4th URSI AT-RASC, Gran Canaria, 19-24 May 2024

Beyond the Crystal Ball: Machine Learning and the Future of Space Weather Prediction

Dario Del Moro(1)
(1) Physics Department, Rome "Tor Vergata"University,

Abstract

In this paper, I will share my personal experience on how (Machine Learning) ML can contribute to the forecast of space weather, presenting my point of view about what ML can still offer us and why it is probably here to stay. Considering how large the ML field is, I can only focus on some aspects, trying to report about recent approaches that show how ML methods can go further than the common "black-box" approach and can also give us a physical insight.

1. Introduction

The capability of humankind to somehow forecast the future is probably the main reason for its success in the survival of the fittest game. We have honed this ability first in the eons of biological evolution, then encoding it in culture, finally by thoroughly applying the scientific method to our everyday life problems. But it is just recently that we have developed the conceptual tools to understand why some events are more difficult to forecast than others.

The development in the last decades in the sciences of complex systems and statistics have allowed us to better define the limits of predictability and -often- to push forward those limits. The rapid progress of computing techniques and capabilities has even further bolstered those two branches of Science to the point that it is now feasible both to compute the trajectory of plasma and magnetic field structures under the Magneto-Hydro-Dynamic (MHD) equations in domains as large as the Heliosphere [1,2], and to try to predict -just by analyzing solar full disk images or the magnetograms- whether a given solar Active Region will release part of its stored energy as high energy photons and particles, or shoot out a coronal mass ejection.

Is it then all solved? Not at all. In some cases,, we can make almost exact predictions; in others, we can have robust predictions with an estimation of the associated errors, and sometimes, we can only speak about probabilities. We are prevented either by intrinsic limitations or by the lack of knowledge about the present state. Unfortunately, space weather predictions usually suffer from both of these effects. We quite often have complex systems to model and only a fragmentary perception of what was actually happening on the Sun and its surroundings. To counter these limitations, we have come up with clever and clever numerical techniques to solve the differential equations that typically describe our problems, and ensemble methods [3,4] to cope with measure errors and unknown variables. We strive to extract from remote measures all the relevant information [5], by applying our understanding of the physics of the problem, to feed this information into our forecasting algorithms. And we are getting better and better at this too.

Nevertheless, the robust forecast of flare eruption still escapes us. And also the apparently simpler problem of the propagation of a coronal mass ejection in the interplanetary medium has not been solved to the limit that we would like, while we fight with the uncertainties associated with the boundary conditions. As a consequence, some of us [e.g., 6,7] turned to the dark side and applied the methods of this hybrid of numerical methods, complex system science and statistics which is usually referred to as Machine Learning (ML).

2. Machine Learning pros and cons

ML approaches have revolutionized problem-solving by offering speed and efficiency that traditional methods, such as solving the MHD equations in models or standard data analysis, often struggle to match. Solving complex differential equations can be computationally intensive and time-consuming, especially in scenarios with intricate variables or vast datasets. ML, on the other hand, excels in identifying patterns without explicit programming, allowing for the rapid analysis of large datasets or to reproduce the outcome of complex simulations. The inherent capacity of ML models to learn from data rather than rely on predefined equations contributes to their swiftness in extracting meaningful insights [8,9]. Moreover, ML's ability to handle nonlinear relationships and adapt to evolving patterns makes it particularly advantageous. In contrast, brute force approaches to data selection involve exhaustive trial and error, consuming substantial time and resources. ML, with its ability to discern relevant features and optimize model performance, accelerates the process.

Moreover, ML literature evolves exceptionally fast due to the rapid pace of research and technological breakthroughs in the several different disciplines where ML is employed. The competitive drive among researchers and the interdisciplinary nature of ML, integrating insights from computer science, mathematics, and various domains, further accelerate literature growth.

However, ML is highly sensitive to the Training Dataset and Loss Functions chosen due to its reliance on pattern recognition. The model learns from the provided data, making the training set's representativeness crucial for generalization. The choice of Loss Functions guides the optimization process, affecting how the model adjusts its parameters. Imbalanced or biased datasets can result in skewed learning, while the selection of inappropriate loss functions may lead to suboptimal model performance. Sensitivity to these factors underscores the importance of thoughtful data curation and precise loss function selection in ensuring ML models achieve accurate and meaningful results [10].

Recently, Deep learning (DL) models (a subset of ML models that specifically involve artificial neural networks with multiple layers, and designed to learn hierarchical representations of data through multiple non-linear transformations) have demonstrated remarkable success in forecasting space weather activity. By leveraging the power of neural networks, particularly recurrent and convolutional architectures, the DL methods efficiently analyze vast datasets of solar and geomagnetic observations. DL excels at capturing intricate patterns in time-series data, enabling accurate predictions of solar flares, geomagnetic storms, and other space weather phenomena [11]. However, they suffer from a common challenge in ML forecasting: the potential detachment from the physical understanding of system dynamics. While DL excels in prediction, it often lacks interpretability regarding how underlying quantities impact outcomes. The “black-box” nature of certain DL models may obscure the cause-and-effect relationships crucial for a comprehensive understanding. Striking a balance between prediction accuracy and preserving a physical perspective is paramount for us, Space Weather physicists, urging us to develop and employ models that not only forecast effectively, but also provide transparent insights into the intricate interplay of relevant quantities within the system.

