JOHNS HOPKINS
APPLIED PHYSICS LABORATORY

DASH

Discovering patterns, imbalanced classification & boundary surfaces in Heliophysics with artificial neural networks

Savvas Raptis[1]

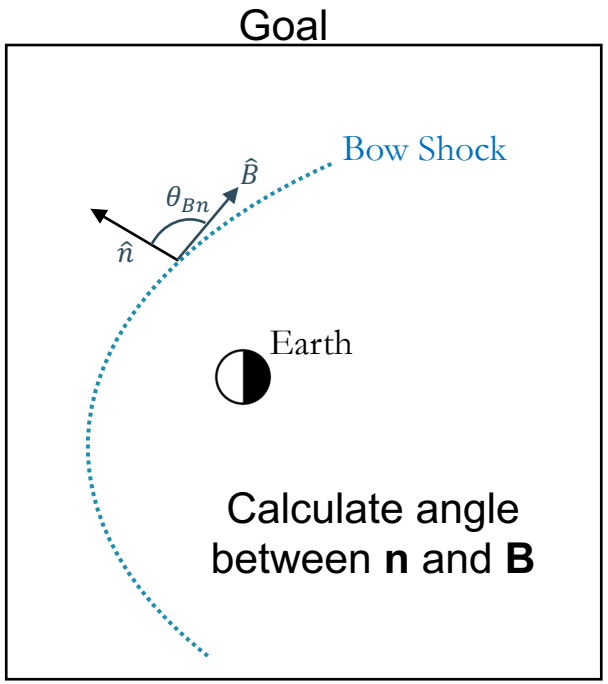
[1]APL/JHU, Laurel, MD, USA

SR acknowledges the support by John Hopkins University Applied Physics Laboratory independent R&D fund

