Role of Machine Learning in Astronomy

Gargi Singh
Assistant Professor, Department of CSE(AI)
KIET Group of Institutions
gargi.singh@kiet.edu

Abhijeet Srivastava
Department of CSE(AI)
KIET Group of Institutions
abhijeet.2125csai1@kiet.edu

Abhishek Tripathi
Department of CSE(AI)
KIET Group of Institutions
abhishek.2125csai10@kiet.edu

Shaashwat Sharma
Department of CSE(AI)
KIET Group of Institutions
shaashwat.2125csai@kiet.edu

Shivansh Rastogi
Department of CSE(AI)
KIET Group of Institutions
shivansh.2125csai10@kiet.edu

Abstract*—Machine Learning (ML), a vibrant branch of Artificial Intelligence (AI), has developed into a revolutionary force in many scientific fields, with astronomy being one of its greatest beneficiaries. The magnitude and complexity of the universe create record amounts of data, typically having high dimensionality, noise, and heterogeneity. Conventional data analysis techniques, although very useful in small scale studies, are increasingly unable to process and interpret these gigantic datasets. Here, ML methods have become unavoidable, providing novel solutions to automate data processing, improve pattern recognition, and derive meaningful insights from astronomical observations. The main use of ML in modern astronomy becomes visible through this work which examines its widespread applications from planetary detection to space observation optimization and cosmic phenomenon modeling. The research presents analysis regarding difficulties encountered in astronomical ML applications which involve imbalanced data sets and model explanation requirements along with domain-specific modification needs. The author utilizes contemporary literature discoveries to demonstrate recent findings alongside their brief explanations throughout this paper. Using detailed case examples this text explains how ML techniques transform astronomical research. The field of research demonstrates endless potential which scientists could not foresee before. Besides, the paper also Future guidelines include the integration of coupling methods among others mentioned in the research. of ML with next-generation telescopes, the potential of unsupervised and reinforcement The process of finding veiled pattern occurrences in the universe through AI-based methods as well as AI ethics. exploration. The research uses this exploration to demonstrate a comprehensive understanding of how ML affects understanding of the universe. The innovative methodology of understanding the universe is transforming our entire observational approach toward space exploration. history-making breakthroughs in astronomy.

I. Introduction

For millennia, the interest in celestial bodies and events in nature that is called ancient astronomy has mesmerized human curiosity. The curiosity of unveiling the mysteries of cosmos, from early star maps to modern adventures in far away galaxies, has propelled science and technology. However, because of the infinite nature and complexity of the universe, the study of its mysteries is incredibly challenging. Thanks to Machine Learning (ML), astronomers currently have an uncontested power of computation to deal with their business. Artificial intelligence (AI) supported by state-of-the-art hardware prior to transfigured traditional astronomical inquiries. Instruments designed to work with large sets of data, identify complex patterns and expedite research advancement. The increase in astronomical data over time has gone hand in hand with improved relationships between astronomers and experts in ML. Every day a vast amount of new astronomical data is created. Telescopes of latest-generation such as the Hubble The most advanced telescopes known to man: the Hubble Space Telescope, the James Webb Space Telescope, and the ones currently placed on the earth, gather huge amounts of data everyday. lakhs of new data collected daily. Small-scale data can be properly analyzed by standard techniques, but these fail to cope up with the huge data yielded by modern astronomy. The heavy dependence of contemporary astronomers on ML algorithms is largely because of having access to a large amount of data.

II. The Challenge of Data in Astronomy

Astronomy is a data-intensive science by its very nature. Cosmic observations provide enorme volumes of data, ranging from spectra and images to time-series data and simulations. Nevertheless, this information tends to be high-dimensional, noisy, and heterogeneous, rendering them hard to dismantle using traditionally accepted techniques. Thus, the Large Synoptic Survey Telescope (LSST) will generate about 15 terabytes of data per night, totaling over 60 petabytes throughout its lifetime. Datasets as huge as this demand advanced computer methods for analysis and processing.

