UNIFYING KNOWLEDGE WITH AI IN EXPLORATION OF UNIVERSE

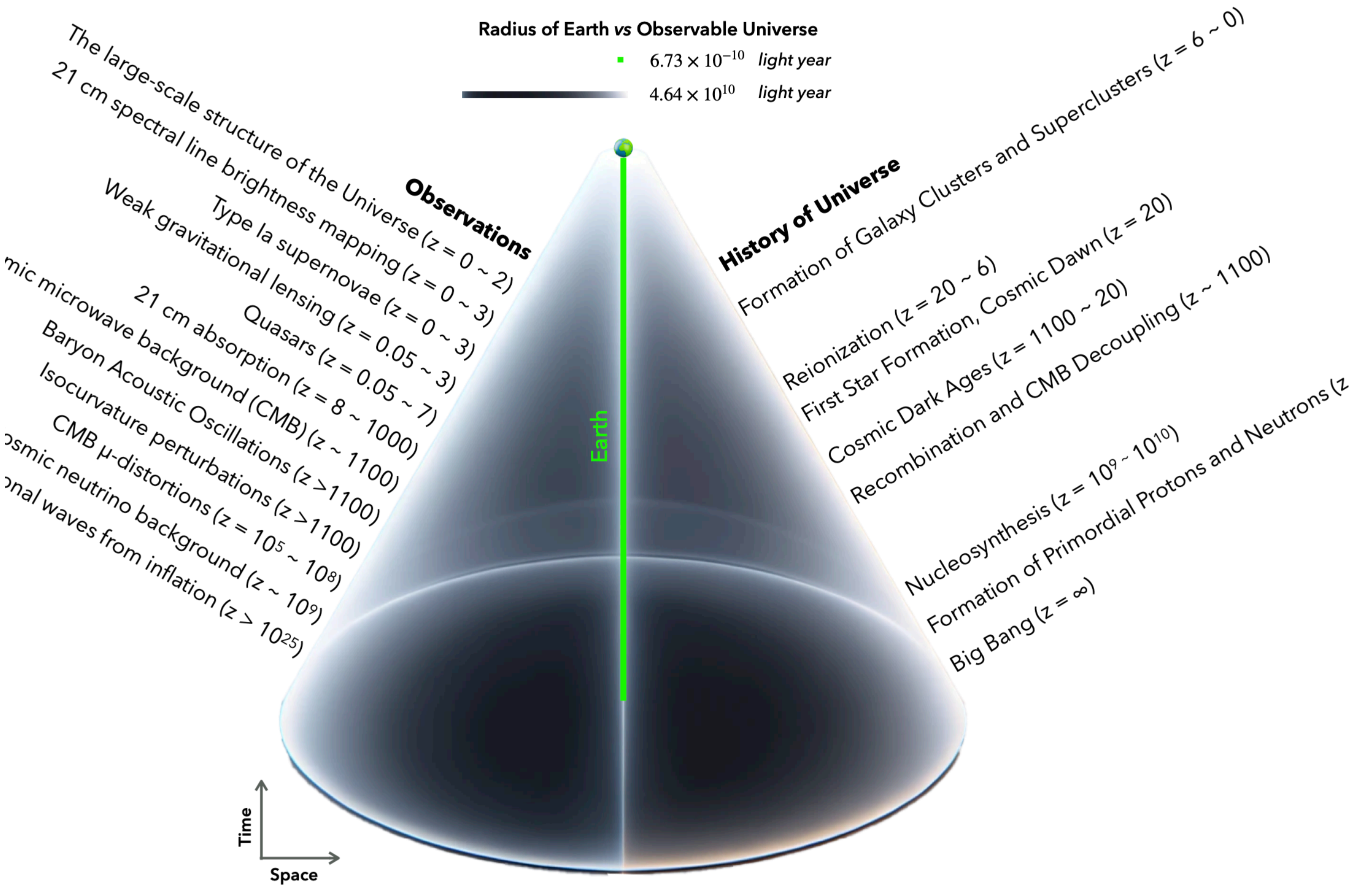
The year 2024 marked a pivotal moment in the advancement of artificial intelligence (AI) and machine learning (ML), particularly in their application to understanding the Universe. As we stand on the cusp of an era defined by the integration of AI into every facet of scientific inquiry, it is crucial to reflect on the profound implications and transformative potential of these technologies. The eight papers summarized in this collection were written in collaboration with ICRANet in 2024, especially the last two, illuminate a future where AI is not just a tool for data analysis but a partner in the quest to uncover the fundamental principles governing the cosmos.

