Space Weather of the Heliosphere: Processes and Forecasts
Proceedings IAU Symposium No. 335, 2017
C. Foullon & O.E. Malandraki, eds.
© International Astronomical Union 2018
doi:10.1017/S1743921317007293

Solar Flare Prediction Using Machine Learning with Multiwavelength Observations

Naoto Nishizuka, Komei Sugiura, Yuki Kubo, Mitsue Den, Shin-ichi Watari and Mamoru Ishii

National Institute of Information and Communications Technology,
4-2-1 Nukui-Kitamachi, Koganei, Tokyo, Japan 184-8795
email: nishizuka.naoto@nict.go.jp

Abstract. We developed a flare prediction model based on the supervised machine learning of solar observation data for 2010-2015. We used vector magnetograms, lower chromospheric brightening, and soft-X-ray data taken by Solar Dynamics Observatory and Geostationary Operational Environmental Satellite. We detected active regions and extracted 60 solar features such as magnetic neutral lines, current helicity, chromospheric brightening, and flare history. We fully shuffled the database and randomly divided it into two for training and testing. To predict the maximum size of flares occurring in the following 24 hours, we used three machine-learning algorithms independently: the support vector machine, the k nearest neighbors (kNN), and the extremely randomized trees. We achieved a skill score (TSS) of greater than 0.9 for kNN. Furthermore, we compared the prediction results in a more operational setting by shuffling and dividing the database with a week unit. It was found that the prediction score depends on the way the database is prepared.

Keywords. Sun: flares, magnetic fields; sunspots, methods: statistical, techniques; image processing

1. Introduction

Solar flares are an origin of space weather phenomena. They originate from the stored magnetic energy around sunspots, and the impulsive release of energy produces flares. Since the magnetic flux emerges from the bottom of the photosphere, solar flares can be predicted by monitoring the photospheric magnetic field in principle. Larger flares tend to occur when a sunspot is larger and more complex (e.g., Leka & Barnes 2003, Bloomfield *et al.* 2012, McCloskey et al. 2016). The repetition of flares is frequently observed. It has been shown by theory and observations that an emerging flux around magnetic neutral lines triggers a flare. Solar flares have been predicted by humans for a long time. However, the prediction accuracy has not significantly improved, despite the increase in the amount of high-resolution observation data available in near real-time.

On the other hand, supervised machine learning, which has recently been developed, is advantageous for dealing with a huge amount of data. New prediction models of solar flares using machine learning have been intensively studied (e.g., Ahmed 2013, Bobra & Couvidat 2015, Nishizuka *et al.* 2017, Raboonik *et al.* 2017, Liu *et al.* 2017, and references therein). Machine learning can tell us which feature is most effective for predicting flares in real-time operation. Several prediction models of flares have been developed (e.g., Barnes *et al.* 2017), but it is still difficult to compare the performance of each prediction model in the same way. There is no standard way of evaluating the prediction models in an operational setting. In this paper, we have attempted to solve these remaining problems.

https://doi.org/10.1017/S1743921317007293 Published online by Cambridge University Press

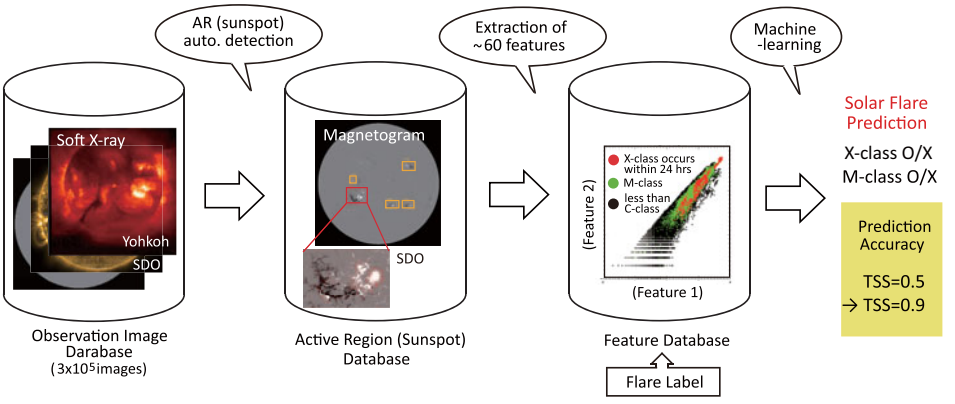

Figure 1. Flow chart of our model of flare prediction.