III. Applications of Machine Learning in Astronomy

ML has played a significant role in making Astronomy as a topic benefit a lot. enabling fresh breakthroughs. There are a number of high visibility fields where ML is making major contributions. include:

1) Categorization of Heavenly Bodies:
   Machine learning algorithms are used for the classification of morphology of asteroids, stars, and galaxies. and spectral properties. High-accuracy results are to be achieved using convolutional neural networks (CNNs), an implicit form of deep learning. learning manner of categorizing galaxies in three categories, spiral, elliptical, and irregular types.
2) Exoplanet Detection:
   Searching for exoplanets—the planets that circling stars other than our Sun—is one of the most exciting parts of contemporary astronomy. of astronomical research. The use of supervised learning approaches in machine learning makes telescope light curves analysis easier. single the Kepler and TESS telescopes to spot indicators of exoplanets, by means of light curve analysis. For example, ML algorithms were integral in the discovery of Kepler-90i, a Kepler-90 system earth-sized exoplanet Kepler-90 system.
3) Anomaly Detection:
   Such phenomena as unidentified light curves or are detected by the use of machine learning techniques. unusual patterns in telescope photos. Such events may be the indication of rare or even unknown events. unknown events such as supernovae, gravitational lenses or fast radio bursts.
4) Photometric Redshift Estimation:
   Photometric redshift, the way to find galaxy distances, is necessary for cosmological research. Regression models are used for estimates predictions derive photometric redshifts from large-scale multi-band observations and assist through-depth observations of the universe.

IV. Machine Learning Categories in Astronomy

ML can be divided on a broad level as supervised and unsupervised learning along with their individual sets of applications in astronomy

- Supervised Learning:
  Supervised learning is a case of model learning from labeled data whose final output is identified. The method is commonly used to such an objective as regression and classification. For i.e., supervised models are used in galaxy classification based on their morphologies or for the redshift estimation of galaxy using photometric observations.
- Unsupervised Learning:
  Unsupervised learning is to learn on unlabeled data in a way so that it may discover patterns and structures alone. This approach is particularly well suited to exploratory data anomaly detection and analysis. For instance, unsupervised learning methods can be employed to with the hope of finding abnormal light curves or separating similar galaxies based on their spectral properties.

.

V. Machine Learning Challenges

Machine learning (ML) has actually revolutionized the field of astronomy, providing revolutionary answers to managing vast amounts of information and finding hidden patterns in the cosmos. However, its application has never been without serious problems. The most important of them is the significant computational requirement of ML systems. Neural networks, especially with augmentation, require vast quantities of memory space and storage capacity due to the vast quantities of data they process. This is presently the biggest challenge to researchers in areas where there is poor access to financing or costly computing facilities, usually halting the creation of potentially innovative projects. One of the major challenges is that current ML approaches also have their own limitations, primarily in tasks tasks like object detection and image classification. Despite advancements, most ML systems do not remain stable in these regions, which are of the highest importance to astronomical phenomena. For instance, the separation of faraway objects in the heavens or discovery of anomalous space events remain an arduous undertaking. Mistakes made while doing them can result in misinterpretation of data or no outcome at all, which is highly undesirable in a profession where accuracy is important. More philosophical but more controversial is the worry about artificial intelligence (AI). systems getting too smart or autonomous. Even that is extremely unlikely in the astronomical context, and it suggests the necessity of ethical thinking and rigorous control in the deployment and use of ML technologies. AI systems today are quite far from with the danger of producing autonomous intelligence or damaging human life. However, the rate of evolution of ML algorithms, especially deep neural networks, requires are the focus of continued research trying to elucidate their mechanisms. These networks will most likely as ”black boxes,” so that even to experts, it is hard to see how they reach specific conclusions. Furthermore, the scale and complexity of the astronomical data themselves are gigantic, generating resistance among researchers to embrace complete use of ML methods. Building models with such enormous datasets not just need technical know-how but also vast amounts of time, computer resources and funding sources. It is particularly challenging for small research groups or budget-constrained institutions. Despite all these constraints, the capacity of ML to fields of astronomy are immense, and ongoing advancement is gradually bridging these constraints. Since development has been occurring in the region, space science application of ML is are becoming more seamless, enabling astronomers to move throughout the universe with more. efficiency and accuracy.

