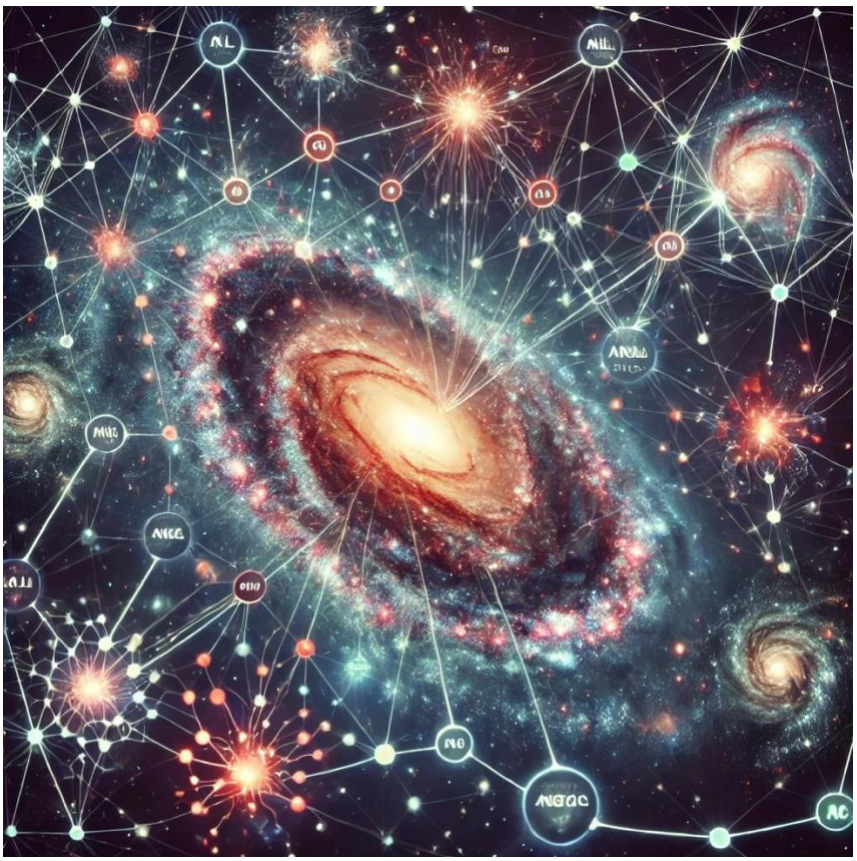
A new study explores the large-scale distribution of galaxies through Interpretable Machine Learning

- Using innovative techniques to make Artificial Intelligence more transparent, this study published in *Physical Review Letters* unveils how neural networks learn from the large-scale distribution of galaxies—demystifying a process that was once considered a "black box."
- The work is led by the Institute for Theoretical Physics (IFT) UAM-CSIC and the University of Chile.



Generated by ChatGPT.

Madrid, February 10, 2025.– Modern cosmology faces several fundamental challenges, such as understanding the nature of dark matter and dark energy, the large-scale structure of the Universe, and probing the early moments of cosmic evolution. This is linked to determining which theoretical model most accurately describes the evolution of the Universe. In a new study, published in the prestigious journal *Physical Review Letters*, a team of physicists has introduced **a key innovation in the application of machine learning to cosmology.** They used interpretability techniques for machine learning, which provide insights into how neural networks make predictions.

The research, conducted by Indira Ocampo, George Alestas, and Savvas Nesseris from the Institute for Theoretical Physics (IFT) UAM-CSIC, along with Domenico Sapone from the University of Chile, shows that the use of neural networks can enhance the analysis of observational (or simulated) data to test models beyond the standard cosmological model (ΛCDM). In this particular case, **to distinguish between ΛCDM and alternative modified gravity models**, such as the Hu-Sawicki f(R) model. More importantly, through the use of interpretable ML tools, they were able to see what the neural network is learning from the data, shedding light on why it was able to classify correctly between the two models. **This interpretability aspect is crucial** in gaining insights into the underlying mechanisms driving the classification process, but also to understand the physics behind "more important regions of data".

