

# What can AI learn about the universe?

May 3 2024, by Matt Williams

---



Illustration of an active quasar. New research shows AI can identify and classify them. Credit: ESO/M. Kornmesser

Artificial intelligence and machine learning have become ubiquitous, with applications ranging from data analysis, cybersecurity, pharmaceutical development, music composition, and artistic renderings.

In recent years, [large language models](#) (LLMs) have also emerged, adding human interaction and writing to the long list of applications.

This includes ChatGPT, an LLM that has had a profound impact since it was introduced less than two years ago. This application has sparked considerable debate (and controversy) about AI's potential uses and implications.

Astronomy has also benefitted immensely, where [machine learning](#) is used to sort through massive volumes of data to look for signs of planetary transits, correct for atmospheric interference, and find patterns in the noise. According to an international team of astrophysicists, this may just be the beginning of what AI could do for astronomy.

In a recent study, the team fine-tuned a Generative Pre-trained Transformer (GPT) model using observations of astronomical objects. In the process, they successfully demonstrated that GPT models can effectively assist with [scientific research](#).

The study was conducted by the International Center for Relativistic Astrophysics Network (ICRANet), an [international consortium](#) made up of researchers from the International Center for Relativistic Astrophysics (ICRA), the National Institute for Astrophysics (INAF), the University of Science and Technology of China, the Chinese Academy of Sciences Institute of High Energy Physics (CAS-IHEP), the University of Padova, the Isfahan University of Technology, and the University of Ferrara.

Their [paper](#), "Can AI Understand Our Universe? Test of Fine-Tuning GPT by Astrophysical Data," was recently posted to the *arXiv* preprint server.

As mentioned, astronomers rely extensively on machine learning algorithms to sort through the volumes of data obtained by modern telescopes and instruments. This practice began about a decade ago and has since grown by leaps and bounds to the point where AI has been

integrated into the entire research process. As ICRA President and the study's lead author Yu Wang told Universe Today via email:

"Astronomy has always been driven by data and astronomers are some of the first scientists to adopt and employ machine learning. Now, machine learning has been integrated into the entire astronomical research process, from the manufacturing and control of ground-based and space-based telescopes (e.g., optimizing the performance of adaptive optics systems, improving the initiation of specific actions (triggers) of satellites under certain conditions, etc.), to data analysis (e.g., noise reduction, data imputation, classification, simulation, etc.), and the establishment and validation of theoretical models (e.g., testing modified gravity, constraining the equation of state of neutron stars, etc.)."

Data analysis remains the most common among these applications since it is the easiest area where machine learning can be integrated. Traditionally, dozens of researchers and hundreds of citizen scientists would analyze the volumes of data produced by an observation campaign.

However, this is not practical in an age where modern telescopes are collecting terabytes of data daily. This includes all-sky surveys like the Very Large Array Sky Survey (VLASS) and the many phases conducted by the Sloan Digital Sky Survey (SDSS).

To date, LLMs have only been applied sporadically to astronomical research, given that they are a relatively recent creation. But according to proponents like Wang, it has had a tremendous societal impact and has a lower-limit potential equivalent to an "Industrial Revolution."

As for the upper limit, Wang predicts that that could range considerably and could perhaps result in humanity's "enlightenment or destruction." However, unlike the Industrial Revolution, the pace of change and

integration is far more rapid for AI, raising questions about how far its adoption will go.

To determine its potential for the field of astronomy, said Wang, he and his colleagues adopted a pre-trained GPT model and fine-tuned it to identify astronomical phenomena:

"OpenAI provides pre-trained models, and what we did is fine-tuning, which involves altering some parameters based on the original model, allowing it to recognize astronomical data and calculate results from this data. This is somewhat like OpenAI providing us with an undergraduate student, whom we then trained to become a graduate student in astronomy.

"We provided limited data with modest resolution and trained the GPT fewer times compared to normal models. Nevertheless, the outcomes are impressive, achieving an accuracy of about 90%. This high level of accuracy is attributable to the robust foundation of the GPT, which already understands data processing and possesses logical inference capabilities, as well as communication skills."

To fine-tune their model, the team introduced observations of various astronomical phenomena derived from various catalogs. This included 2,000 samples of quasars, galaxies, stars, and broad absorption line (BAL) quasars from the SDSS (500 each). They also integrated observations of short and long gamma-ray bursts (GRBs), galaxies, stars, and black hole simulations. When tested, their model successfully classified different phenomena, distinguished between types of quasars, inferred their distance based on redshift, and measured the spin and inclination of black holes.

"This work at least demonstrates that LLMs are capable of processing astronomical data," said Wang. "Moreover, the ability of a model to

handle various types of astronomical data is a capability not possessed by other specialized models. We hope that LLMs can integrate various kinds of data and then identify common underlying principles to help us understand the world. Of course, this is a challenging task and not one that astronomers can accomplish alone."

Of course, the team acknowledges that the dataset they experimented with was very small compared to the data output of modern observatories. This is particularly true of next-generation facilities like the Vera C. Rubin Observatory, which recently received its LSST camera, the largest digital camera in the world!

Once Rubin is operational, it will conduct the 10-year Legacy Survey of Space and Time (LSST), which is expected to yield 15 terabytes of data per night! Satisfying the demands of future campaigns, says Wang, will require improvements and collaboration between observatories and professional AI companies.

Nevertheless, it's a foregone conclusion that there will be more LLM applications for astronomy in the near future. Not only is this a likely development, but a necessary one considering the sheer volumes of data astronomical studies are generating today. And since this is likely to increase exponentially in the near future, AI will likely become indispensable to the field of study.

**More information:** Yu Wang et al, Can AI Understand Our Universe? Test of Fine-Tuning GPT by Astrophysical Data, *arXiv* (2024). [DOI: 10.48550/arxiv.2404.10019](https://doi.org/10.48550/arxiv.2404.10019)

Provided by Universe Today

Citation: What can AI learn about the universe? (2024, May 3) retrieved 26 June 2024 from <https://phys.org/news/2024-05-ai-universe.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.