

Article

Not peer-reviewed version

---

# Solar Flare Prediction from Extremely Imbalanced Multivariate Time Series Data using Minimally Random Convolutional Kernel Transform

---

[Kartik Saini](#)<sup>\*</sup>, [Khaznah Alshammari](#)<sup>\*</sup>, [Shah Muhammad Hamdi](#)<sup>\*</sup>, [Soukaina Filali Boubrahimi](#)<sup>\*</sup>

Posted Date: 5 March 2024

doi: 10.20944/preprints202403.0210.v1

Keywords: Multivariate Time Series; Solar Flare; Space Weather; Imbalanced Data



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

# Solar Flare Prediction from Extremely Imbalanced Multivariate Time Series Data Using Minimally Random Convolutional Kernel Transform

Kartik Saini <sup>1,†</sup> , Khaznah Alshammari <sup>2,\*</sup> , Shah Muhammad Hamdi <sup>1,\*</sup>   
and Soukaina Filali Boubrahimi <sup>1,\*</sup> 

<sup>1</sup> Utah State University, Logan, UT 84322, USA; s.hamdi@usu.edu

<sup>2</sup> Northern Border University, Rafha, Saudi Arabia; kalshamm@nmsu.edu

\* Correspondence: kalshamm@nmsu.edu (K.A.); s.hamdi@usu.edu (M.H.)

† Current address: Utah State University, Logan, UT 84322, USA.

‡ These authors contributed equally to this work.

**Abstract:** Solar flares are characterized by sudden bursts of electromagnetic radiation from the Sun's surface, and caused by the changes in magnetic field states in solar active regions. Earth and its surrounding space environment can suffer from various negative impacts caused by solar flares ranging from electronic communication disruption to radiation exposure-based health risks to the astronauts. In this paper, we address the solar flare prediction problem from magnetic field parameter-based multivariate time series (MVTs) data using multiple state-of-the-art machine learning classifiers that include MINImally RandOm Convolutional Kernel Transform (MINIROCKET), Support Vector Machine (SVM), Canonical Interval Forest (CIF), Multiple Representations SEquence Learner (Mr-SEQL), and Long Short-Term Memory (LSTM)-based deep learning model. Our experiment is conducted on the on the Space Weather ANalytics for Solar Flares (SWAN-SF) benchmark data set, which is a partitioned collection of MVTs data of active region magnetic field parameters spanning over 9 years of operation of the Solar Dynamics Observatory (SDO). The MVTs instances of the SWAN-SF dataset are labeled by GOES X-ray flux-based flare class labels, and attributed to extreme class imbalance because of the rarity of the major flaring events (e.g., X and M). As a performance validation metric in this class-imbalanced dataset, we used the true skill statistic (TSS) score. Finally, we demonstrate the advantages of the MVTs learning algorithm MINIROCKET, which outperformed the aforementioned classifiers without the need for essential data preprocessing steps such as normalization, statistical summarization, and class imbalance handling heuristics.

**Keywords:** multivariate time series; solar flare; space weather; imbalanced data

## 1. Introduction

Solar flares are strong outbursts of radiation that result from the sudden release of magnetic energy that has been stored in the Sun. The length of time that a solar flare lasts can range anywhere from a few minutes to many hours. Since 1974, the Geostationary Operational Environmental Satellite (GOES) that are operated by the National Oceanic and Atmospheric Administration (NOAA) have been able to detect and categorize the X-ray flux that is produced by flare events in the 1-8 wavelength range. Based on their peak soft X-ray emission in this range, flares are grouped logarithmically as A, B, C, M, and X, ascending from less powerful to more powerful, starting  $10^{-8} \text{ W m}^{-2}$  [1]. As a direct consequence of this, the peak X-ray flux of an X-class flare is typically one hundred times stronger than that of a C-class flare and ten times stronger than that of an M-class flare. Each class can be broken down into one of nine sub-classes. When the X-ray level is high, it is sometimes difficult or even impossible to detect flares of the A and B classes. However, flares of the C class and higher are identified the vast majority of the time, particularly above level C2. Because of the potential damage they could cause, flares of the M and X classes, which are the most severe flares, are typically the focus of space weather forecasting.

Flares of the X-class and the M-class have the potential to cause radio blackouts across the Earth and initiate persistent radiation storms in the upper atmosphere. Astronauts, flight attendants, and passengers could be exposed to significant risks. As highlighted in the study by [2], the destruction resulting from a solar flare might incur repair and replacement costs exceeding one trillion dollars. However, by implementing appropriate safety measures and deploying a reliable system for predicting solar flares, it is possible to substantially reduce the extent of the damage.

