

JGR Space Physics

COMMENTARY

10.1029/2019JA027502

Key Point:

- Discussion around a new community focused on machine learning in heliophysics

Correspondence to:

E. Camporeale,
enrico.camporeale@noaa.gov

Citation:

Camporeale, E., & The Scientific Organizing Committee of ML-Helio (2020). ML-Helio: An emerging community at the intersection between heliophysics and machine learning. *Journal of Geophysical Research: Space Physics*, 125, e2019JA027502. <https://doi.org/10.1029/2019JA027502>

Received 6 OCT 2019

Accepted 18 DEC 2019

Accepted article online 29 JAN 2020

ML-Helio: An Emerging Community at the Intersection Between Heliophysics and Machine Learning

Enrico Camporeale¹  and The Scientific Organizing Committee of ML-Helio

¹CIRES, University of Colorado, Boulder, CO, USA

Abstract The advancements and breakthroughs achieved in the last 5–10 years in Artificial Intelligence and machine learning (ML) have not gone unnoticed in the scientific community. The body of literature that borrows techniques from ML has steadily grown in all fields of physics. Space physics is particularly well posed to exploit ML due to the large amount of (often under scrutinized) data accumulated over the last few decades. Indeed, ML techniques can offer insights into the data that might enhance our understanding of physical mechanisms. Many of the pioneering studies on the use of ML in Space Physics have been led by several individuals who have independently taken the burden of moving out of their comfort zone to climb the steep slope of learning new jargon, new methodologies, and new coding skills. Such early adopters have recently convened in Amsterdam for the first conference on machine learning in heliophysics. The conference has laid the foundation for a new emerging community, and this commentary summarizes the discussions and steps taken to make such community flourish.

The first conference on machine learning in heliophysics (ML-Helio 2019) took place in Amsterdam on 16–20 September 2019. About 170 researchers from 28 countries met a few steps away from Amsterdam's Seventeenth-Century Canal Ring, a UNESCO World Heritage site, to present and discuss their research, network and above all to establish a new, vibrant interdisciplinary community focused on leveraging the huge possibilities offered by approaching heliophysics with the most modern ML techniques.

The audience was very heterogeneous, with about a quarter of the attendees being early-career researchers and skill levels in machine learning varying widely from beginners to everyday practitioners. The program was packed with 43 oral presentations, 100 posters, and 2 tutorials; however, it also allowed plenty of time for discussions and networking. In particular, on the last day, an open, lively discussion was held, whose main points are summarized here.

The conference spanned all aspects of heliophysics, including studies in solar physics, interplanetary medium, planetary science, the magnetosphere, ionosphere, and space weather. From the perspective of machine learning, the presentations focused on forecasting, feature extraction and engineering, data set development, and anomaly detection. The majority of studies fell into the category of supervised learning, but a few used less traditional techniques, such as information theory, unsupervised clustering methods, and physics-informed machine learning. The data set most utilized was the large amount of high-quality solar images provided by the Solar Dynamic Observatory. Images from solar observatories (including STEREO and SOHO) were incorporated in several projects as well. Other papers focused on in situ solar wind data (from, e.g., ACE, DSCOVR, and OMNI), ground-based magnetometers, magnetospheric missions (Van Allen Probes and MMS), and geomagnetic indices.

The need to create curated data sets and standardized metrics for comparing different models was one of the major issues discussed. Learning from the wider ML community, it has become clear that establishing a set of well-defined benchmarks and accessible data set helps enormously in evaluating and comparing different approaches and, ultimately, stimulates research on given problems. Recently, the teams working at the NASA Frontier Development Lab have used Google Cloud as a data sharing and compute platform (FDL is a public-private AI incubator for space science. See <https://frontierdevelopmentlab.org/>). This is not, however, seen by the community as a long-term, affordable solution. In this sense the ML-Helio community intends to bring to the attention of funding agencies the urgent need for sufficient computational resources, such as large data storage and high throughput GPU-based infrastructure, to enable these methods to be adopted by the wider community, thereby increasing their scientific return.

With more and more works combining expertise in space physics and machine learning being published in the scientific literature, the problem of reproducibility of results and availability of source codes was also discussed. The ML-Helio community supports and encourages open-source software and aims to establish a solid community practice in making published results transparent and reproducible and in giving proper credit to software developers. This translates into the publication of papers devoted to an in-depth discussion of the techniques and motivation behind the models/analyses as is done when new missions are launched.

