Predicting Solar Energetic Particle Event Occurrences Using Explainable AI

Yuta Kato[1], Kanya Kusano[2], Chihiro Mitsuda[1], Yasuhide Ishihara[1]
1. Cross Industry Business Office, Solution Transformation Unit, Fujitsu Limited, Kanagawa, Japan
2. Institute for Space-Earth Environmental Research, Nagoya University, Aichi, Japan

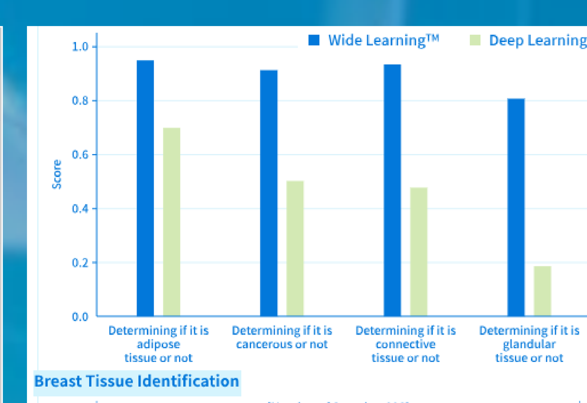
Summary

We conducted a classification task using Wide Learning™, an explainable AI developed by Fujitsu, to explore the conditions of Solar flares with/without SEP Events.
We created 57 features from GOES and SDO satellites data and the physics-based three-dimensional extrapolated magnetic fields developed by Kusano et al. (2020).
We classified Solar flares that meet > 10 MeV, > 10 pfu in the NOAA SWPC database during Solar Cycle 24 as positive samples, and all other ≧C1.2 Solar flares as negative samples.
We conducted 100 trials with random replacements and our model demonstrates a True Skill Statistic (TSS) approximately about 0.4.
For positive flares, we identified multiple useful conditions with numerical ranges which consist of the longitude, duration time, and history of SFs.
These results indicate the potential for real-time step by step SEP Event alerts and the ability to reference past cases that align with identified conditions.

1.Introduction

While existing SEP event forecasting models, including AI models, are discussed in Whitman et al. (2022), there are few studies utilizing explainable AI. We employed Wide Learning™, an explainable AI developed by Fujitsu, to perform classification tasks on SEP productive flares.

	Wide Learning™	Deep Learning
Operating principle	Process of scientific discovery	Simulation of a neural network
Suitable data type	Tabular data	Images and sound
Amount of data	From several dozen to several hundred records	At least one thousand to several tens of thousands of records
Explainability	XAI (explainable AI)	Black box
Output	Classifications, forecasting, and action plans	Classifications and forecasting
Hardware requirements	Computer with a generic CPU (Even a notebook can be used)	Parallel computers such as GPUs



Benchmark scores

$$h_\theta(g(x)) = \frac{1}{1+e^{-\boxed{\theta^T g(x)}}}$$

Logistic function

2.Methods

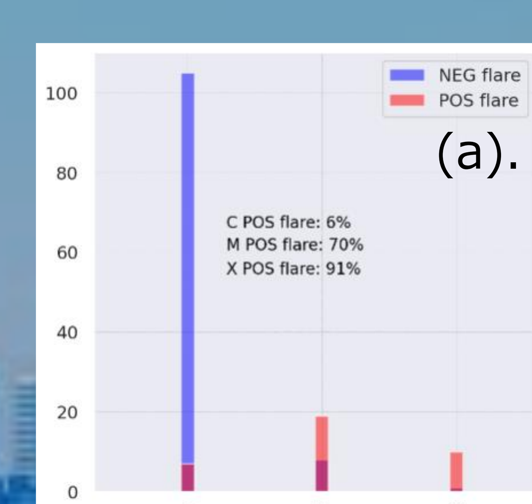
Data: Solar Cycle 24 (2010/1～2017/6)
+SEP POS flare: >10 MeV, >10 pfu NOAA/SWPC list 42 samples
+SEP NEG flare: GOES flare catalogue (≧C1.2) 6780 sample