2. Flare Prediction Model and Observation Database

Figure 1 shows a flow chart of our flare prediction model. First, we compiled an original observation database for the period 2010-2015 by downloading the Solar Dynamics Observatory (SDO; Pesnell *et al.* 2012) data archives, i.e., the line-of-sight and vector magnetograms taken by the Helioseismic and Magnetic Imager (HMI; Scherrer *et al.* 2012) on board SDO, the lower chromospheric brightening taken by the 1600 Å filtergram of the Atmospheric Imaging Assembly (AIA; Lemen *et al.* 2012) on board SDO, and the soft-X-ray emission taken by the Geostationary Operational Environmental Satellite (GOES). Next, we automatically detected active regions from full-disk line-of-sight magnetograms using a threshold of 1.4×10^{-2} T (see Nishizuka *et al.* 2017 more in detail) and attached IDs to them. We identified a rotating active region in the time sequence with the same ID. The original data were reduced by employing one hour cadence. We attached flare labels to each active region in the time sequence, for which a larger flare than a GOES M-class flare occurred in the following 24 hours. Finally, we predicted the maximum size of flares occurring in the following 24 hours by using the following supervised machine-learning algorithms independently: the support vector machine (SVM; Cortes & Vapnik 1995), the k nearest neighbors (k-NN; Dasarathy 1991), and the extremely randomized trees (ERT; Geurts *et al.* 2006).

For each active region, we extracted ~60 features, including the sunspot area, the total magnetic flux, the length and number of magnetic neutral lines, the current helicity, the area of chromospheric brightening, the X-class and M-class flare histories, and their time evolutions (see Nishizuka *et al.* 2017 more in detail). We show some examples of two-dimensional distributions of four of the ~60 extracted solar features in Figure 2. The data for less than 24 hours before an X-class (M-class) flare are colored in red (green) and the non-flare data (including C-class flares) are colored in black. It was found that X-class flares tend to occur with a stronger magnetic field (>0.1 T), a larger sunspot area, a larger unsigned magnetic flux, and a larger number of magnetic neutral lines.

3. Evaluation Methods and Prediction Results

To evaluate the prediction results, we prepared the two datasets for training and testing in two ways (Figure 3). First, after standardizing the database of extracted features, we fully shuffled it and divided it into two for training and testing in the ratio of 7 to 3. Second, to investigate the performances of prediction models in a more operational setting, we shuffled the database with a one week unit. This is because while fully shuffling

https://doi.org/10.1017/S1743921317007293 Published online by Cambridge University Press

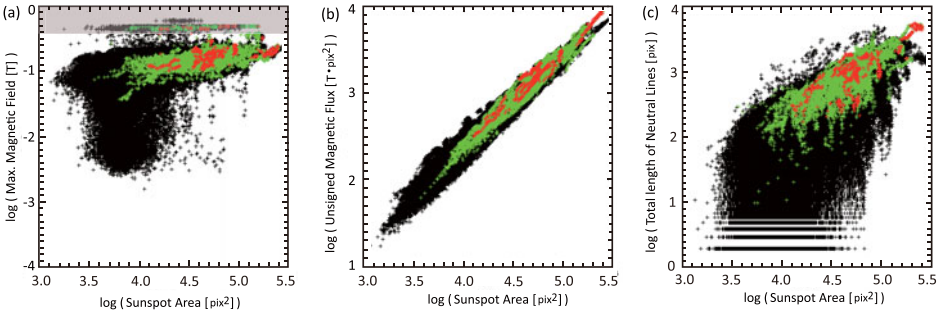

Figure 2. Examples of two-dimensional distributions of four of the ∼60 extracted solar features in our model. The data for less than 24 hours before an X-class/M-class flare are colored in red/green. The non-flare data (including C-class flares) are colored in black. Data in the shaded area, which are affected by the observation fitting error, were removed from the database for machine learning. Here, the length unit of one pixel of SDO images is 0.5" or equal to about 362 km on the Sun.

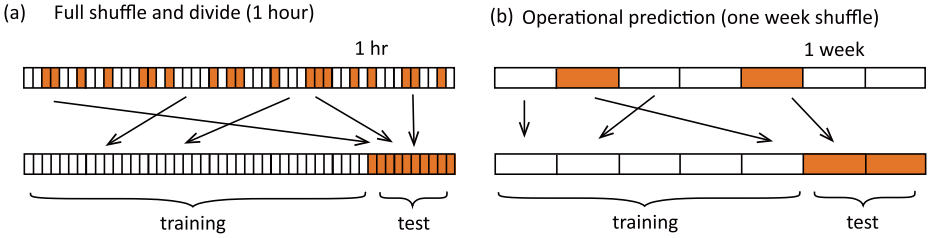

Figure 3. Two ways of producing the training and testing datasets: (a) full shuffle and divide and (b) one week shuffle and divide for investigation in a more operational setting.

and dividing the database, consecutive labeled data in a time sequence before a flare can be separated into both training and testing datasets, which results in an increase in the score.

The contingency tables of the prediction results for X-class flares are shown in Figure 4. We evaluated the prediction results using the skill score named the true skill statistic (TSS; Hanseen & Kuipers 1965). It is determined by $TSS = TP/(TP+FN) - FP/(FP+TN)$, where TP, FN, FP, and TN are the numbers of true positives, false negatives, false positives, and true negatives, respectively. Using the fully shuffled and divided database, all three methods achieved TSS greater than 0.8. It was found that, among the three algorithms, k-NN with k=1 exhibited the best performance with TSS greater than 0.9 (for reference, TSS=0.76 in Bobra & Couvidat 2015). In a more operational setting with a one week shuffle, k-NN still exhibited the highest score of greater than 0.9, but the scores of SVM and ERT decreased to 0.60 and 0.52, respectively.