savvas.raptis@jhuapl.edu / https://savvasraptis.github.io

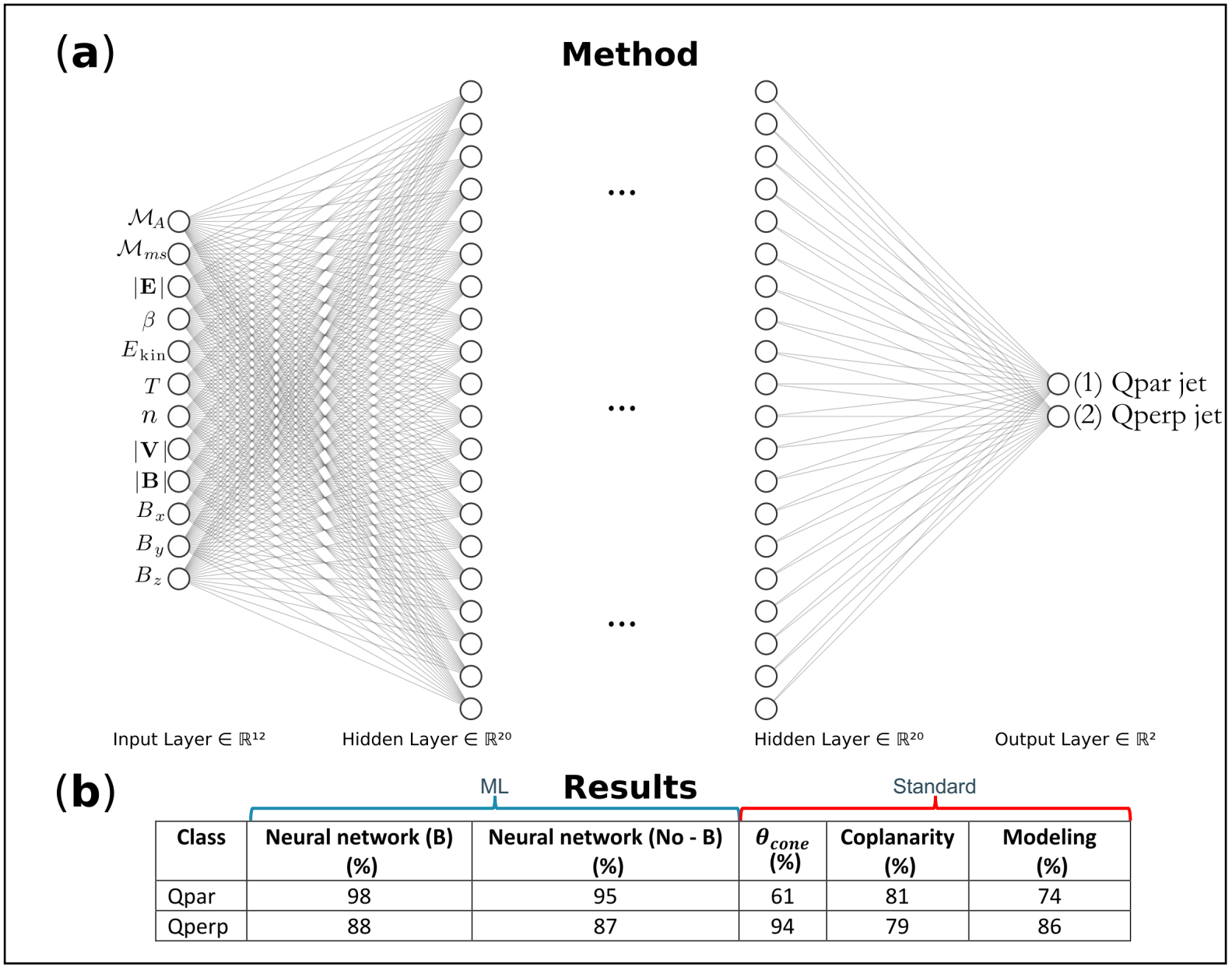
CGS | CENTER FOR GEOSPACE STORMS

Discovering patterns



Discussion

1. NN > standard techniques
2. Works without (**B**) → Patterns found with SW classes & properties



(b) **Results**

Class	Neural network (B) (%)	Neural network (No - B) (%)	θ_{cone} (%)	Coplanarity (%)	Modeling (%)
	ML		Standard		
Qpar	98	95	61	81	74
Qperp	88	87	94	79	86

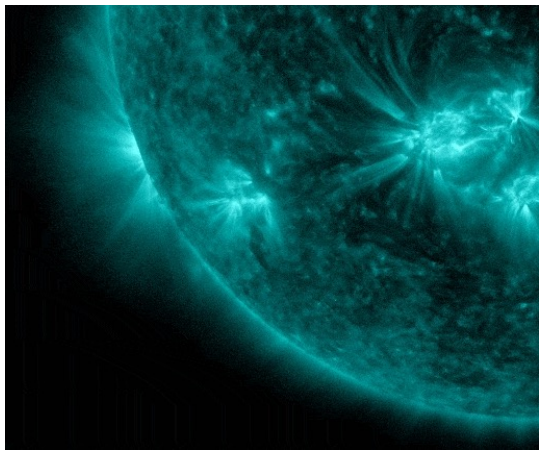
Raptis, Aminalragia-Giamini et al. (2020) | Front. Astron. Space Sci

Imbalanced learning

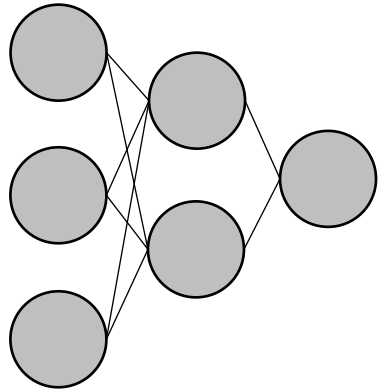
Goal

SEP forecasting associated with a flare using GOES SXR.