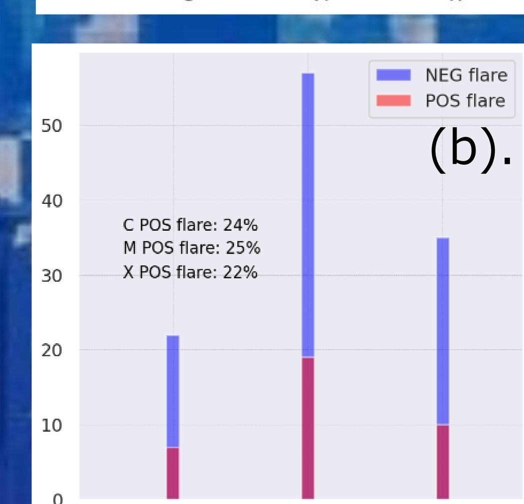
		True	
		POS	NEG
Prediction	POS	True Positive (TP)	False Positive (FP)
	NEG	False Negative (FN)	True Negative (TN)

$$TSS = \frac{TP}{TP+FN} - \frac{FP}{FP+TN}$$

Catalogue Creation:
- we set the ratio of positive to negative cases at 1:3, keeping the positive cases constant and creating 10 catalogs with (a) randomly sampled for each flare class, (b) 1:3 ratio for each C, M, and X flare class
- Each catalog was divided into a 4:1 ratio for training and testing, and testing was conducted 10 times.
- Evaluation method: We carried out evaluations using the True Skill Statistic (TSS).

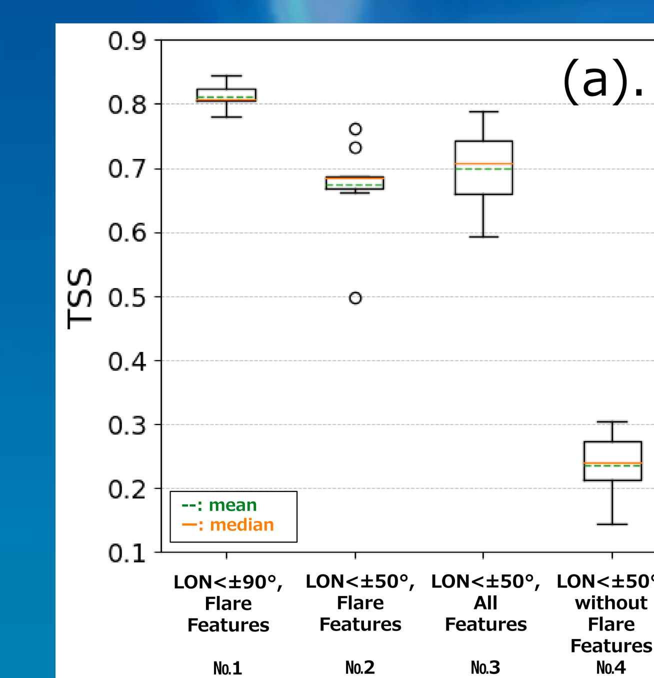
①Flare feature: 15 (#3)	②Active Region: 10 (#4)	③SHARP: 16 (#5)	④Kappa･PIL: 16 (#6)
<eruption> Flare durations Flare stand Flare decay SXR_MAX xrsb_duration xrsb_stand xrsb_decay <1day history> Xhis1d Mhis1d Chis1d Flarehis1d <all history> Xhis Mhis Chis Flarehis	<1day property> Latitude Longitude AREA of ARs CL PEN DIST LL NN HALE <eruption> xrsb_4h	<Active Regions> usflux meangam meangbt meangbz meangbh meanjzd totusjz meanalp meanjzh totusjh absnjzh savncpp meanpot totpot meanahr Shrgt45	<non-potential field> Er_Max Er_0<rc<1 Er_0<rc<2 Er_0<rc<3 Er_0<rc<4 Er_0<rc<5 Er_5<rc<10 E_HiFER_max E_HiFER_tot <Polarity Inversion Line> N_500 Lmax_500 Lt_500 N_1000 Lmax_1000 Lt_1000

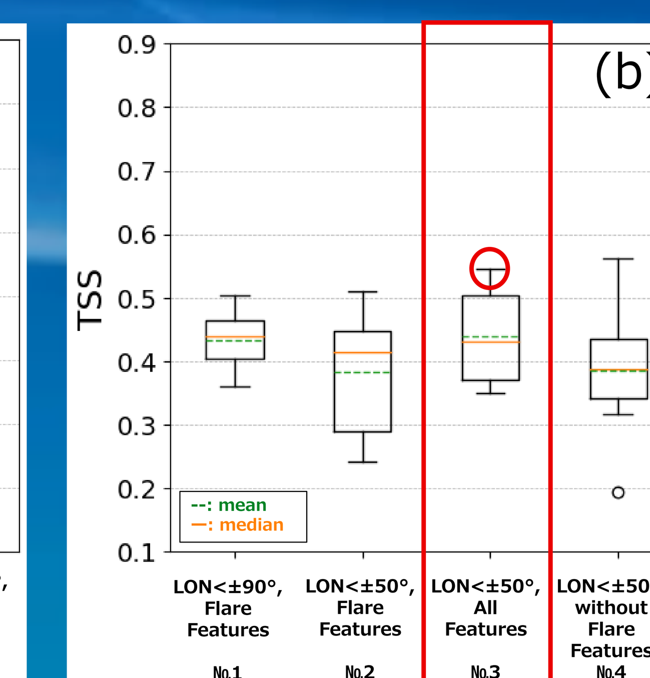




3.Results

Our model demonstrates a TSS of approximately 0.8 for case (a), and about 0.4 for case (b). In case (a), the X-ray peak intensity emerged as a significant weighted hypothesis for positive examples, while in case (b), the flare longitude, duration and flare history emerged as significant weighted hypotheses.





