Machine Learning as a Flaring Storm Warning Machine: Was a Warning Machine for the 2017 September Solar Flaring Storm Possible?

F. Benvenuto1, C. Campi1, A. M. Massone2, and M. Piana2

1 Dipartimento di Matematica, Università di Genova, Genova, Italy; benvenuto@dima.unige.it, campi@dima.unige.it
2 Dipartimento di Matematica, Università di Genova and CNR—SPIN, Genova, Italy; massone@dima.unige.it, piana@dima.unige.it

Received 2020 July 5; revised 2020 October 28; accepted 2020 October 29; published 2020 November 19

Abstract

Machine learning is now one of the methodologies of choice for flare forecasting, and supervised techniques, in both their traditional and deep versions, are becoming more frequently used for prediction in this area of space weather. Most studies assess the prediction effectiveness of machine-learning methods by computing confusion matrices, which are typically highly non-diagonal, particularly in applications concerning the forecasting of X-class flares. The present study suggests that the reliability of the outcomes of a supervised machine-learning method could be better assessed by using it as a warning machine, sounding binary alerts unrolled over time, and by comparing the number of alerts sounded by the machine in specific time windows with the number of events actually observed in those time windows. Indeed, when applied to the prediction of the events associated with the 2017 September solar storm, a hybrid LASSO algorithm was able to sound alerts every day a flare actually occurred; it also identified the corresponding flare class. In addition, the machine was able to predict with some accuracy a reliable proxy of the energy budget daily released by magnetic reconnection during the entire course of the storm. Finally, the analysis shows that the combination of sparsity-enhancing machine learning and feature ranking could allow the identification of the prominent role that the Ising energy played as an active region property in the forecasting process.

Unified Astronomy Thesaurus concepts: Solar flares (1496); Solar active regions (1974); Astronomy data modeling (1859)

1. Introduction

2017 September was the most flare-productive period of solar cycle 24: between September 6 and 10, 27 M-class flares and four X-class flares were emitted by the Sun, which correspondingly released several powerful coronal mass ejections (CMEs) and bursts of high-energy protons.

On the one hand, the impact of the 2017 September CMEs on Earth’s magnetosphere was timely and correctly predicted by the Space Weather Prediction Center of the National Oceanic and Atmospheric Administration (SWPC-NOAA).3

Between September 4 and 13 two R2 and three R3 alerts were released concerning radio blackouts and, in the same time range, the service sent one G3 (strong level) and one G4 (severe level) watch for CME-induced geomagnetic storms.

Solar physics (Tandberg-Hanssen & Emslie 1988) explains that solar flares are the major trigger of space weather. In the case of the 2017 September storm, SolarMonitor4 provided some probabilistic predictions of the sequence of flares that occurred during the storm by relying on McIntosh classification (Gallagher et al. 2002; Bloomfield et al. 2012; McCloskey et al. 2018) available for long-term data sets. SolarMonitor provides flare-occurrence probabilities starting from September 4 and, focusing on X-class events, it computed probability values around 9% until September 6, the day when an X2.2 flare and an X9.9 flare occurred. Then the SolarMonitor probabilities increased to around 20% on September 7 and 8 (although no X-class flares occurred on the latter day).

SolarMonitor predictions are very consolidated and should be considered as a reference landmark for flare forecasting. However, in the present Letter we test the capabilities of machine learning when applied to recently available data for the prediction of the 2017 September storm. Specifically, we consider whether it would have been possible to set up a warning machine able to sound binary alerts about the occurrence of the flares triggering the looming superstorm. We do this by applying supervised machine-learning techniques exploiting the rich (but more time-limited) information about the solar magnetic field configuration recorded by the Heliosseismic and Magnetic Imager on board the Solar Dynamics Observatory (SDO/HMI; Scherrer et al. 2012). Since 2010 February SDO/HMI has been providing both line-of-sight and vector magnetograms of the full solar disk at a (vector magnetogram) cadence of 12 minutes. In particular, for space weather applications, the Space-Weather HMI Active Region Patches (SHARP) process (Bobra et al. 2014) has been developed in order to automatically segment active regions (ARs) in HMI frames. The Solar Monitor Active Region Tracker algorithm (Higgins et al. 2011) has been developed in order to extract AR properties from SDO/HMI line-of-sight magnetograms. Other procedures that rely on different methodological approaches to determine physical, geometrical, and topological predictors have recently been introduced (Pariat et al. 2017; Guerra et al. 2018; Kontogiannis et al. 2018); such predictors have been extracted exploiting pattern recognition algorithms developed within the framework of the Horizon 2020 Flare Likelihood and Region Eruption Forecasting (FLARECAST) project (http://flarecast.eu).