A recurring theme in these works is the remarkable versatility of modern AI systems. Unlike traditional approaches that rely on specialized models tailored for specific tasks, recent advancements demonstrate the power of unifying diverse scientific inquiries under a single framework. The fine-tuning of large language models (LLMs) such as GPT exemplifies this trend. By training one model to perform tasks ranging from classifying celestial objects to inferring black hole parameters, researchers have shown that AI can generalize across domains, effectively serving as a universal instrument for astrophysical exploration. This ability is not just a technical milestone; it represents a philosophical shift in how we approach the integration of knowledge across disciplines.

In particular, the reflections on “Can AI Understand Our Universe?” delve deeply into the nature of understanding itself. Two concepts—intuition and causality—are identified as cornerstones of human cognition. The authors argue that AI, through technologies like Transformers, chain-of-thought reasoning, and multimodal processing, is beginning to exhibit elements of these capabilities. While AI does not possess human-like intuition rooted in sensory experience, it develops what can be described as “super-intuition.” By processing vast and complex datasets, AI can identify patterns and relationships that are invisible to the human mind. For example, AI's ability to “perceive” high-dimensional data and integrate information across spatial and temporal scales offers a new lens for understanding phenomena ranging from gamma-ray bursts to cosmic evolution.

Equally compelling is AI’s potential to uncover causality. The discussion highlights how AI employs statistical models and causal inference techniques to establish relationships that go beyond simple correlations. This capability is particularly significant in fields where human intuition often falls short, such as the analysis of complex systems or the exploration of phenomena at the edges of observational capabilities. AI’s ability to test counterfactuals and simulate alternative scenarios introduces a new dimension to scientific reasoning, enabling researchers to probe the mechanisms underlying observed events with unprecedented depth and rigor.

The implications of these developments extend far beyond astrophysics. The methodologies pioneered in these studies are adaptable to other domains, including climate science, medicine, and economics. For instance, AI’s capacity to integrate multimodal data—from images and text to time-series measurements—positions it as a powerful tool for tackling global challenges that require interdisciplinary approaches. Moreover, the reflections on AI’s role in understanding underline its potential to bridge the gap between data-driven insights and theoretical frameworks, fostering a more holistic view of complex systems.



Phenomena across different periods of the observable universe, with the expectation that AI will integrate various existing and future observations for a comprehensive understanding.

However, these advances also compel us to confront fundamental questions about the nature of understanding and the limits of machine cognition. While AI can process information at scales and speeds that far exceed human capabilities, it lacks the self-awareness and contextual grounding that define human understanding. This raises critical ethical and philosophical considerations. As we increasingly rely on AI to guide scientific discovery, we must remain vigilant about the boundaries of its applicability and the risks of over-reliance on its outputs.

In conclusion, the works presented here illustrate a trajectory of progress that is both inspiring and thought-provoking. They challenge us to rethink traditional approaches to scientific inquiry and embrace a future where AI and ML serve not merely as tools but as collaborators in the pursuit of knowledge. As we move forward, the integration of AI into scientific research will undoubtedly reshape our understanding of the Universe, revealing not only its hidden laws but also new questions that we have yet to imagine. The journey ahead is as much about exploring the potential of AI as it is about deepening our connection to the mysteries of existence.