The rarity of the major flaring events, however, is the most challenging in addressing solar flare classification. As stated by NASA, the solar cycle, which lasts around 11 years, influences the frequency with which solar flares occur. During times of solar maximum, it may occur multiple times each day, and during times of solar minimum, it may occur less than once per week. In addition, stronger flares occur less frequently than their less intense counterparts. For instance, flares of the X10 class, which are considered severe, occur on average approximately eight times every cycle, making it a very rare phenomenon. Flares of the M1 class, which are considered small, occur on average around two thousand times per cycle. Since there is a significant imbalance in the class distribution, traditional classifiers struggle to predict the minority class with high accuracy. The majority of classification-based machine learning algorithms were built on the assumption of an equal ratio of samples for each class ([3]).

The fact that the dataset comprises a range of time series parameters obtained from solar photospheric magnetograms makes flare forecasting even more challenging, in addition to NOAA's record of flares in active regions (Refer to table 2). It also includes physics-based magnetic field parameters, originally acquired through the Space Weather HMI Active Region Patches (SHARP) data product ([4]). The high dimensionality of the time series poses another challenge because of the curse of dimensionality and the fact that many feature vectors can be noisy. Classifiers that achieve credible accuracy on imbalanced time series data are quite computationally expensive. Even with small data sets, they can take a substantial amount of time to train.

The field of astrophysics lacks a specific physical theory that comprehensively explains the mechanism behind the occurrence of solar flares, which limits the ability to forecast and classify them [5]. While various groups of physicists are actively researching to unveil a definitive theory for flare prediction, the likelihood of success remains uncertain. Given the rapid advancements in AI and machine learning, the most promising approach is to adopt a data-driven strategy using the active region parameters observed by the Solar Dynamics Observatory. The goal is to develop a model that can establish an empirical relationship between AR parameters and flare occurrences.

[1] comprehensively presented the challenges involved with handling the SWAN-SF Dataset, the largest data up-to-date on solar flares based on MVTs-based photospheric magnetic field parameters of solar active regions. They discussed the extreme class imbalance in the data as well as the temporal coherence and proposed different remedies to tackle the problems. They started with extracting the statistical features of each magnetic field parameter time series, such as median, standard deviation, skewness, and kurtosis followed by the last value of each time series, which also reduced the dimensionality of the data set and made it scalable. They used a SVM classifier to test their flare prediction performance. For class imbalance, they leveraged undersampling and oversampling techniques as mandatory preprocessing steps. At the classifier level, they tuned the misclassification weighting parameter to keep false positives and false negatives at minimum. For temporal coherence, they used 20 pairs of testing and training data from different partitions to avoid the overlapping of the sampled MVTs sequence. As forecast metrics, they used True Skill Statistics (TSS) and realizations of the Heidke Skill Score (HSS; HSS2) to report how robustly the SVM model performed. However, the experimental settings utilized by [1] had some limitations. Firstly, they calculated only five statistical features; these features may not accurately represent the complete properties of the time series data, resulting in inaccurate predictions. Furthermore, they performed an important data preprocessing step which consists of undersampling or oversampling of training data. While the previous methods

relied on preprocessing through normalization and balancing, the proposed algorithm MINIROCKET can achieve higher performance without the preprocessing steps.

The goal of this study is as following:

- This study aims to assess the effectiveness of the MINIROCKET ([6]) time series classifier, based on the MINImally RandOm Convolutional Kernel Transform, for real-time prediction of solar flares with minimal data manipulation. MINIROCKET, an efficient variant of the ROCKET classifier ([7]), achieves high precision at reduced computing costs by employing random convolutional kernels to transform input time series. The transformed features are then used to train a linear classifier. MINIROCKET, being a (nearly) deterministic reformulation of ROCKET, exhibits significantly faster performance on larger datasets while maintaining comparable accuracy.
- In this study, we compare MINIROCKET's performance with Canonical Interval Forest (CIF) ([8]), Multiple Representations SEquence Learner (Mr-SEQL), support vector machine (SVM) and Long Short-Term Memory (LSTM) models.
- The evaluation metrics, TSS and HSS2 ([9]), are selected for comparison because they are the most commonly used metrics for flare prediction with class imbalance data. To address data overlapping, we implement the 20-partition pair strategy proposed by [1].

## 2. Related Work

Theo, [10], was among the first systems to predict flares. It was an expert system that required human input. It used a set of sunspots and magnetic field parameters to forecast different flare classifications. Rule-based flare prediction using Theo was adopted by National Oceanic and Atmospheric Administration's (NOAA) Space Environment Center (SEC) in 1987. The current methods of flare prediction are data-driven and are divided into two categories: linear statistical and nonlinear statistical. They can be further divided into line-of-sight magnetogram-based models and vector magnetogram-based models. The continuous stream of vector magnetograms is always considered a better means for parameterizing the active regions as they contain the full-disk magnetic field data as mentioned in [11]. However, it was not easily available before the launch of the Solar Dynamics Observatory (SDO) by the National Aeronautics and Space Administration (NASA) in 2010, and solar physicists had to depend on line-of-sight magnetic data for flare prediction.