As far as computation is concerned, several data analysis approaches in the last decade refer to machine learning as the general framework within which to pick up prediction algorithms (Li et al. 2007, 2008, 2020; Barnes & Leka 2008; Wang et al. 2008; Colak & Qahwaji 2009; Yu et al. 2009; Ahmed et al. 2013; Bobra & Couvidat 2015; Barnes et al. 2016; Murray et al. 2017; Sadykov & Kosovichev 2017; Benvenuto et al. 2018; Florios et al. 2018; Huang et al. 2018;
Massone et al. 2018; Park et al. 2018). The assessment of performance for these flare-forecasting algorithms is typically made by computing skill scores (Bloomfield et al. 2012), whose determination relies on confusion matrices. However, the values of these scores in the many experiments described so far in the literature, they are surprisingly systematically, distinctly (and often significantly) smaller than one, which indicates a far-from-optimal performance (Campi et al. 2019)

There are three likely reasons for these disappointing achievements of machine learning in flare forecasting.

First, the largest majority of properties extracted from HMI images include overlapping information, and this implies that the data sets of corresponding features are highly redundant for flare prediction (Campi et al. 2019). Second, solar-flaring storms rarely occur; more specifically, in the HMI archive there are very few of the X-class flares that often characterize this kind of extreme event (for example: the fraction of X-class flares in the historical HMI database corresponding to the range between 2012 September and 2016 April is 0.3% with respect to the total amount of C-class, M-class, and X-class flares). This implies that the training sets with which supervised machine-learning networks can be optimized are strongly imbalanced against X-class flares and specific approaches to generate more balanced training sets should be applied (Al-Ghraibah et al. 2015; Ahmadzadeh et al. 2019). Third, few machine- and deep-learning methods have been designed so far to exploit the dynamical information contained in AR data and, in any case, their performances typically suffer from the imbalanced nature of the training sets (Huang et al. 2018; Park et al. 2018, 2020; Angryk et al. 2020; Liu et al. 2019).

The main message of the present study is that there is an alternative way to interpret the prediction results of machine learning in flare forecasting. This is to consider them as the outcomes of a warning machine that takes point-in-time feature vectors as input data and sounds flare/no-flare alerts unrolled chronologically, thus using point-in-time information to realize dynamical predictions. Specifically, we show in the case of the 2017 September flaring storm that the performances of this warning machine can be assessed by comparing the number of sounded alerts with the number of observed events in specific time windows and, furthermore, that such information can be processed to predict, with some degree of accuracy, a proxy of the energy budget released day by day during the storm.

This Letter is organized as follows. Section 2 describes the forecasting experiment and provides some details about the features utilized in the training/prediction steps and the data analysis approach. Section 3 presents and discusses the data analysis results. Our conclusions are offered in Section 4.

2. Methods

This section describes the prediction experiment performed in the case of the 2017 September flaring storm, with specific focus on the way the training set is constructed, on the features extracted from the HMI images, and on the machine-learning method utilized in the data processing.

2.1. Setup of the Prediction Experiment

The forecasting experiment was conceived according to the following steps.

1. We constructed four sets of HMI images recorded at four issuing times in the range of days between 2012 September 14 and 2016 April 30. The four issuing times were 00:00 UT, 06:00 UT, 12:00 UT, and 18:00 UT.
2. We applied feature-extraction methods (Guerra et al. 2018) in order to extract 19 features from each of the ARs in the four sets described in the previous item. We added to them the features representing the longitude and latitude of each AR, a binary feature about flare occurrence in the previous 24 hr period, and the possible accumulated flare peak magnitude in the previous 24 hr period. Then, for each issuing time, we labeled with a “1” 23 dimensional feature vectors for which Geostationary Operational Environmental Satellites (GOES) recorded a flare of GOES class C1 and above (C1+ flares) in the 24 hr after the issuing time. We labeled with a “0” each feature vector for which GOES did not record any event in the 24 hr after the issuing time. This procedure allowed for the construction of four training sets for the C1+-class flares, each one for a specific issuing time. The same procedure was applied in order to construct four training sets for the prediction of flares of class M1 and above (M1+ flares) and of flares of class X1 and above (X1+ flares).
3. A machine-learning regularization network for classification was trained by means of the four training sets constructed for the prediction of C1+ flares. Specifically, we utilized the hybrid LASSO method discussed in Benvenuto et al. (2018). This same training step was repeated for the prediction of M1+ and X1+ events.