PAPERS

1) FNet II: Spectral Classification of Quasars, Galaxies, Stars, and Broad Absorption Line (BAL) Quasars

Summary:

This study introduces FNet II, an enhanced version of a convolutional neural network (CNN) designed for spectral classification of astrophysical objects, including quasars, galaxies, stars, and BAL-quasars. The research focuses on data from the Sloan Digital Sky Survey (SDSS) Data Release 17 (DR17), a vast dataset comprising 1.4 million spectra.

FNet II uses a ResNet-based architecture, allowing it to learn both local and global spectral patterns autonomously. Unlike other methods that rely on identifying specific spectral lines, FNet II employs a self-learning approach, making it adaptable to diverse and ambiguous spectra. The network achieves exceptionally high classification completeness:

- **Galaxies:** 99.00% ± 0.20
- **Quasars:** 98.50% ± 0.30
- **BAL-quasars:** 99.00% ± 0.18
- **Stars:** 98.80% ± 0.20

FNet II's performance is comparable to or better than existing models such as QuasarNET, but with broader applicability due to its design. Its ability to classify spectra without predefining emission or absorption lines represents a significant innovation in automating large-scale astronomical surveys.

Significance:

This work addresses critical challenges in processing the massive datasets generated by modern astrophysical surveys. By automating spectral classification with high precision, FNet II reduces the reliance on manual inspection, enabling faster and more scalable analysis of cosmic objects. The method demonstrates potential for use in upcoming large-scale surveys, contributing to our understanding of the Universe's structure and evolution.

2) Advancing Gamma-Ray Burst Identification through Transfer Learning with Convolutional Neural Networks

Summary:
This study presents a novel approach to identifying Gamma-Ray Bursts (GRBs) using a combination of convolutional neural networks (CNNs), data augmentation, and transfer learning. GRBs are transient astronomical events that release immense energy in short durations, making their identification critical for multi-messenger and multi-band astrophysics. The study specifically uses data from the GECAM-B satellite and enhances identification by leveraging pre-trained models on a larger dataset from the Fermi/GBM satellite.

Key components of the approach include:

1. **A multi-scale feature cross fusion module (MSCFM)** integrated into a ResNet-based CNN to improve feature extraction.
2. **Transfer learning** to adapt models trained on Fermi/GBM data to the GECAM-B dataset, addressing challenges of limited labeled data.
3. **Data augmentation techniques** that increase the diversity of training samples by creating synthetic GRB light curves with varying signal-to-noise ratios (SNRs).

The model achieved:

- **Accuracy:** 96.41% on the GECAM-B dataset
- **Precision:** 91.07%
- **Recall:** 89.47%

Three previously unidentified GRBs were discovered in GECAM-B data, validated through manual analysis and comparison with the Fermi/GBM catalog.

Significance:

This research advances the automation of GRB identification, a critical need for modern astrophysical studies. The transfer learning and data augmentation methods significantly enhance the adaptability of machine learning models to limited datasets. These innovations have the potential to improve multi-satellite GRB detection, enabling more efficient follow-up observations of afterglows, host galaxies, and other counterparts. The proposed methodology offers scalability to future missions, enhancing the study of high-energy transients.

3) Application of Machine Learning to Background Rejection in Very-high-energy Gamma-Ray Observation

Summary:

This paper investigates the use of advanced machine learning (ML) techniques to distinguish between very-high-energy (VHE) gamma rays and the cosmic-ray hadronic background, a critical challenge in gamma-ray astronomy. Using simulated data from the LHAASO-KM2A experiment, the study evaluates various ML models for gamma/hadron classification across four energy ranges (10 to 10 eV). Eight high-level features were extracted from the simulation data, focusing on physical properties like the muonic and electromagnetic components of extensive air showers.

Key methods include logistic regression, support vector machines, decision trees, random forests, XGBoost, CatBoost, deep neural networks (DNNs), and a stacking ensemble algorithm that combines these models. The stacking ensemble demonstrated the best overall performance, achieving high accuracy and recall across energy bins.

Key Results:

- The stacking ensemble model outperformed traditional methods and standalone ML models, particularly in lower energy ranges (10-10 eV).
- Gamma/proton discrimination was quantified using metrics like accuracy (95.18% to 99.39%), F1 scores, and the Q-factor, with the stacking model achieving a Q-factor of 21.19 in the highest energy bin.