Occurrence of SEP based on X-rays of flares



24 features



	SEP predicted always YES	SEP predicted always NO
SEP occurred YES	191/220 [86.81 %]	19/220 [8.63%]
SEP occurred NO	[7.77 %]	[92.23%]

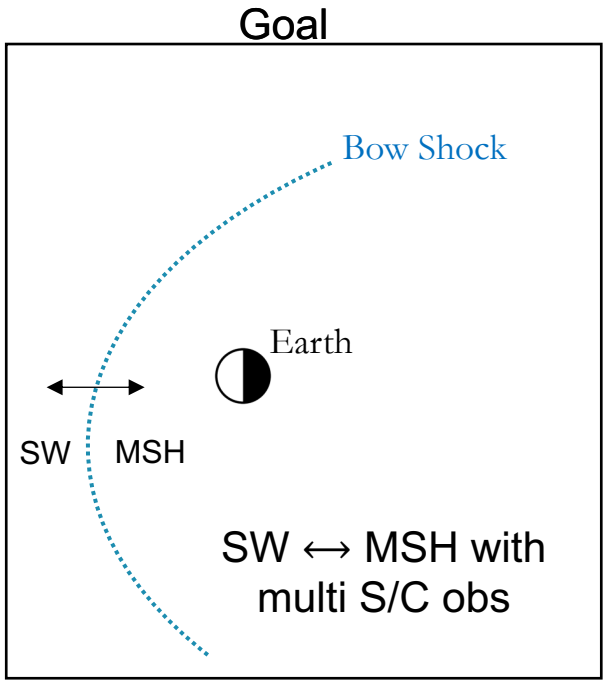
https://github.com/SavvasRaptis/machine-learning-examples