Figure 1 shows a sample of HMI frames recorded from 2017 August 29 to September 10. In this sequence, the image corresponding to August 30 has been highlighted in order to point out the appearance of AR 12673, a rather extended AR that exited the telescope’s field of view on September 9. Starting on August 30 at 00:00 UT (the first day that the AR was recorded in the NOAA database), we began extracting the 23 features from AR 12673, feeding the corresponding feature vector into the hybrid LASSO algorithm separately trained by using the training sets corresponding to the 00:00 UT issuing time for the prediction of C1+, M1+, and X1+ flares, and annotating whether a flare of corresponding class was predicted in the subsequent 24 hr period. We did the same for the remaining three issuing times, and then we continued with the 00:00 UT frame of August 31, and so forth. In this way, machine learning worked as a warning machine characterized by four time windows per day in which to decide whether or not to send an alert, and able to predict three possible lower bounds (C1+, M1+, X1+) for the energetic budget associated with each possible predicted event. We point out that this binary flare/no-flare behavior of the warning machine is made possible by the characteristics of its computational core. Indeed, hybrid LASSO is intrinsically a classification algorithm, which performs an automatic clustering of the LASSO regression outcomes in a way that depends on the historical data set used in the training phase, but not on skill scores (i.e., in such a way that no optimization of any specific skill score is needed).

2.2. Data and Data Features

Our analysis relied on SHARP data products in the HMI database. These data comprise 2D images of continuum intensity, the full three-component magnetic field vector, and the line-of-sight component of each HARP’s photospheric
extent. We then made use of property-extraction algorithms developed within the FLARECAST Horizon 2020 project to obtain 19 features; finally, four more properties were added, corresponding to the longitude and latitude of the AR, a binary label encoding the presence of a flare in the previous 24 hr period, and the flare index over the previous 24 hr period. Therefore, the machine-learning analysis was performed against 23 features overall, i.e., the ones presented in the first column of the left table in Figure 4.

2.3. Prediction and Feature Ranking Methods

The machine learning utilized for data analysis in this Letter is a hybrid version of the LASSO regression method (Benvenuto et al. 2018; Campi et al. 2019).

This approach applies a fuzzy clustering algorithm to the outcomes of a regression method, namely LASSO (Tibshirani 1996). In this way binary classification is realized in a way that is independent of the optimization of a specific skill score, i.e., fuzzy clustering automatically classifies the flare/no-flare events by means of a data-adaptive and skill-score-independent thresholding procedure. Further, LASSO is a sparsity-promoting algorithm and this supports feature selection. Indeed, once the machine-learning method has been applied to the input data, predictors are ranked by using Recursive Feature Elimination (RFE). This iterative procedure can be summarized in three recursive steps (Guyon et al. 2002): the training of the classifier; the computation of the ranking for all features; and the elimination of the feature with smallest ranking.

3. Results

Once the storm ended, we began assessing the effectiveness of our warning machine. The first task was to compare the alerts sounded by the warning machine with the actual events observed by GOES. The top-left panel of Figure 2 puts on a time axis the 24 hr warnings raised by the algorithm, distinguished with respect to the three flare classes and the four issuing times: therefore, this panel is the pictorial representation of how the warning machine works. The top-right panel of Figure 2 contains the actual observations of events recorded by GOES along the same time range, together with the corresponding intensities. If we superimpose the two panels, we obtain, in the bottom panel of Figure 2, the pictorial representation of the possible matching between predictions and observations.

Standard practice suggests that these data need to be computed according to the corresponding confusion matrices for prediction and observations; the results of this can be found in Table 1. These numbers show that the confusion matrices for X1+ flares are highly non-diagonal, which can be explained by the notable imbalance of the training set and which systematically implies sub-optimal skill scores values. Yet, solely considering confusion matrix values does not do justice to the forecasting ability of machine learning. Indeed, from Figure 2, we can state the following.

