Space Weather prediction
 USING VARIOUS TECHNIQUES, INCLUDING MACHINE LEARNING

Prof. Dario Del Moro
Dept of Physics
University of Rome “Tor Vergata”
Forecasting has always been a powerful tool

High efficiency vs Loss of explicability

ML is cheap + ML evolves extremely fast

Quick, nice results vs data limits & Deep Learning

We need to do better + we need to understand

The next steps
Forecasting has always been a powerful tool

We have honed this ability first in the eons of the biological evolution, then encoding it in our culture, ‘recently’ by applying the scientific method to our everyday life problems.

The mathematical tools to understand why some events are more difficult to forecast than others.

The development in the last decades in sciences of complex systems and statistics have allowed us to better define the limits of predictability and - often- to extend those limits.

BUT! The robust forecast of flare eruptions 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.
Scientific approach vs Machine Learning

- It is now feasible to compute the trajectory of plasma and magnetic field structures under the MHD equations in domains as large as the Heliosphere.

- Or try to predict -just by analyzing 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.

Solar flares originate from magnetically active regions (ARs)
Scientific approach vs Machine Learning

- Empirical, data-driven models and data fitting now enrolled into ML...


Scientific approach vs Machine Learning

- Exact space weather predictions are prevented either by intrinsic limitations or by the lack of knowledge about the present state.
- 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.
Scientific approach vs Machine Learning

- In the absence of a definitive physical theory explaining the mechanisms of an AR, the best hope for forecasting solar eruptions lies in finding an empirical relationship between some well chosen features of ARs and the solar flares and CMEs.

- As a consequence, we have 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).


ML is cheap + ML evolves extremely fast


ML is cheap + ML evolves extremely fast

ML: Quick, Nice results vs Correct Dataset use

- Neural Networks
- SVM
- Random Forest
- Gaussian models
- Convolutional NN
- ...

---

BEWARE: sometimes there are errors in:

- The hyperparameter tuning process
- How the samples are used
- How the cross-validation is used

---


Flares: Next steps?

- More info from B → by using topological describers
- Add the Time dimension → use B movies
- Higher layers info → use B + chromosphere + corona
- Adapt the Loss Functions → redefine error types
- Explicability of Deep Learning → analysis of attention frames
- Change of paradigm (e.g. Physics-Informed ML, ... )
Use of topological describers

Topologically complex ARs are strongly correlated to flare emissions.

Therefore a topological descriptor that counts the number of separated PILs fragments in the ARs.

**Feature ranking** analysis tells us it is extremely relevant to reach high skill scores → it brings relevant information on the AR flaring potential.

Analysis of time sequences

**Long-term Recurrent Neural Network**

LRCN = CNN + LSTM

BUT they get **similar TSS** values as more standard methods

→ HMI data do not contain enough information?

→ flares’ stochasticity hampers the possibility of binary predictions, in favor of a probabilistic prediction?

Considering the “memory” of the process

Guastavino, S., Marchetti, F., Benvenuto, F., Campi, C., & Piana, M. (2022). Implementation paradigm for supervised flare forecasting studies: A deep learning application with video data. Astronomy & Astrophysics, 662, A105. [https://doi.org/10.1051/0004-6361/202243617](https://doi.org/10.1051/0004-6361/202243617)

Use of more layers

Each layer used contributes to improved performances

Models combining SDO/AIA EUV images as inputs show improved performances compared to employing SDO/HMI photospheric magnetograms alone

Redefine error types/Performance Scores

Classical skill scores do not take into account the temporal distribution of the prediction

A possible solution:

- define new skill scores which allow ranking the prediction errors on the basis of their distribution along time

Two forecasts: same performance, but totally different predictive value.

Interpretability of CNN

Identify the most class-discriminative regions + the pixels contributing the most to the last convolutional layer
Change of paradigm

Regression instead of Classification

+ Put the Physics in the ML model


Beyond the Black Box: What is next?

- ML is cheap!
- ML is growing fast!
- ML is fixing its (major) drawbacks
- The data(-sets) are more and more important
- ML is here to stay?

Thank You!
CONCEPTS OF CONVOLUTIONAL NN

NN designed to work with images and/or regularly sampled data

Components:

- Convolutional layer
- Non-linear activation functions
- Pooling/downsampling
- Fully connected layer for classification
CONCEPTS OF LONG-TERM SHORT-TERM MEMORY

\[ h_{t-1} \] - hidden state at previous timestep t-1 (short-term memory)
\[ c_{t-1} \] - cell state at previous timestep t-1 (long-term memory)
\[ x_t \] - input vector at current timestep t
\[ h_t \] - hidden state at current timestep t
\[ c_t \] - cell state at current timestep t

- vector pointwise multiplication
- vector pointwise addition

- tanh activation function
- sigmoid activation function
- concatenation of vectors

Dobilasi LSTM Recurrent Neural Networks — How to Teach a Network to Remember the Past, Towards Data Science (towardsdatascience.com)
Forcing the Physics into ML approach

A complex signal, with many different timescales embedded.

→ We can try to mix different proxies to reproduce/forecast the signal.

OR:

→ We separate the signal's relevant scales into different modes, then reproduce/forecast the different modes mixing the proxies' modes (at or near the same timescales). Then, we reproduce/forecast the signal by combining the reproduced/forecast modes.

Applied to Black Carbon

An exercise: we know what we should get, we know how to interpret the results

We can verify which quantities and which modes are more relevant to reproduce the different timescales of the original signal by permutation analysis.

Applied to Thermospheric Density

A test: we expect the outcome and we interpret the results

We can learn which quantities and which modes are more relevant to reproduce the different timescales of the original signal by permutation analysis.