4. Summary and Discussion

We developed a model of flare prediction employing machine learning using the multi-wavelength observation data of line-of-sight and vector magnetograms, the lower chromospheric brightening, and the soft-X-ray emission for the period 2010-2015. We applied three machine-learning algorithms for X-class flare prediction, i.e., SVM, k-NN, and ERT, and achieved TSS values of 0.93 for the fully shuffled datasets and 0.91 for the one-week-shuffled datasets in a more operational setting. k-NN exhibited the best score among the three methods, but in a more operational setting, the SVM and ERT scores decreased

https://doi.org/10.1017/S1743921317007293 Published online by Cambridge University Press

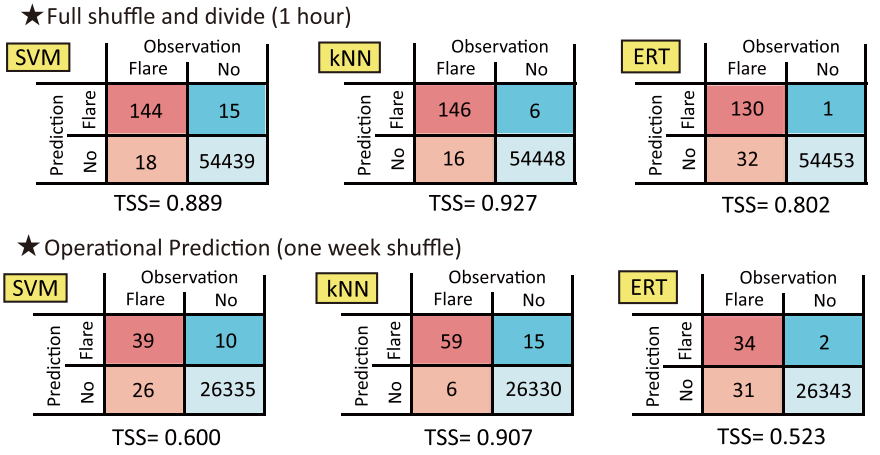

Figure 4. Prediction results of X-class flares occurring in the following 24 hours obtained by three machine-learning algorithms for the fully shuffled datasets and the one-week shuffled datasets in a more operational setting.

markedly. This indicates that the prediction scores strongly depend on the preparation method of the training and test datasets, as well as the evaluation method. Currently, there is no standard evaluation method for different prediction models and current operations (Crown 2012, Devos *et al.* 2014, Murray *et al.* 2017, Kubo *et al.* 2017). To ensure a fair comparison, we need to find a standard way of evaluating prediction models in a more operational setting.

References

Ahmed, O. W., Qahwaji, R., Colak, T., *et al.* 2013, *Sol. Phys.*, 283, 157
Barnes, G., Leka, K. D., Schrijver, C. J., *et al.* 2016, *Astrophys. J.*, 829, 89
Bloomfield, D. S., Higgins, P. A., McAteer, R. T. J. & Gallagher, P. 2012, *ApJ*(Letters), 747, L41
Bobra, M. G. & Couvidat, S. 2015, *Astrophys. J.*, 798, 135
Cortes, C. & Vapnik, V. 1995, *Mach. Learn.* 20, 273
Crown, M. D. 2012, *Space Weather*, 10, S06006
Dasarathy, B. V. 1991, *Nearest Neighbor (NN) Norms: NN Pattern Classification Techniques* (Los Alamitos, CA: IEEE Computer Society Press), 447
Devos, A., Verbeeck, C. & Robbrecht, E. 2014, *J. Space Weather Space Clim.*, 4, A29
Guerts, P., Ernst, D. & Wehenkel, L. 2006, *Mach. Learn.*, 63, 3
Hanssen, A. J. & Kuipers, W. J. 1965, *Meded Verhand*, 81, 2
Kubo, Y., Den, M. & Ishii, M. 2017, *J. Space Weather Space Clim.*, 7, A20
Leka, K. D. & Barnes, G. 2003, *Astrophys. J.*, 595, 1296
Lemen, J., Title, A. M., Akin, D. J., *et al.* 2012, *Sol. Phys.*, 275, 17
Liu, C., Deng, N., Wang, J. T. L. & Wang, H. 2017, *Astrophys. J.*, 843, 104
McCloskey, A. E., Gallagher, P. T. & Bloomfield, D. S. 2016, *Sol. Phys.*, 291, 1711
Murray, S. A., Bingham, S., Sharpe, M. & Jackson, D. R. 2017, *Space Weather*, 15, 577
Nishizuka, N., Sugiura, K., Kubo, Y., *et al.* 2017, *Astrophys. J.*, 835, 156
Pesnell, W., Thompson, B. J., Chamberlin, P. C., *et al.* 2012, *Sol. Phys.*, 275, 3
Raboonik, A., Safari, H., Alipour, N. & Wheatland, M. S. 2017, *Astrophys. J.*, 834, 11
Scherrer, P. H., Schou, J., Bush, R. I., *et al.* 2012, *Sol. Phys.*, 275, 207

https://doi.org/10.1017/S1743921317007293 Published online by Cambridge University Press