1. The warning machine was completely quiet until 2017 September 2 at 12:00:00 UT.
2. At that time, the machine began sending warnings concerning flares with different intensities.
3. On 2017 September 9, 00:00 UT, the algorithm stopped sending warnings (but note that on that day AR 12673 exited from the field of view of HMI).
4. GOES began recording flares at the end of the day on 2017 September 3, and kept on annotating flaring explosions with different flare classes every day that AR 12673 was visible to HMI, and also for the following
days. Specifically, two X1+ flares were caught on September 6 and September 7.

In summary, this means that the warning machine correctly stayed quiet for all days in which no events occurred, and did not miss a single day in which a flare did occur. This result is consistent with the fact that AR 12673 is an emerging AR evolving rather rapidly into more complex configurations.

The warning machine can provide more quantitative information if one focuses on the prediction of lower bounds for the energetic budget associated with the emission. To this

Table 1
Confusion Matrices Concerning the Prediction of the 2017 September Flaring Storm According to the Three Different Flaring Classes and the Four Issuing Times Considered for the Training Step

| Flare Observed | above C |  |  |  |  |  |  |
|----------------|---------|---------|---------|---------|---------|---------|---------|
|                | 00:00:00 UT | 06:00:00 UT | 12:00:00 UT | 18:00:00 UT |
| Flare predicted | yes | no | yes | no | yes | no | yes | no |
| yes | 5 | 1 | 6 | 0 | 6 | 1 | 7 | 0 |
| no | 0 | 5 | 0 | 5 | 0 | 4 | 0 | 4 |

| above M | 00:00:00 UT | 06:00:00 UT | 12:00:00 UT | 18:00:00 UT |
| Flare predicted | yes | no | yes | no | yes | no | yes | no |
| yes | 5 | 1 | 5 | 1 | 5 | 2 | 5 | 2 |
| no | 0 | 5 | 0 | 5 | 0 | 4 | 0 | 4 |

| above X | 00:00:00 UT | 06:00:00 UT | 12:00:00 UT | 18:00:00 UT |
| Flare predicted | yes | no | yes | no | yes | no | yes | no |
| yes | 0 | 4 | 1 | 2 | 1 | 3 | 0 | 2 |
| no | 0 | 7 | 0 | 8 | 0 | 7 | 0 | 9 |
aim, in the top panel of Figure 3 we present GOES data in the 1–8 A channel (which may serve as a proxy of the released energy) and, for each day from 2017 September 4 to 9 we summed up the peak flux of all C1+ flares that occurred on that day, measured in watts per square meter. In the bottom panel of Figure 3 we compare this peak flux profile with the alerts sounded by machine learning for an X1+ flare: this panel shows that, with the exception of 2017 September 8, alerts sounded by the warning machine on all other days in the range from 2017 September 4 to 9 the X1+ corresponded to overall daily peak fluxes measured by GOES greater than $10^{-4}$ W m$^{-2}$, i.e., the ones associated to X-class flares.

The final part of our study was devoted to investigating the role played by the different features in the prediction process. We applied RFE to the LASSO outcomes twice: first, by considering the same training step utilized for the forecasting obtained in Figure 2; then, by considering a slightly enriched training set, in which we also added the features extracted from AR 12673 recorded by HMI during the time range from 2017 August 30 to September 9. The results of this feature-ranking process are illustrated in Figure 4, which compares the average positions of each feature in the two conditions and the corresponding standard deviations, where average and standard deviation values are computed with respect to the ranks obtained by each feature over the four issuing times considered in the analysis. The two tables in the figure show the notable robustness of the asset of features utilized for training, which is an important confirmation of the reliability of the data analysis process. Interestingly, feature number 11, corresponding to the Ising energy, has the greatest difference in the ranks obtained in the case of the two training sets: specifically, this feature performs five steps forward (from rank 11 to rank 6) when the data concerning the flaring storm are included in the training set. A possible interpretation of this fact is that, when the training set contains feature vectors associated to intense events like the ones that occurred during the 2017 September flaring storm, the weight with which the Ising energy feature contributes to the prediction increases rather significantly.