SEP Positive flare conditions (b-№3. TSS=0.55)	weight	POS samples	NEG samples
LON≥-43 ∧ Flare duration≥2910 ∧ Chis<38	1.03	13/17	0/39
LON≥-43 ∧Flare Stand≥390 ∧ Xhis<1 ∧ Chis1d<9	0.86	17/17	11/39
Flare Stand≥390∧ Xhis<1 ∧ Mhis<18 ∧ Chis1d<9	0.55	17/17	12/39
LON≥-18 ∧ Flare Stand≥390∧ Xhis<1 ∧ Chis1d<9	0.52	16/17	7/39

Based on the training data, we identified 1148 important combinations out of approximately 6 billion that represent positive cases and weighted 6 conditions as useful for prediction (displaying the top 4).

SEP Negative flare conditions (b-№3. TSS=0.55)	weight	NEG samples	POS samples
Mhis<15 ∧ meangam≥48.8 ∧ meanpot<17000 ∧ Lmax_500≥77	-1.86	33/39	0/17
Chis<58 ∧ meangam<59.2 ∧ meangbz≥98.0 ∧ N_1000≥2	-1.08	36/39	1/17
Flare duration<2910 ∧ N_500<180 ∧ N_1000≥2	-1.06	34/39	0/17
SXR<4.6e-05 ∧ meangbh≥48.8 ∧ Shrgt45<54.6 ∧ L_1000≥2	-0.68	30/39	0/17

Based on the training data, we identified 1005 important combinations out of approximately 6 billion that represent negative cases and weighted 23 conditions as useful for prediction (displaying the top 4).

	TP	FP	FN	TN	ACC	F1	Prec.	Rec.	AUC	PC	TSS	HSS	POD	FAR
LR	211	**117**	189	**993**	**0.80**	0.54	0.60	0.53	0.80	**79.80**	0.42	0.43	0.53	0.40
RF	197	**117**	203	**993**	0.79	0.53	**0.62**	0.49	**0.81**	78.80	0.39	0.41	0.49	**0.38**
XGB	227	178	173	932	0.77	0.55	0.57	0.57	0.77	76.79	0.41	0.40	0.57	0.43
LGBM	**240**	199	**160**	911	0.76	**0.56**	0.55	**0.60**	0.77	76.31	0.42	0.40	**0.60**	0.45
WL	224	135	176	975	0.79	**0.56**	**0.62**	0.56	0.80	79.44	**0.44**	**0.44**	0.56	**0.38**

Skill scores with b-№3 100 random replacements

4.Discussion

We can predict SEP event occurrences with each POS/NEG fitted condition and refer past similar cases, which indicate the potential for materials to consider real-time step by step SEP event alerts.

Examples).

AR＼Condition		№.1	№.2	№3	...	№.N	NEG Prob.	POS Prob.	Past similar cases
AR13xxx	54	1	1	1	1	0	1.00	0.00	AR11xxx(№.1,2,3)
AR13xxy	39	1	1	1	1	0	0.99	0.01	AR11xxx(№.2,3,4)
...	38	1	1	1	1	0	0.91	0.09	AR11xxx(№.3,4,5)
	20	0	0	0	0	0	0.02	0.98	AR11429(№.6,7,8)
AR13xxz	68	0	0	0	0	0	0.97	0.03	AR11xxx(№7,8,9)

Red Box: SEP Productive conditions
LON≥-43 ∧ Flare duration≥2910 ∧ Chis<38
Flare Stand≥390∧ Xhis<1 ∧ Mhis<18 ∧ Chis1d<9
Mhis<15 ∧ meangam≥48.8 ∧ meanpot<17000 ∧ Lmax_500≥77
SXR<4.6e-05 ∧ meangbh≥48.8 ∧ Shrgt45<54.6 ∧ L_1000≥2
Blue Box: SEP non-productive conditions
POS Weight > NEG Weight
prediction: SEP POS flare (98%)

1. Whitman et al. (2022)
2. Iwashita et al. (2020)
3. ftp://ftp.swpc.noaa.gov/pub/warehouse
4. Marroquin et al. (2023)
5. Bobra et al. (2014)
6. Kusano et al. (2020)

Contact: kato.yuta@fujitsu.com