Beyond the Standard Model

The ΛCDM (Lambda Cold Dark Matter) model has been the dominant reference for explaining the evolution of the Universe. This model successfully describes the accelerated expansion of the Universe, the formation of large-scale structures, and the properties of the cosmic microwave background radiation. However, it presents some discrepancies with recent observations, such as the determination of the Hubble constant, which describes the rate of expansion of the Universe, and anomalies in the distribution of matter on large scales, particularly irregularities observed in how matter is distributed across the Universe, especially at vast distances. An alternative that is under exploration to solve these discrepancies is studying **beyond ΛCDM models.**

The ΛCDM model is the standard in cosmology, explaining the accelerated expansion of the Universe through dark energy (represented by the cosmological constant Λ) and cold dark matter (CDM). On the other hand, an interesting alternative class of models are the so-called f(R) models, which modify Einstein's theory of general relativity, the foundation of our understanding of gravity. At smaller scales that are within human reach (for example the Solar System), the f(R) models can recover General Relativity, but can also mimic dark energy or dark matter by altering Einstein's equations on cosmological scales.

The Revolution of Machine Learning in Cosmology

Testing new ideas about the universe usually involves comparing predictions from different models with actual observations. Recently, machine learning (ML) has gained a lot of attention either to speed up complex calculations or to help categorize different astronomical objects, showing amazing results. However, some concerns have arisen in the scientific community because **it is not always clear how these computer tools make their decisions.**

The method implemented by the researchers from the Institute for Theoretical Physics in Madrid and the University of Chile used artificial intelligence to analyze simulated data from the large-scale distribution of galaxies and successfully distinguished between the two cosmological models: ΛCDM and the f(R) model, with a very high accuracy. But more importantly, **to address the challenges of transparency, they have turned to interpretable Machine Learning techniques.** In particular, they used **LIME** (Local Interpretable Model-agnostic Explanations), a methodology that allows understanding which features of the data have a greater influence on the predictions made by the neural network. Physicists find this crucial for decision-making when validating any new theoretical approach, as Indira Ocampo, co-author of the study, explains: "Most of the methods known to date were developed due to their growing urgency in fields such as medicine, economics, and earth sciences. **In cosmology, interpretability is equally important**, as we rely on machine learning models to analyze vast and complex datasets, such as large-scale galaxy distributions or fluctuations in the cosmic microwave background".

A New Horizon for Computational Cosmology

The use of interpretable machine learning tools not only improves the accuracy of cosmological model selection but also **lays the foundation for future applications in the exploration of the Universe**. Even more crucial, are interpretability tools that can help us enhance our understanding of the fundamental physics behind the cosmological phenomena. As galaxy surveys and other astronomical observations generate increasingly large volumes of data, these techniques will be essential for extracting relevant information and advancing our understanding of the cosmos.

Ocampo, I., Alestas, G., Nesseris, S., & Sapone, D. (2025). Enhancing cosmological model selection with interpretable machine learning. *Physical Review Letters, 134*(4), 041002. https://doi.org/10.1103/PhysRevLett.134.041002

IFT

The Institute for Theoretical Physics (IFT) UAM-CSIC was officially created in 2003 as a joint research center belonging to the Spanish National Research Council (CSIC) and the Autonomous University of Madrid (UAM). It is the only Spanish center dedicated entirely to research in Theoretical Physics. The IFT members develop research in the frontiers of Elementary Particle Physics, Astroparticles and Cosmology, in order to understand the fundamental keys of Nature and the Universe. They are also leading many research projects, both at the national and international level. The IFT is part of the strategic line `Theoretical Physics and Mathematics´ of the Campus of International Excellence (CEI) UAM+CSIC established in 2009. Since 2012, it is credited as Severo Ochoa Centre of Excellence. Besides purely scientific activity, in the IFT is also conducted intensive training tasks of young researchers and professionals through the graduate program in Theoretical Physics with mention of excellence from the CEI and the Ministry of Education. In addition, the Institute carries out the important task of transferring knowledge to society through several outreach programs.

For more information and interviews, please contact:

Laura Marcos Mateos
laura.marcos@csic.es
comunicacion@ift.csic.es
912999879