The validation of results concerned with the prediction of the 2017 September flares is made difficult by the fact that there is little HMI data corresponding to flaring storms and that, as previously mentioned, training sets relying on the HMI database are imbalanced against X-class flares. However, we have repeated the same analysis in the case of a sequence of flaring emissions that occurred in the time range from 2015 March 7 and March 17 and associated with AR 12297 (note that GOES revealed flares associated with that AR starting from March 5 but, due to technical reasons, the FLARECAST database did not contain any data corresponding to that day or to March 6). We applied to those data the same hybrid LASSO network, trained this time on the data corresponding to the time range from 2012 September 14 to 2015 March 4. On the one hand, we can see that the confusion matrices for X-class flares presented in Table 2 are even more dramatically non-diagonal than the ones associated with the 2017 September storm. On the other hand, the left panel in Figure 5 shows that the warning
machine sent alerts on all of the days in the time range, consistent with GOES observations. As far as the prediction of the released energy budget is concerned, the machine warns about the above X-class flares in the four-day range from 2015 March 9 to March 12, i.e., for days in which, according to GOES, the Sun produced an overall energetic activity equivalent to an X1+ flare. However, from 2015 March 13 to March 17 the machine predicted X1+ flares each day, while the energy budget proxy measured by GOES was clearly smaller. We point out that the training set used for this second analysis is even more imbalanced than the one used in the 2017 September storm case: indeed, the frequency of X-class flares is 0.1%, averaged over the four issuing times, compared to 0.3% corresponding to the training set used for the previous prediction.

4. Conclusions

This study suggests that, in the case of flaring events that last for some days, the information extracted by machine learning from HMI data may be accurate enough to set up a warning machine able to sound reliable alerts on a daily scale, in a way that is similar to what is done by the SWPC-NOAA watches for geomagnetic storms. In 2017 September, i.e., toward the minimum of solar cycle 24, a huge flaring storm unexpectedly occurred. If this warning machine had been available at that time, it would have behaved in a rather reliable way, staying quiet and sounding warnings consistent with the Sun’s activity in the subsequent 24 hr period. Also, the machine would have been able to estimate, at the same daily scale, a lower bound for the total amount of energy released by the flaring events.

The machine-learning tool utilized in our warning machine is a hybrid LASSO algorithm, i.e., a method that allows the automatic clustering of the regression outcomes in order to realize binary classification and the selection of data features that most greatly impact the prediction process. Hybrid LASSO performs binary classification in a way that does not depend on the optimization of some specific skill score; of course, there may be applications where it is more appropriate to adopt thresholding techniques relying on a specific risk assessment. The feature-ranking process identified the Ising energy as the image property with the highest sensitivity for the training step, and this confirms that forecasting methods should consider the energy budget stored in the magnetic field configuration as a crucial parameter of interest. Therefore, it is probably true that the intrinsic stochasticity of the flaring phenomenon and the notable redundancy of information contained in the observations at disposal prevent satisfactory prediction performances on a point-in-time basis (Campi et al. 2019). However, it also seems true that the representation of the machine-learning outcomes as a set of binary alerts unrolled over time allows the use of the algorithm as a reliable warning machine in the case of flaring storms. In addition to this, the result of the feature-ranking-process suggests that there is a promising research direction for machine-learning experts. This is the invention of reliable multi-task flare prediction.

Figure 4. Result of the feature-ranking process performed by applying RFE on the LASSO regression outcomes. Left panel: the training set used for prediction, i.e., SHARP data in the time range between 2012 September 14 and 2016 April 30. Right panel: the training set obtained by enriching the previous data set with SHARP data of AR 12673 in the time range between 2017 August 30 and September 9. Feature 11 is the only one able to perform five steps forward, from rank 11 to rank 6.

