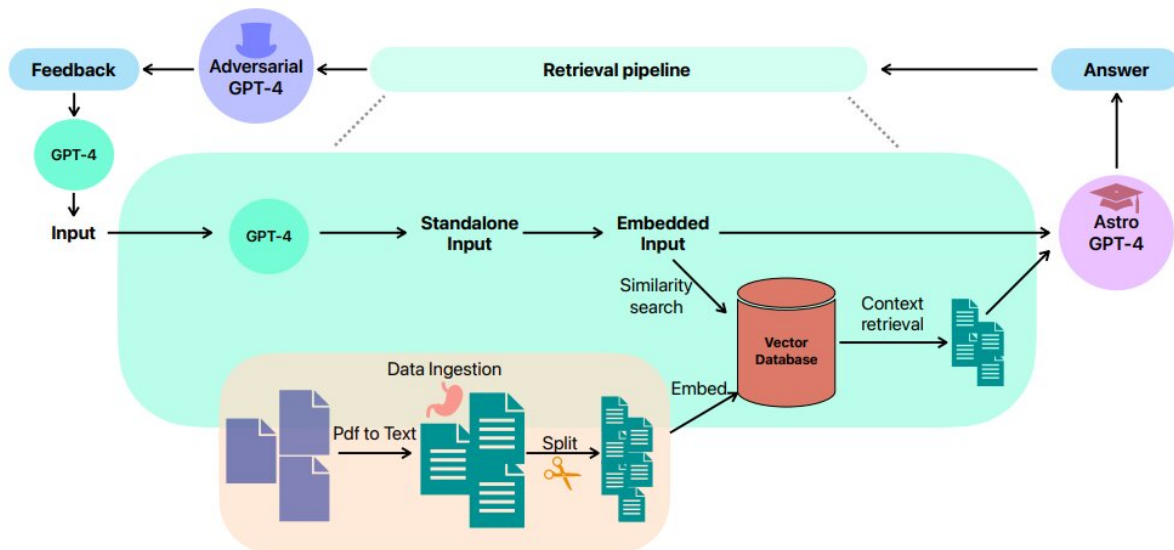


AI could help astronomers rapidly generate hypotheses

June 28 2023, by Andy Tomaswick



The adversarial in-context prompting workflow using OpenAI’s GPT-4 model. The procedure begins with the pre-processing and embedding of Galactic Astronomy papers. A similarity search is conducted on the embedded query, and relevant document chunks are retrieved. A further contextual compression is performed to remove irrelevant information from the chunks. These compressed texts serve as input to a GPT-4 instance, which generates an idea. This idea is then critiqued by a second GPT-4 model, and the feedback is moderated by a third GPT-4 model. Credit: *arXiv* (2023). DOI: 10.48550/arxiv.2306.11648

Almost anywhere you go on the internet, it seems nearly impossible to

escape articles on AI. Even here at UT, we've published several. Typically they focus on how a specific research group leveraged the technology to make sense of reams of data. But that sort of pattern recognition isn't all that AI is good for. In fact, it's becoming pretty capable of abstract thought. And one place where abstract thought can be helpful is in developing new scientific theories. With that thought in mind, a team of researchers from ESA, Columbia, and the Australian National University (ANU) utilized an AI to come up with scientific hypotheses in astronomy.

Specifically, they did so in the sub-field of galactic astronomy, which specializes in research surrounding the formation and physics of galaxies. A recently published paper on the *arXiv* pre-print server mentions that they selected this sub-field because of its "integrative nature," which requires "knowledge from diverse subfields."

That sounds exactly like what AI is already good at. But a standard [large language model](#) (LLM) like those that have become most familiar recently (ChatGPT, Bard, etc.) wouldn't have enough subject knowledge to develop reasonable hypotheses in that field. It might even fall prey to the "hallucinations" that some researchers (and journalists) warn are one of the downsides of interacting with the models.

To avoid that problem, the researchers, led by Ioana Ciucă and Yuan-Sen Ting of ANU, used a piece of code known as an [application programming interface](#) (API), which was written in Python, known as Langchain. This API allows more advanced users to manipulate LLMs like GPT-4, which serves as the latest basis for ChatGPT. In the researchers' case, they loaded over 1,000 scientific articles relating to galactic astronomy into GPT-4 after downloading them from NASA's Astrophysics Data System.

One of the researchers' experiments was to test how the number of

papers the model had access to affected its resulting hypotheses. They noticed a significant difference between the suggested hypotheses it developed having access to only ten papers vs. having access to the full thousand.

But how did they judge the validity of the hypotheses themselves? They did what any self-respecting scientist would do and recruited experts in the field. Two of them, to be precise. And they asked them to judge the hypotheses based on originality of thought, the feasibility of testing the hypotheses, and the scientific accuracy of its basis. The experts found that, even with a limited data set of only ten papers to go off of, the hypotheses suggested by Astro-GPT, as they called their model, were graded only slightly lower than a competent Ph.D. student. With access to the full 1,000 papers, Astro-GPT scored at a "near-expert level."

A critical factor in determining the final hypotheses that were presented to the experts was that the hypotheses were refined using "adversarial prompting." While this sounds aggressive, it simply means that, in addition to the program that was developing the hypotheses, another program was trained on the same data set and then provided feedback to the first program about its [hypotheses](#), thereby forcing the original program to improve their logical fallacies and generally create substantially better ideas.

Even with the adversarial feedback, there's no reason for astronomy Ph.D. students to give up on coming up with their own unique ideas in their field. But, this study does point to an underutilized ability of these LLMs. As they become more widely adopted, scientists and laypeople can leverage them more and more to come up with new and better ideas to test.

More information: Ioana Ciucă et al, Harnessing the Power of Adversarial Prompting and Large Language Models for Robust

Hypothesis Generation in Astronomy, *arXiv* (2023). [DOI: 10.48550/arxiv.2306.11648](https://doi.org/10.48550/arxiv.2306.11648)

Provided by Universe Today

Citation: AI could help astronomers rapidly generate hypotheses (2023, June 28) retrieved 28 April 2024 from <https://phys.org/news/2023-06-ai-astronomers-rapidly-generate.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.