- Feature importance analysis identified parameters such as NfiltM (filtered muon count) and Base (ratio of muons to electrons) as critical for classification.

Significance:
This work demonstrates that ensemble ML techniques can significantly enhance gamma-ray background rejection. The results are crucial for future large-scale gamma-ray observatories, improving sensitivity and detection efficiency. Moreover, the interpretability of ML models provides insights into the underlying physical processes in cosmic-ray showers, contributing to multimessenger astrophysics.

4) Deep Learning for Identification and Characterization of Ca II Absorption Lines: A Multitask Convolutional Neural Network Approach

Summary:
This study introduces a multitask convolutional neural network (CNN) for automating the detection and characterization of Ca II absorption lines in quasar spectra. These lines are crucial for understanding the distribution of gas, dust, and metals in galaxy halos and disks, as well as their evolution over cosmic time. Traditionally, identifying these lines required manual Gaussian fitting, which is time-consuming and prone to errors.

The novel ResNet-CBAM model, which combines residual learning with an attention mechanism, demonstrated outstanding performance. The model achieved a classification accuracy of 99.77% and excelled in predicting equivalent width (EW) and full width at half maximum (FWHM), with average correlation coefficients of 0.98 and 0.85, respectively. By applying the model to Sloan Digital Sky Survey (SDSS) data (DR7 and DR12), it rediscovered 321 known Ca II absorbers and identified 381 new candidates, increasing the sample size significantly.

Significance:

- **Efficiency and Precision:** The model automates the detection process, significantly reducing manual inspection workloads while improving accuracy and reliability.

- **Scientific Impact:** The discovery of new Ca II absorbers enables more detailed studies of the interstellar medium, galaxy formation, and cosmic evolution.
- **Future Applications:** The methodology can be adapted for other spectral features and integrated into next-generation astronomical surveys.

This research showcases how deep learning can revolutionize the analysis of quasar absorption lines, paving the way for more efficient and scalable astrophysical investigations.

5) Modeling Blazar Broadband Emission with Convolutional Neural Networks: Synchrotron Self-Compton Model

Summary:
This work introduces a novel approach to modeling the spectral energy distributions (SEDs) of blazars, a subclass of active galactic nuclei known for their high luminosity and variability. By employing convolutional neural networks (CNNs), the study overcomes the computational challenges associated with traditional methods, which are often resource-intensive and time-consuming.

The research demonstrates the effectiveness of CNNs in reproducing blazar spectra generated by sophisticated numerical models. The trained CNN provides fast and accurate predictions of SEDs, significantly reducing computational time to milliseconds. This allows for real-time fitting of multiwavelength observational data. The authors showcase the framework's capability by modeling the SEDs of two blazars, Mrk 421 and 1ES 1959+650, and obtaining detailed posterior parameter distributions.

Significance:

1. **Efficient Computational Framework:** The CNN-based method addresses a longstanding challenge in astrophysics—balancing computational efficiency with the complexity required for accurate modeling. By drastically reducing the time needed for SED modeling, this approach opens new avenues for analyzing large datasets from blazar observations.
2. **Enhancing Multi-Messenger Astronomy:** The framework lays the groundwork for incorporating additional astrophysical processes, including hadronic interactions and external Compton models. This paves the way for a more comprehensive

understanding of blazars within the context of multi-messenger astronomy, combining electromagnetic and neutrino observations.

3. **Accessible and Scalable Tools:** The development of a publicly available interface, hosted by the Markarian Multiwavelength Data Center, will enable researchers to model SEDs interactively. This democratizes access to advanced modeling techniques, fostering broader collaboration and innovation in the field.

Broader Implications:

The work exemplifies how AI technologies can revolutionize the interpretation of complex astrophysical phenomena. By providing a scalable and accurate tool for modeling SEDs, this study enhances our ability to investigate the mechanisms powering some of the Universe's most energetic objects. The methodology is expected to have applications beyond blazars, including gamma-ray bursts and kilonovae, underscoring its versatility in multiwavelength and time-domain astronomy.