Table 2
Confusion Matrices Concerning the Prediction of X1+ Flares for the 2015 March Event in the Case of the Four Issuing Times Considered for the Training Step

| Flare Observed | 00:00:00 | 06:00:00 | 12:00:00 | 18:00:00 |
|---------------|---------|---------|---------|---------|
|               | UT      | UT      | UT      | UT      |
| Flare predicted | yes | no | yes | no | yes | no | yes | no |
| yes           | 0      | 5      | 0      | 8      | 1      | 6      | 0      | 0      |
| no            | 0      | 5      | 0      | 2      | 0      | 3      | 0      | 10     |

Table 2
Confusion Matrices Concerning the Prediction of X1+ Flares for the 2015 March Event in the Case of the Four Issuing Times Considered for the Training Step

| feature | rank (mean) | rank (sd) |
|---------|-------------|-----------|
| Falconer’s gradient-weighted length of neutral line | 3 | 2 |
| heliographic longitude of SHARP centroid | 3 | 2 |
| heliographic latitude of SHARP centroid | 4 | 2 |
| binary flag for flare occurrence in previous 24 hr | 5 | 2 |
| accumulated GOES flare peak magnitudes in previous 24 hr | 6 | 2 |
| total length magnetic pol. inv. lines (MPLs) | 7.3 | 4.2 |
| multifractal structure function integral range index | 7.3 | 1.9 |
| total unsigned flux around all MPLs | 7.5 | 1.7 |
| Schrödinger’s R (log10 form) | 8.3 | 1.9 |
| sum of horizontal magnetic gradient | 8.5 | 7.5 |
| Ising energy (pixelwise) | 10.5 | 5.7 |
| maximum length of single MPL | 10.8 | 0.9 |
| multifractal gen. correlation dimension spectrum | 11.8 | 0.9 |
| Ising energy (calculated using Belft flux partitions) | 12.8 | 0.9 |
| sep. distance between leading and following polarity subgroups | 14.5 | 0.6 |
| Ising energy (calculated using Belft flux partitions) | 13.8 | 0.6 |
| Ising energy (calculated using Belft flux partitions) | 13.8 | 0.6 |
| Ising energy (calculated using Belft flux partitions) | 13.8 | 0.6 |
| Ising energy (calculated using Belft flux partitions) | 13.8 | 0.6 |
| Ising energy (calculated using Belft flux partitions) | 13.8 | 0.6 |

| feature | rank (mean) | rank (sd) |
|---------|-------------|-----------|
| heliographic longitude of SHARP centroid | 2.5 | 2.4 |
| heliographic latitude of SHARP centroid | 3.5 | 2.4 |
| binary flag for flare occurrence in previous 24 hr | 4.5 | 2.4 |
| Falconer’s gradient-weighted length of neutral line | 5 | 1 |
| accumulated GOES flare peak magnitudes in previous 24 hr | 5.5 | 2.4 |
| Ising energy (pixelwise) | 7.3 | 6.7 |
| total length magnetic pol. inv. lines (MPLs) | 7.3 | 4.2 |
| multifractal structure function integral range index | 7.3 | 1.9 |
| total unsigned flux around all MPLs | 7.5 | 1.7 |
| Schrödinger’s R (log10 form) | 8.3 | 1.9 |
| maximum length of single MPL | 10.8 | 0.9 |
| multifractal gen. correlation dimension spectrum | 11.8 | 0.9 |
| Ising energy (calculated using Belft flux partitions) | 12.8 | 0.9 |
| sep. distance between leading and following polarity subgroups | 14.3 | 0.5 |
| Ising energy (calculated using Belft flux partitions) | 13.8 | 0.6 |
| Ising energy (calculated using Belft flux partitions) | 13.8 | 0.6 |
| Ising energy (calculated using Belft flux partitions) | 13.8 | 0.6 |
| Ising energy (calculated using Belft flux partitions) | 13.8 | 0.6 |
| Ising energy (calculated using Belft flux partitions) | 13.8 | 0.6 |
| Ising energy (calculated using Belft flux partitions) | 13.8 | 0.6 |
| Ising energy (calculated using Belft flux partitions) | 13.8 | 0.6 |
| Ising energy (calculated using Belft flux partitions) | 13.8 | 0.6 |

The Astrophysical Journal Letters, 904:L7 (7pp), 2020 November 20 Benvenuto et al.
algorithms that take as input a selection of image properties related to the magnetic energy stored in the magnetograms and provide as output both the binary flare/no-flare classification and some quantitative prediction of the amount of energy actually released during a temporal scale, ranging from some hours to a whole day.

We acknowledge financial contribution from the agreement ASI-INAF No. 2018-16-HH.0.