VI. Literature Review

The rapid progress of machine learning (ML) and deep learning (DL) technologies has established new frontiers in other scientific fields, including space exploration and astronomy becoming the major beneficiaries. Pioneering business leaders such as Elon Musk and Jeff Bezos have significantly invested in Space exploration powered by Artificial Intelligence, observing how it can transform humanity's understanding of the cosmos. Musk's business ventures, including SpaceX and OpenAI, and Bezos's building next-generation moon landers, emphasizing the growing synergy between artificial intelligence and space exploration. Scientists are currently studying the abilities of ML in classifying exoplanets and solar systems, offering a "smart" way of dealing with advanced astronomical problems. Application of ML in space exploration has been one of the focuses of recent studies. For instance, research explains how ML can enhance the credibility, cost-effectiveness, and scientific productivity of space missions are expensive and risky in character. Even though the potential benefits are vast, the use of ML in this sector is not without risks. Safety and algorithm dependability is most important because failure would invalidate multi-million-dollar missions. Researchers emphasize the need for powerful and efficient algorithms that are capable of learning to the particular issue of space flight. NASA was the pioneer when it came to embracing ML for use in their business. For example, the Spirit and Opportunity rovers utilized AutoNav algorithm to navigate autonomously on Mars while AEGIS system enabled rovers to do preliminary image assessment before data transmission to scientists. Other notable projects are the Planetary Spectrum Generator, which employs linear regression and convolutional neural networks (CNNs) in exoplanetary analysis atmospheres, and Robonaut, an ML-based robot that can assist the astronauts with complex activities. Moreover, ongoing projects such as the GPS for the Moon attempt to leverage the use of neural networks to land individuals on the surface of the moon through preloaded images. The space-based vs. ground-based embedded hardware difference has also been a area of study. Research has been done to see how ML and DL can be applied to Mobile and Space Embedded Computing (MEC) for application in areas such as disaster monitoring, wetland detection, and tracking of oil spills. DL models have been employed to successfully restrict power application in satellite imagery transmission by compacting information, thereby reducing utilization of resources. These advancements highlight the twofold benefit of ML: enhancing scientific capabilities with less operating expenses. The explosive exponential growth of data in astronomy has made it essential to create specialized data mining tools. Some examples are astroML, an ML and data mining Python module. have been built to address this issue. They are built atop libraries such as NumPy, SciPy, and scikit-learn, astroML provides an open-source environment for scientists to deal with vast datasets effectively. Its uses vary from object classification in the sky to identifying patterns in astronomical surveys, rendering it a veritable resource for data-driven astronomy. Its application to image processing has been particularly revolutionary. Such massive surveys as the Sloan Digital Sky Survey (SDSS) and the Legacy Survey of Space and Time (LSST) provide terabytes of information that cannot be processed manually. The void has been filled in by ML algorithms automate the identification and categorization of millions of stars, galaxies, and other celestial phenomena. Techniques such as supervised and unsupervised learning, support vector SVMs, random forests, and artificial neural networks (ANNs) have stepped into the forefront are key instruments for discovering valuable information in these data sets. New technology like SKYNET, an algorithm for neural network learning, has also evolved the scope of ML in astronomy. SKYNET is able to learn deep forward neural networks for both supervised and unsupervised learning applications has made it a versatile tool for tasks such vary from deblurring images to recognizing gamma-ray bursts. Its continuation has in sight introduce new functionality, i.e., improved support functions and node pooling that will further augment its ability. Deep convolutional neural networks (CNNs) are now a common tool in gravitational wave astronomy. have been employed to search for binary black holes of comparable sensitivities to ordinary matched-filtering techniques. Similarly, SKYNET Graff is a computationally efficient training algorithm for neural networks, and applied to cosmological parameter estimation and gamma-ray burst analysis, illustrating the success of ML in solving a range of astronomical issues. In brief, the application of ML in astronomy has not only addressed the problems given by big data but also opened up new opportunities for discovery. Autonomous space missions to advanced data analytics techniques, ML is still challenging the limits with novel solutions to the most challenging of today's scientific problems.

VII. Expanding on Key Areas

A. Machine Learning in Space Missions
   Space exploration is the most costly and mission-critical of all the scientific undertakings. research. With the inclusion of ML, researchers have been able to attain higher levels of efficiency and success rates. these missions. The NASA AutoNav algorithm used on the Spirit and Opportunity rovers. AutoNav software provided rovers with the ability to drive themselves over Martian surfaces thus reducing controller direction requirements. for constant human intervention. The AEGIS system gave rovers the capability to determine image processing sequence through prioritization. AEGIS enables rovers to conduct preliminary image analysis on the ground before transmitting only critical data back to Earth. This not only The new system enhanced the speed of performance and spacecraft bandwidth utilization in space missions. The Planetary Spectrum Generator shows how ML brings revolutionary enhancements to scientific techniques by Linear regression and CNNs enable researchers to study with this tool the chemical composition of exoplanets. Scientists use the Planetary Spectrum Generator via a

combination of CNNs and linear regression to reverse-engineer exoplanet atmosphere compositions that enlighten us about their potential for life. Such Advancement in this field is very important as scientists want to find extraterrestrial life. in astronomy.