3. The need to understand of a Physicist

In recent approaches within physics and space weather, there's a concerted effort to bridge the gap between ML predictions and the physical understanding of system dynamics. Researchers are exploring hybrid models that combine the predictive power of ML algorithms with traditional physics-based models. Additionally, efforts are directed towards developing explainable ML techniques, enabling practitioners to comprehend and trust DL models' decisions. This convergence of ML and physics reflects a nuanced approach, seeking to leverage the strengths of both disciplines for more robust and insightful predictions in complex systems like, of course, space weather.

3.1 DL interpretability

A possible method is to interpret and explain the decisions made by DL models, particularly in the context of neural networks equipped with attention mechanisms. Attention frames in this context refer to the parts of the input data that the model "focuses on" or deems most relevant when making predictions. Analyzing attention frames involves scrutinizing which parts of the input data the model pays attention to during different stages of its decision-making process. This analysis provides insights into how the model processes information, contributing to its overall explicability or interpretability, therefore addressing the "black box" problem. See [12] for an exemplar application of such an approach on the flare forecasting problem. By understanding which features of the input images the model prioritizes, we can gain a better understanding of why a particular decision was made, making DL models more transparent, and allowing us to focus on those quantities that the DL model deemed most relevant to the flaring process.

3.2 Physics-informed ML

Physics-informed ML (PI-ML) is an approach that combines principles from traditional physics-based modeling with the flexibility and predictive power of ML techniques. In this methodology, known physical laws and equations are incorporated into ML algorithms to guide and constrain their learning. This integration allows ML models to benefit from existing domain knowledge, ensuring that predictions align with the fundamental principles governing the system. PI-ML is particularly valuable in scenarios where data might be limited, noisy, or expensive to obtain. By infusing physics-based constraints, the models become more interpretable, and predictions become more trustworthy, as they adhere to the known laws of the system. This approach is being increasingly explored in various scientific fields, including fluid dynamics and climate science [13], where a balance between data-driven insights and physical principles is crucial.

3.3 An example of a PI-ML

In the following, I hypothesize a possible use of a combination of Complex Systems theory and Machine Learning to forecast time series with multiple embedded timescales. This is particularly valuable in domains like climate science and space weather, where diverse timescales often coexist due to various underlying processes acting at the same time. Integrating Complex Systems principles into the time series ingestion process enhances the capability of ML models to discern intricate patterns and relationships within the data. Specifically, we propose employing a technique to assess diverse intrinsic timescales and decompose the input time series into multiple modes, as in e.g. [14]. Subsequently, various methods, such as Recurrent Neural Networks as in e.g. [15], Long Short-Term Memory networks as in e.g. [16], or AutoRegressive Moving Average-based architectures, can be applied to capture dependencies and the varied timescales present in different modes. The selection of the optimal approach is made based on the distinctive characteristics and intrinsic timescales of each mode, allowing for a tailored and effective modeling strategy. As a result of this decomposition, ML models can predict each mode, each offering a distinct forecasting horizon [17]. Given the system's non-linearity, a mere summation of forecast time series might not capture the system's complexity entirely. Consequently, composing the modes might be better achieved using a suitable ML algorithm rather than a direct summation.

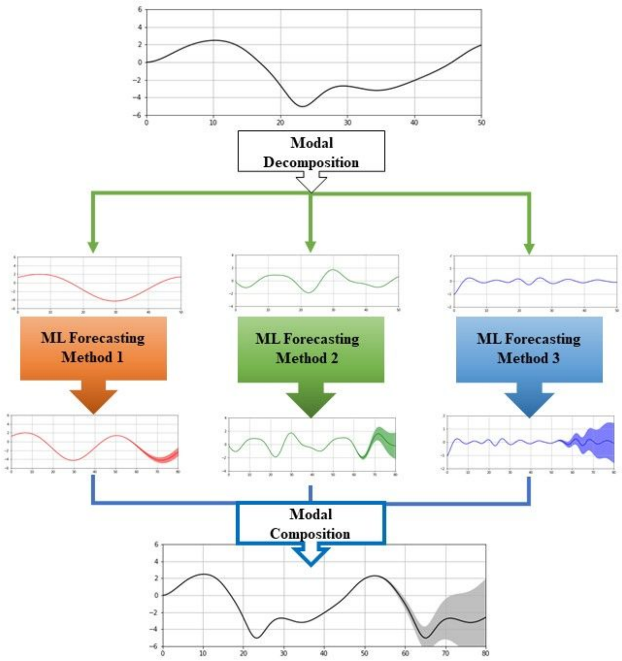

Figure 1. Architecture of the Modal Decomposition process. The input time series is decomposed into different modes with different timescales. The modes are sorted and the intrinsic time scales are identified via the relevant algorithm. The ML approach then forecast separately the various mode, defining different forecasting horizons. The various forecasts are then suitably aggregated to provide the forecast of the whole time series, with the reliability changing as a function of the forecasting time.

