

# Solar Flare Prediction using the Machine-learning and Operational Evaluation Method

\*Naoto Nishizuka<sup>1</sup>, Komei Sugiura<sup>1</sup>, Yuki Kubo<sup>1</sup>, Mamoru Ishii<sup>1</sup>

1. National Institute of Information and Communications Technology

We have developed a flare prediction model using solar observation data and machine-learning techniques. From the full-sun images, we extracted features such as photospheric magnetic field, chromospheric brightening and X-ray activities in each active region, and then we predicted the maximum class of flares occurring in the following 24 hours (Nishizuka et al., 2016, oral presentation in JpGU). However, a standard evaluation method of flare prediction models has not been established. Moreover, under the sever condition that the test dataset is completely independent from the training dataset in an operational setting, we could not predict solar flares with high accuracy.

In this presentation, we introduce a method of the time-series cross-validation (CV) to evaluate flare prediction models in an operational setting, though the k-fold (10-fold) CV has ever been used in the previous studies. In some sense, these two methods are reasonable and available. However, when we focus on the operational usage, the time-series CV is superior to the k-fold CV. Furthermore, we used a machine-learning algorithm called the Gradient Boosted Trees for the first time. The boosting is a method to minimize the loss function by sequentially adding weak classifiers, or decision trees in our model. This is used to achieve a better prediction, by repeating learning of the calculation of the gradient when optimizing parameters in each step. We applied this algorithm to the flare prediction and performed the time-series CV. As a result, we succeeded in improving our prediction score, a skill score called the true skill statistic, from 0.2 to 0.6 for X-class flares and to 0.8 for M-class flares. We also compared the performance of other five different machine-learning algorithms to predict flares, and we found that the ranking of the performance of the algorithms completely differs according to the CV method.

Keywords: Solar Flare , Prediction, Modeling , Machine-learning, Evaluation Method