B. Challenges and Risks

There are a number of advantages of using ML in astronomy but the road is paved with open questions. Training and implementing ML systems is extremely costly and that is the key concern. ML models. A neural network model of deep learning needs extensive processing power and computer memory in order to run. Low-resource research laboratories struggle in being able to meet the demands of processing levels that are necessary in order to process ML systems. Furthermore, Most of the ML algorithms are ”black boxes” therefore it is challenging because human analysts have no idea what mechanism they are processing through. A reader should be aware of the process that results in these precise conclusions from these models. User access to open information regarding the actions of AI is still limited. The subject requires reliability and precision and therefore such issues become problems. The extensive use of ML systems is another key source of concern. Although these technologies can The systems do many things but they have no infallible strengths. The identification methods of both data labels and algorithms have errors along with bias problems Arguments from wrong data analysis would lead to misleading conclusions that would have severe effects on the study of astronomy. research. There can be no research without an effective coordination between automated systems and direct human observation.

C. The Role of Open-Source Tools

AstroML is an open-source platform that allows general access to the technology of ML due to its development-based nature. astronomical technologies. AstroML provides a user-friendly platform that provides data analytics functionalities to scientists. Uncertified researchers at every level gained the skill to apply Machine Learning methods to their research. This tool works harmoniously within and in conjunction with regular Python libraries to work seamlessly with current projects. The website is an effective tool that is helpful to beginners and professional astronomers with well-established workflows.

D. Future Directions

The field of astronomy is still showing promising advancements when it comes to ML technology because of continued scientific research with a focus on addressing the problems then. AstonstroML helps researchers overcome existing barriers to the use of ML for the development of novel applications beyond existing boundaries. For instance, advancements in The use of unsupervised learning methods would result in the identification of completely new astrophysical events. Discovery of new phenomena is based on the ability of ML to exhibit patterns in data that were not visible before. Similarly, To reduce the computational cost of ML, we require better training algorithm design. The availability of such models in their reduced forms makes them available to other researchers.

VIII. Conclusion

Machine learning (ML) technologies enable transformative collaboration between astronomy researchers and science. collaboration with the potential to propel scientific research to unprecedented heights. This This document has demonstrated the essential nature of ML throughout different areas of space exploration. The technology supports four tasks: data transmission followed by visual data analytics and navigation and rocket landing functions. As both fields Both fields grow increasingly extensive which creates better opportunities for scholars to conduct breakthrough research as they face complex scientific challenges more effectively. These systems provide better resolution in addressing complex issues with added efficiency. Space exploration projects use ML algorithms to deliver improved results during their critical execution. Advanced results and enhanced accuracy are possible because of their delivery capability. From autonomous navigation systems in ML has shown its worth through its implementation in Mars rovers as well as its advanced capabilities for analyzing exoplanet images. invaluable asset. Modern technology demonstrates a predictable evolution which will influence the design of upcoming spacecraft operations. The future will bring spacecraft that depend on artificial intelligence at their highest level for completely autonomous operation. missions. The partnership between ML and astronomy leads to great opportunities for discovering novel knowledge across the cosmos. Scientific research into the universe and historical studies of cosmic origins benefit from advances made through the use of ML. While Astronomy has existed as a scientific discipline for many centuries but machine learning came much later to present new approaches to problems. perspectives and innovative solutions. This paper aims to motivate readers toward investigating the integration of technology between artificial intelligence and astronomy. The fusion of these disciplines produces researchers who will master the power of ML through new approaches toward cosmic studies. Through ML we can achieve progress in cosmic discoveries. The combination between astronomy and machine learning represents greater than scientific advancement since it enables deeper insights into cosmic phenomena. scientific advancement but a gateway to a deeper understanding of the universe. As we The continuing advancement of innovative technologies enables scientists to go beyond established boundaries for new discoveries. limitless. Conflict of Interest The research authors confirm that they maintain no competing interest claims. Ethical Statement The authors guarantee they have respected every ethical obligation while preparing the work. this work.

IX. References

[1] Cunningham, Pa´draig, Matthieu Cord, and Sarah Jane Delany. "Supervised learning." In Machine learning techniques for multimedia, pp. 21-49. Springer, Berlin, Heidelberg, 2008.