6) Modeling Blazar Broadband Emission with Convolutional Neural Networks: External Compton Model

Summary:

This paper builds upon prior work by extending a convolutional neural network (CNN)-based approach for modeling the spectral energy distributions (SEDs) of blazars to include the External Inverse Compton (EIC) mechanism. The EIC model accounts for interactions between relativistic electrons in the jet and external photon fields originating from the accretion disk, broad-line region, and dusty torus, explaining the high-energy emission in blazars. This is a significant step forward, as EIC modeling involves more parameters and greater spectral diversity compared to previous synchrotron self-Compton (SSC) models.

The CNN was trained on a large dataset of numerically simulated spectra and demonstrated high accuracy in reproducing the SEDs. It was applied to two well-known flat-spectrum radio quasars (FSRQs), 3C 454.3 and CTA 102, during their flaring states. This enabled detailed modeling of their broadband emissions and derivation of posterior parameter distributions.

Significance:

1. **Efficiency and Scalability:** The CNN dramatically reduces computation time for SED modeling, allowing complex radiative processes to be modeled in seconds instead of months. This enables real-time analysis of large datasets from modern multiwavelength astronomical surveys.
2. **Comprehensive Framework:** By incorporating external photon fields, the model captures a more complete picture of the radiative mechanisms in FSRQs. It enables comparisons between SSC and EIC scenarios, offering insights into the location and physical properties of the emission regions in blazar jets.
3. **Advancing Multi-Messenger Astronomy:** The methodology is adaptable for future studies of lepto-hadronic models, including interactions that produce high-energy neutrinos, bridging electromagnetic and neutrino observations in multi-messenger astrophysics.
4. **Public Access:** The model is made publicly available through the Markarian Multiwavelength Data Center (MMDC), promoting broader access to advanced tools for the scientific community.

Broader Impact:
This study highlights the growing role of AI in addressing computational bottlenecks in astrophysics. The integration of CNNs with complex numerical models provides a powerful framework for exploring the physics of relativistic jets and the environments around supermassive black holes. It exemplifies how AI can facilitate the interpretation of increasingly large and complex astronomical datasets, contributing to a deeper understanding of the Universe's most energetic phenomena.

7) Can AI Understand Our Universe? Fine-Tuning GPT with Astrophysical Data

Summary:
This paper explores the groundbreaking potential of artificial intelligence (AI) in unifying diverse astrophysical tasks under a single framework. By fine-tuning GPT, a large language model, with data from galaxies, quasars, gamma-ray bursts, and black holes,

the study demonstrates the model's ability to interpret, classify, and extract meaningful information from different types of astrophysical phenomena. Remarkably, one model proves capable of handling tasks as varied as object classification, redshift estimation, gamma-ray burst categorization, and black hole parameter inference.

Significance:

1. **Versatility in Understanding:** The ability of a single AI model to process and interpret multiple types of astronomical data underscores a significant leap in how science can approach complex datasets. This adaptability mirrors the way humans apply universal principles to diverse problems, hinting at AI's potential to synthesize insights across different domains.
2. **Discovery of Universal Laws:** By identifying patterns and relationships in vast datasets, the AI model offers a path to uncovering deeper, universal principles governing the cosmos. This work suggests that AI may one day assist in the discovery of laws that underpin seemingly disparate phenomena, paving the way for a more unified understanding of the Universe.
3. **Future of Scientific Exploration:** The study exemplifies how AI could become an indispensable tool for next-generation scientific facilities. With their ever-expanding data volumes, such tools will be essential for analyzing information at a scale and speed far beyond human capacity, unlocking discoveries that were previously inaccessible.

Broader Meaning:

This research represents a shift in how humanity approaches scientific inquiry. By demonstrating that one model can integrate and understand diverse astrophysical datasets, it signals the potential for AI to go beyond data analysis to uncovering the fundamental laws of nature. The study not only advances our ability to explore the Universe but also inspires a future where AI serves as a collaborator in revealing the interconnectedness of all cosmic phenomena.