Although machine learning is currently enjoying considerable coverage in the media, the space physics community remains relatively conservative, and thus, ML-based studies continue to receive a higher level of scrutiny and, often, skepticism. In this regard, a crucial aspect of new publications will be to include a well-thought justification of the proposed machine learning approach, by comparing new results against simpler statistical methods. This point cannot be overstated. Some of the attendees of the conference have noted that space physics is not new to the waves of artificial intelligence and some of the pioneering studies in the field were discussed. Indeed, more than 25 years have passed since the International Workshop on Artificial Intelligence Applications in Solar-Terrestrial Physics took place in Lund, Sweden, at a time when the term machine learning was not even invented! (Joselyn et al., 1993) Interestingly, the majority of those studies were not well received by the larger community, at least at the time. In spite of this, it appears that the space physics community may now be ready to embrace and enjoy the fruits borne by machine learning approaches.

With many attendees being early ML practitioners, the problem of how to find the best educational resources and how to navigate across the multitude of offerings available online and in the press was also particularly stressed. Indeed, there are only a few resources that are tailored to the space community: An online book that contains a collection of interactive Jupyter notebook (Bobra & Mason, 2019), a monograph on machine learning techniques for Space Weather (Camporeale et al., 2018), and, in the field of Space Weather, a recent review paper (Camporeale, 2019).

In order to help all of these ideas to become reality and to offer the members of the ML-Helio community a central hub where to share resources, information, advertise activities, and link to various projects, a github repository and wiki page is now live (<https://github.com/ml-helio>). Everybody is welcome to join and contribute.

Also, the presentations (oral and posters) discussed during the conference are available on the website (<https://ml-helio.github.io/>).

Finally, the presentations at the conference will be collected in a special issue of *Frontiers in Astronomy and Space Sciences* (<https://www.frontiersin.org/research-topics/10384/machine-learning-in-heliophysics>). The call for papers is open to all contributors and not limited to conference participants.

In conclusion, machine learning will increasingly and systematically be used in more and more scientific disciplines, and space physics is no exception. As for any interdisciplinary approach, there is an obvious barrier that the community faces in entering the field. The first important steps to overcome such barriers have been taken at the ML-Helio 2019 conference where active participants of this emerging community have met and identified important practices that will ensure that the field will be supported by an enthusiastic, open-minded, welcoming community yet following the most rigorous and well-established scientific procedures.

The ML-Helio conference will be held every 2 years, rotating venues between continents.

Appendix A: ML-Helio Author Names and Affiliations

Hazel Bain, CIRES, University of Colorado, Boulder, CO, USA

Monica Bobra, Stanford University, CA, USA

Jacob Bortnik, University of California, Los Angeles, CA, USA

Enrico Camporeale, CIRES, University of Colorado, Boulder, CO, USA

Mark Cheung, Lockheed Martin Solar and Astrophysics Laboratory, USA

Veronique Delouille, Royal Observatory of Belgium

Farzad Kamalabadi, University of Illinois
 Michael Kirk, NASA Goddard, USA
 Giovanni Lapenta, Katholieke Universiteit Leuven, Belgium
 Stefan Lotz, South African National Space Agency, South Africa
 Sophie Murray, Trinity College Dublin, Ireland
 Bala Poduval, University of New Hampshire, USA
 Pete Riley, Predictive Science Inc., USA
 Simon Wing, APL, Johns Hopkins, USA

Acknowledgments

We kindly acknowledge the financial support provided by the following institutes and agencies: NASA, INRIA (Institut National de Recherche en Informatique et en Automatique), KNMI (Royal Netherlands Meteorological Institute), ESA, Origins Center, NDNS+, Frontiers In Astronomy and Space Science, and Springer Nature Applied Science. No data were used.

References

- Bobra, M., & Mason, J. (2019). Machine learning, statistics, and data mining for heliophysics. <https://doi.org/10.5281/zenodo.1412824>
- Camporeale, E. (2019). The challenge of machine learning in space weather: Nowcasting and forecasting. *Space Weather*, 17(8), 1166–1207. <https://doi.org/10.1029/2018SW002061>
- Camporeale, E., Wing, S., & Johnson, J. (Eds.) (2018). *Machine learning techniques for space weather*. Amsterdam, Netherlands: Elsevier.
- Joselyn, J., Lundstedt, H., & Trollinger, J. (Eds.) (1993). Proceedings of the international workshop on artificial intelligence applications in solar-terrestrial physics, Lund, Sweden, 22-24 September 1993.