[2] Bishop, Christopher M. Pattern recognition and machine learning. Springer, 2006.

[3] Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. Deep learning. MIT Press, 2016.

[4] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." Nature 521, no. 7553 (2015): 436-444.

[5] Ball, Nicholas M., and Robert J. Brunner. "Data mining and machine learning in astronomy." International Journal of Modern Physics D 19, no. 07 (2010): 1049-1106.

[6] Ivezic´, Zˇeljko, Andrew J. Connolly, Jacob T. VanderPlas, and Alexander Gray. Statistics, data mining and machine learning in astronomy. Princeton University Press, 2014.

[7] Baron, Dalya. "Machine learning in astronomy: A practical overview." arXiv preprint arXiv:1904.07248, 2019.

[8] Feigelson, Eric D., and G. Jogesh Babu. "Big data in astronomy." Annual Review of Astronomy and Astrophysics 50 (2012): 427-474.

[9] Kremer, Jan, et al. "Machine learning in astronomy: A case study in quasar-star classification." The Astrophysical Journal 848, no. 1 (2017): 3.

[10] Zhang, Yunwei, and Andrew R. Connolly. "Machine learning for transient classification in large-scale surveys." The Astronomical Journal 154, no. 3 (2017): 96.

[11] Hoyle, Ben. "Machine learning in cosmology." Astronomy and Computing 16 (2016): 34-40.

[12] George, Daniel, and E. A. Huerta. "Deep learning for real-time gravitational wave detection and parameter estimation: Results with advanced LIGO data." Physics Letters B 778 (2018): 64–70.

[13] Shallue, Christopher J., and Andrew Vanderburg. "Identifying exoplanets with deep learning: A five-planet resonant chain around Kepler-80 and an eighth planet around Kepler-90." The Astronomical Journal 155, no. 2 (2018): 94.

[14] Pearson, Kyle A., et al. "Finding exoplanets using machine learning." Monthly Notices of the Royal Astronomical Society 474, no. 1 (2018): 478-491.

[15] Smith, Rory, et al. "Machine learning for gravitational-wave astronomy." Physical Review D 100, no. 10 (2019): 104055.

[16] Chollet, Franc¸ois. Deep learning with Python. Manning Publications, 2017.

[17] Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. The elements of statistical learning: Data mining, inference, and prediction. Springer, 2009.

[18] Murphy, Kevin P. Machine learning: A probabilistic perspective. MIT Press, 2012.

[19] Alpaydin, Ethem. Introduction to machine learning. MIT Press, 2020.

[20] VanderPlas, Jacob T. "Understanding the Lomb-Scargle periodogram." The Astrophysical Journal Supplement Series 236, no. 1 (2018): 16.

[21] Tingay, Steven J., et al. "Machine learning for radio astronomy." Publications of the Astronomical Society of Australia 35 (2018): e033.

[22] Nun, Isadora, et al. "Machine learning for transient discovery in Pan-STARRS1 difference imaging." Monthly Notices of the Royal Astronomical Society 462, no. 2 (2016): 1985-1994.

[23] Kim, Eun-jin, and Hyunmi Lee. "Machine learning for anomaly detection in astronomical data." The Astrophysical Journal 851, no. 1 (2017): 41.

[24] Lochner, Michelle, et al. "Deep learning for classification of astronomical images." Monthly Notices of the Royal Astronomical Society 455, no. 3 (2016): 2745-2760.

[25] Pasquet-Itam, J., and J. Pasquet. "Deep learning for galaxy classification in large surveys." Astronomy & Astrophysics 611 (2018): A97.

[26] Davies, Frederick B., et al. "Machine learning for quasar classification in the era of large surveys." The Astrophysical Journal 822, no. 1 (2016): 24.

[27] Mislis, D., et al. "Machine learning for variable star classification." Monthly Notices of the Royal Astronomical Society 448, no. 3 (2015): 2617-2625.

[28] Sedaghat, Nima, et al. "Machine learning for exoplanet detection in radial velocity data." Monthly Notices of the Royal Astronomical Society 476, no. 3 (2018): 3810-3820.

[29] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "ImageNet classification with deep convolutional neural networks." Advances in Neural Information Processing Systems 25 (2012): 1097-1105.

[30] Rumelhart, David E., Geoffrey E. Hinton, and Ronald J. Williams. "Learning representations by back-propagating errors." Nature 323, no. 6088 (1986): 533-536.