8) Reflections on "Can AI Understand Our Universe?

Summary:
This article extends the discussion on whether artificial intelligence (AI), particularly large language models like GPT, can achieve a level of "understanding" that goes beyond data processing to deducing the fundamental principles of our Universe. It focuses on two key aspects of understanding—intuition and causality—and explores how current AI technologies, such as Transformers, chain-of-thought reasoning, and multimodal processing, contribute to this pursuit.

The author highlights how a single AI model can unify tasks across diverse astrophysical datasets, such as classifying celestial objects, estimating quasar redshifts, and inferring black hole parameters. This ability to generalize across domains signifies a departure from traditional approaches that rely on task-specific models.

Significance:

1. **AI as a Tool for Unified Understanding:** The potential of one model to analyze diverse data types and uncover relationships offers a path toward discovering universal principles that govern the cosmos. This integration points to a future where AI might bridge gaps between seemingly unrelated scientific phenomena.
2. **Expanding Human Perception:** AI technologies extend beyond human sensory and cognitive limitations. By processing high-dimensional data, perceiving across scales, and synthesizing multimodal information, AI can identify patterns and causal relationships inaccessible to humans. This capability enables new insights into complex systems, from molecular biology to cosmic evolution.
3. **Advancing Scientific Exploration:** The reflections underscore the transformative role AI will play in the era of large-scale, multidisciplinary data. As observational datasets grow exponentially, AI's ability to integrate and analyze this information efficiently will accelerate discovery and deepen our understanding of the Universe.

Broader Implications:
The author posits that AI's "understanding" is not merely about replicating human cognition but developing unique perspectives informed by its strengths—such as super-intuition and superior causal reasoning. These abilities could lead AI to propose novel hypotheses, reveal hidden laws of nature, and even reshape our scientific paradigms.

However, the piece also emphasizes the philosophical and technical challenges ahead. AI's current limitations, such as its dependence on extensive training data and lack of self-awareness, suggest that achieving human-level understanding—or surpassing it—remains a long-term goal.

REFERENCE

1) Moradi, R., F. Rastegarnia, Y. Wang, and M. T. Mirtorabi, Monthly Notices of the Royal Astronomical Society 533, no. 2 (2024): 1976-1985

2) Zhang Peng, Bing Li, Ren-zhou Gui, Shao-lin Xiong, Yu Wang, Yan-qiu Zhang, Chen-wei Wang, Jia-cong Liu, Wang-chen Xue, Chao Zheng, Zheng-hang Yu, Wen-long Zhang, arXiv preprint arXiv:2408.13598 (2024)

3) Jie Li, Hongkui Lv, Yang Liu, Jiajun Huang, Yu Wang, and Wenbin Lin, The Astrophysical Journal Supplement Series 276, no. 1 (2025): 24

4) Yang Liu, Jie Li, Linqing Gao, Haotong Zhang, Zhenghua Xu, Yu Wang, and Wenbin Lin, The Astrophysical Journal Supplement Series 276, no. 2 (2025): 37

5) Damien Bégué, Narek Sahakyan, Hüsne Dereli-Bégué, Paolo Giommi, Sargis Gasparyan, Mher Khachatryan, Andrea Casotto, and Asaf Pe'er, The Astrophysical Journal 963, no. 1 (2024): 71

6) N. Sahakyan, D. Bégué, A. Casotto, H. Dereli-Bégué, P. Giommi, S. Gasparyan, V. Vardanyan, M. Khachatryan, and A. Pe'er, The Astrophysical Journal 971, no. 1 (2024): 70

7) Yu Wang, Shu-Rui Zhang, Aidin Momtaz, Rahim Moradi, Fatemeh Rastegarnia, Narek Sahakyan, Soroush Shakeri, and Liang Li, arXiv preprint arXiv:2404.10019 (2024)

8) Yu Wang, International Journal of Modern Physics D, 2024