**ORCID iDs**

F. Benvenuto [https://orcid.org/0000-0002-4776-0256](https://orcid.org/0000-0002-4776-0256)

**References**

Ahmadzadeh, A., Aydin, B., Kempton, D. J., et al. 2019, in 2019 18th IEEE Int. Conf. On Machine Learning And Applications (ICMLA) (Piscataway, NJ: IEEE), 1814

Ahmed, O. W., Qahwaji, R., Colak, T., et al. 2013, *SoPh*, 283, 157

Al-Ghraibah, A., Boucheron, L. E., & McAteer, R. J. 2015, in 2015 IEEE Int. Conf. on Data Mining Workshop (ICDMW) (Piscataway, NJ: IEEE), 557

Angryk, R. A., Martens, P. C., Aydin, B., et al. 2020, *NatSD*, 7, 227

Barnes, G., & Leka, K. D. 2008, *ApJL*, 688, L107

Barnes, G., Leka, K. D., Schrijver, C. J., et al. 2016, *ApJ*, 829, 89

Benvenuto, F., Piana, M., Campi, C., & Massone, A. M. 2018, *ApJ*, 853, 90

Bloomfield, D. S., Higgins, P. A., McAteer, R. T. J., & Gallagher, P. T. 2012, *ApJL*, 747, L41

Bobra, M. G., & Couvidat, S. 2015, *ApJ*, 798, 135

Bobra, M. G., Sun, X., Hoeksema, J. T., et al. 2014, *SoPh*, 289, 3549

Campi, C., Benvenuto, F., Massone, A. M., et al. 2019, *ApJ*, 883, 150

Colak, T., & Qahwaji, R. 2009, *SpWea*, 7, S06001

Florios, K., Kontogiannis, I., Park, S. H., et al. 2018, *SoPh*, 293, 28

Gallagher, P. T., Moon, Y. J., & Wang, H. 2002, *SoPh*, 209, 171

Guerra, J. A., Park, S. H., Gallagher, P. T., et al. 2018, *SoPh*, 293, 1

Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. 2002, *Machine Learning*, 46, 389

Higgins, P. A., Gallagher, P. T., McAteer, R. T., & Bloomfield, D. S. 2011, *AdSpR*, 47, 2105

Huang, X., Wang, H., Xu, L., et al. 2018, *ApJ*, 856, 7

Kontogiannis, I., Georgoulis, M. K., Park, S. H., & Guerra, J. A. 2018, *SoPh*, 293, 1

Li, R., Cui, Y., He, H., & Wang, H. 2008, *AdSpR*, 42, 1469

Li, R., Wang, H.-N., He, H., Cui, Y.-M., & Du, Z.-L. 2007, *ChJAA*, 7, 441

Li, X., Zheng, Y., Wang, X., & Wang, L. 2020, *ApJ*, 891, 10

Liu, H., Liu, C., Wang, J. T. L., & Wang, H. 2019, *ApJ*, 877, 121

Massone, A. M., Piana, M. & The Flarecast Consortium 2018, *Machine Learning Techniques for Space Weather* (Amsterdam: Elsevier)

McCloskey, A. E., Gallagher, P. T., & Bloomfield, D. S. 2018, *JSWSC*, 8, A34

Murray, S. A., Bingham, S., Sharpe, M., & Jackson, D. R. 2017, *SpWea*, 15, 577

Pariat, E., Leake, J. E., Valori, G., et al. 2017, *A&A*, 601, A125

Park, T., Moon, Y. J., Shin, S., et al. 2018, *ApJ*, 869, 91

Park, S. H., et al. 2020, *ApJ*, 890, 124

Sadykov, V. M., & Kosovichev, A. G. 2017, *ApJ*, 849, 148

Scherrer, P. H., Schou, J., Bush, R. L., et al. 2012, *SoPh*, 275, 207

Tandberg-Hanssen, E., & Emslie, A. G. 1988, *The Physics of Solar Flares* (Cambridge: Cambridge Univ. Press)

Tibshirani, R. 1996, *Journal of the Royal Statistical Society, Series B (Methodological)*, 58, 267

Wang, H. N., Cui, Y. M., Li, R., Zhang, L. Y., & Han, H. 2008, *AdSpR*, 42, 1464

Yu, D., Huang, X., Wang, H., & Cui, Y. 2009, *SoPh*, 255, 91