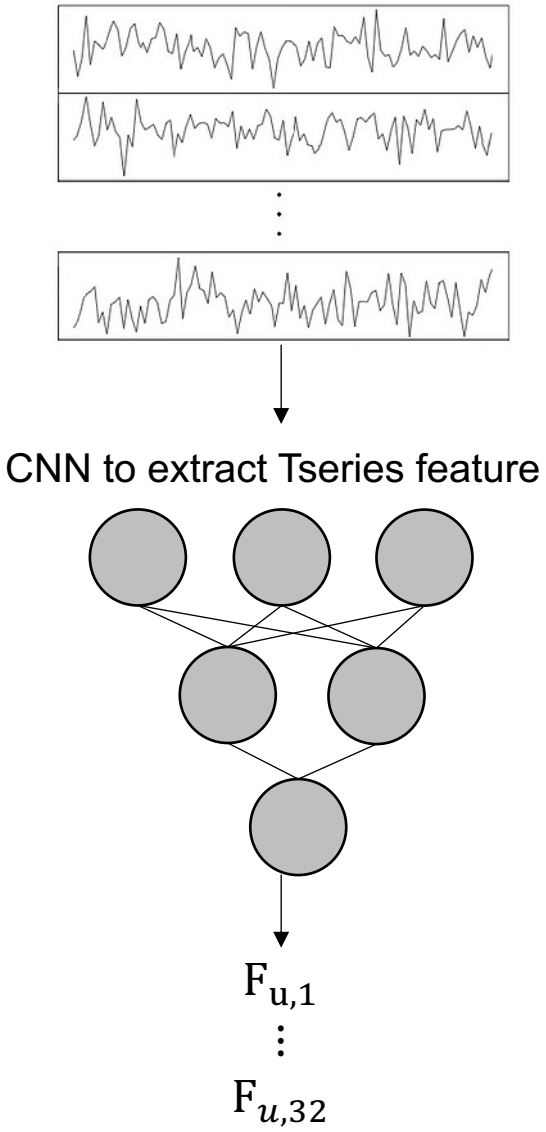
Discussion

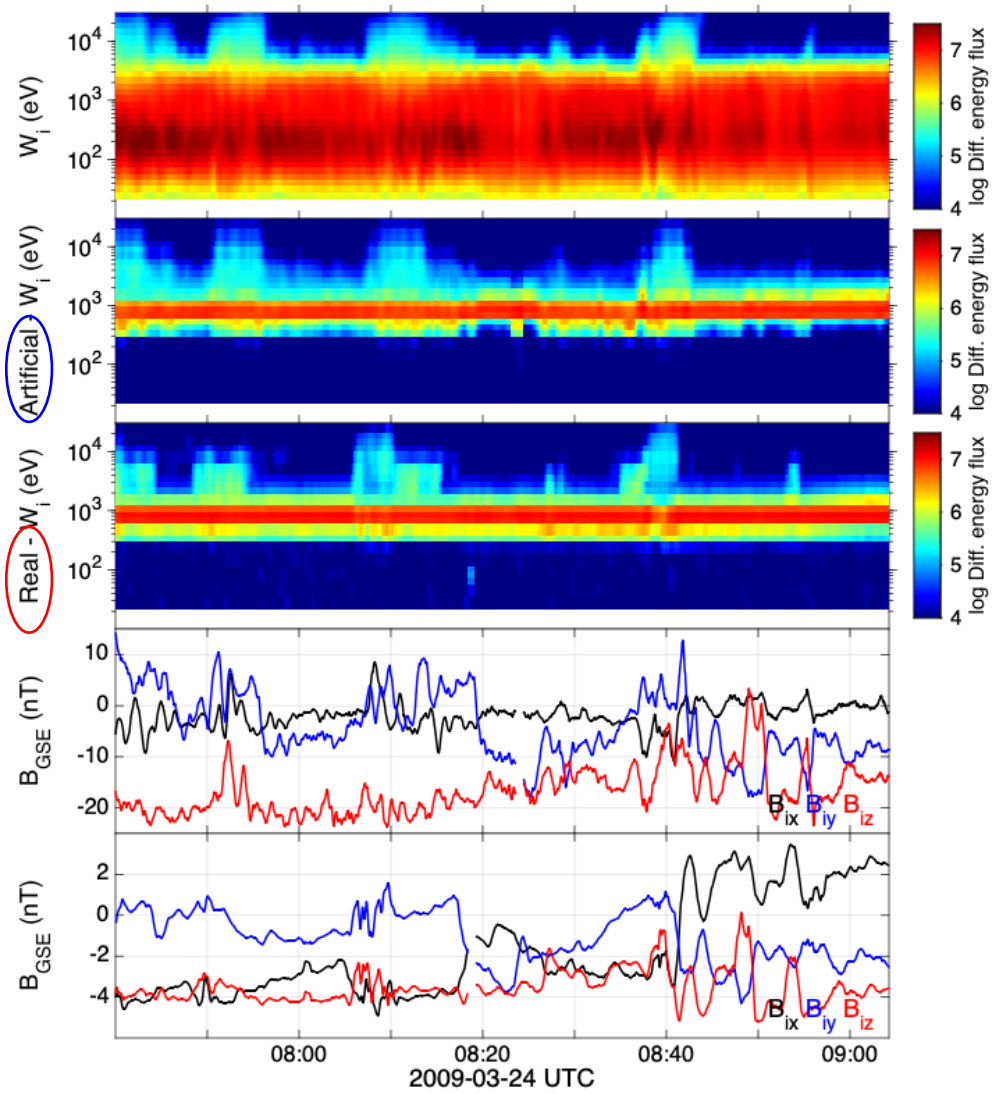
1. https://imbalanced-learn.org/stable/
2. Undersample / Oversample
3. Synthetic Data
4. Weight training / Weight loss function

Aminalragia-Giamini, **Raptis** et al. (2021) | J. Space Weather Space Clim

Evaluating boundary surfaces







Discussion

1. How do we treat variable time series?
2. What are some good evaluation metrics in these case?
3. (not shown) – Preprocess of data (90% of work)

Raptis et al. (ongoing)

Summary

1. Using various features & evaluating their effect → insights to unexplored patterns

2. Imbalanced learning has advanced a lot over the last years. Careful treatment of dataset& choice of methodology → Much better results.

3. How do we evaluate our results ? Experiment with different datasets, scales and phenomena needs special treatment. Pre-process is the most crucial aspect, let's discuss this more.