4. Conclusions

In conclusion, employing PI-ML methods for forecasting space weather or ionospheric activity presents a promising avenue with notable advantages. The integration of physical principles into ML models enhances interpretability, aligning predictions with our understanding of underlying mechanisms. This fusion allows for more transparent predictions, crucial in applications where the consequences of inaccuracies can be severe, such as in space-based technologies and communication systems. The success of this approach hinges on the accurate formulation of physical constraints within the model, demanding an in-depth understanding of the system's dynamics.

Striking the right balance between incorporating physics and exploiting the flexibility of ML poses a continuous challenge. Yet, the rewards may be substantial. PI-ML not only improves forecasting accuracy but also facilitates a deeper comprehension of space weather phenomena. As advancements continue in both physics and ML, the synergy between these domains holds the key to unlocking new insights and refining our ability to predict and mitigate the impact of space weather events on Earth and technological infrastructures orbiting it.

5. Acknowledgements

This research work has been a part of the Space Weather Awareness Training NETwork (SWATNet) project. SWATNet has received funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie Grant Agreement No. 955620. This research has been also carried out in the frame-work of the CAESAR project, supported by the Italian Space Agency and the National Institute of Astrophysics through the ASI–INAF n.2020-35-HH.0 agreement for the development of the ASPIS prototype of the scientific data centre for Space Weather. D.D.M. is grateful to the Italian Space Weather Community (SWICo).

References

[1] Odstrcil D. Modeling 3-D solar wind structure. Advances in Space Research. 2003 Aug 1;32(4):497-506.
[2] Pomoell J, Poedts S. EUHFORIA: European heliospheric forecasting information asset. Journal of Space Weather and Space Climate. 2018;8:A35.
[3] Napoletano G, Forte R, Del Moro D, Pietropaolo E, Giovannelli L, Berrilli F. A probabilistic approach to the drag-based model. Journal of Space Weather and Space Climate. 2018;8:A11.
[4] Dumbović M, Čalogović J, Vršnak B, Temmer M, Mays ML, Veronig A, Piantschitsch I. The drag-based ensemble model (DBEM) for coronal mass ejection

propagation. The Astrophysical Journal. 2018 Feb 26;854(2):180.

[5] Leka KD, Barnes G. Photospheric magnetic field properties of flaring versus flare-quiet active regions. IV. A statistically significant sample. The Astrophysical Journal. 2007 Feb 20;656(2):1173.

[6] Bobra MG, Couvidat S. Solar flare prediction using SDO/HMI vector magnetic field data with a machine-learning algorithm. The Astrophysical Journal. 2015 Jan 8;798(2):135.

[7] Cicogna D, Berrilli F, Calchetti D, Del Moro D, Giovannelli L, Benvenuto F, Campi C, Guastavino S, Piana M. Flare-forecasting algorithms based on high-gradient polarity inversion lines in active regions. The Astrophysical Journal. 2021 Jul 1;915(1):38.

[8] Benvenuto F, Campi C, Massone AM, Piana M. Machine learning as a flaring storm warning machine: Was a warning machine for the 2017 September solar flaring storm possible?. The Astrophysical Journal Letters. 2020 Nov 19;904(1):L7.

[9] Chierichini F, Campi C, Massone AM, Piana M. A Bayesian approach to the drag-based modelling of ICMEs. Journal of Space Weather and Space Climate 2024 Accepted.

[10] Guastavino S, Piana M, Benvenuto F. Bad and good errors: value-weighted skill scores in deep ensemble learning. IEEE transactions on neural networks and learning systems. 2022 Jul 1.

[11] Chierichini S, Jiajia L, Korsos MB, Del Moro D., Erdelyi, R. CME arrival Modelling with Machine Learning. The Astrophysical Journal. 2024 Accepted

[12] Francisco G, Berretti M, Chierichini S, Mugatwala R, Fernandes J, Barata T, Del Moro D., Erdelyi, R. Limits of solar flare forecasting models and new deep learning approach. JGR: Machine Learning and Computation. 2024 In preparation

[13] Kashinath K et al. Physics-informed machine learning: case studies for weather and climate modelling. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2021 Apr; 379: 2194

[14] Dragomiretskiy K, Zosso D Variational Mode Decomposition. IEEE Transactions on Signal Processing, 2014 Feb;62(3):531

[15] Shakir MM, Othman Z, Bakur AA Variational Mode DecompositionTEC Forecasting using Optimized Variational Mode
Decomposition and Elman Neural Networks. nternational Journal of Advanced Computer Science and Applications, 2022;13(7):482

[16] He F, Zhou Feng ZK, Liu G, Yang Y. A hybrid short-term load forecasting model based on variational mode decomposition and long short-term memory networks considering relevant factors with Bayesian optimization algorithm. Applied energy. 2019 Mar 1;237:103-16.

[17] Materassi M, Alberti T, Migoya-Orué Y, Radicella SM, Consolini G. Chaos and Predictability in Ionospheric Time Series. *Entropy*. 2023; 25(2):368. https://doi.org/10.3390/e25020368