Linear statistical studies aim to identify the AR magnetic properties that are correlated with flares. [12] used line-of-sight magnetograms to parameterize active regions and studied the correlation between AR parameters and flare occurrences. From many SOHO/MDI longitudinal magnetograms, they evaluated three physical measures: the maximum horizontal gradient, the length of the neutral line, and the number of singular points. Properties of the photospheric magnetic field, such as non-potentiality and complexity, thought to be highly related to solar flares, have been identified using these evaluated measures. Their statistical analysis concluded that solar flare productivity increases with non-potentiality and complexity. In a similar study, using line-of-sight Michelson Doppler Imager (MDI) magnetograms of 89 active regions and Solar Geophysical Data (SGD) flare reports, [13] assessed the magnitude-scaling correlations between three parameters of magnetic fields and the flare productivity of solar active regions. The mean value of spatial magnetic gradients at strong-gradient magnetic neutral lines (NL), the length of strong-gradient magnetic neutral lines (LGNL), and the total magnetic energy were the parameters studied. Active region MDI magnetograms used in their research were found to be relatively close to the solar central meridian. In particular, they revealed strong positive linkages between the parameters and both the total flare productivity of active regions and the potential of following flare production. Their findings confirmed the dependence of flare productivity on the degree of non-potentiality of active regions. [14] was the first to determine AR parameters from vector magnetograms. They conducted statistical tests based on discriminant analysis on a wide variety of photospheric magnetic parameters in order to identify those properties that are critical for the production of energetic events such as solar flares. They concluded that while the

factors evaluated singly had minimal power to differentiate between flaring and flare-quiet groups, the populations could be separated using multi-variable combinations.

Nonlinear statistical models are commonly implemented using traditional machine learning classifiers. In the context of classification models, several approaches have been explored. [15] utilized a C4.5 decision tree, while [16] employed logistic regression. [17] opted for an artificial neural network, and [18] utilized a relevance vector machine. Additionally, [19] investigated the performance of three classifiers—k-NN, SVM, and Extremely Randomized Tree—utilizing both line-of-sight and vector magnetograms.

The pioneering work by [20] marked the first instance of employing machine learning algorithms on HMI vector magnetograms. They employed a Support Vector Machine (SVM) classifier, leveraging four years of data from the Solar Dynamics Observatory's (SDO) Helioseismic and Magnetic Imager (HMI) to forecast M- and X-class solar flares. Their work was groundbreaking as it involved a vast dataset of vector magnetograms for flare prediction. The authors created a catalog of flaring and non-flaring active regions from a database containing 2071 active regions, comprising 1.5 million active region patches of vector magnetic field data. Using 25 parameters, each active region was classified. Additionally, they employed a feature selection algorithm to identify the most effective features for distinguishing between flaring and non-flaring active zones. To address the class imbalance problem, they utilized a cost function to limit false negatives.

Efficiently addressing the solar flare prediction task, [21] framed solar flare classification as a binary classification problem, distinguishing between flaring and non-flaring active regions. They meticulously extracted time series samples of the active region parameters and developed a flare prediction method utilizing k-NN classification on univariate time series. Interestingly, their research revealed that employing a statistical summarization technique on a specific active Region parameter known as "total unsigned current helicity" outperformed using all active Region parameters at a single point in time. Furthermore, by exploring the time series properties of the AR parameters, the researchers identified the most influential parameter, thereby simplifying the problem to a single-variate time series classification. They proposed a novel approach of using a statistical summarization method on the time series, allowing the top AR parameter to serve as the vector-based representation of flaring/non-flaring active regions. By applying the k-nearest neighbors (k-NN) classifier within this reduced vector space, they achieved significant computational and time savings. Importantly, they also demonstrated that including C-class flares in the positive class did not improve classification performance.

Angryk and colleagues[22] presented a comprehensive multivariate time series (MVTs) dataset derived from solar photospheric vector magnetograms in the Spaceweather HMI Active Region Patch (SHARP) series. The dataset encompassed 4,098 MVTs data instances collected from active regions between May 2010 and December 2018. It included 51 flare-predictive parameters and over 10,000 flare reports. The dataset served as a valuable testbed for solar physicists and machine learning practitioners, providing a cleansed, integrated, and readily available dataset with data verified from multiple sources. The study incorporated data from the GOES flare catalog, SSW and XRT flares, and NOAA AR locations to enhance, verify, and cleanse the dataset. The authors recalculated magnetic field parameters from individual region patches and transformed them into multivariate time series spanning the entire length of a given HARP series. They further addressed dataset cleaning, accounting for empty SHARPs, location-based filtering, and missing values. Subsequently, the dataset was partitioned into target classes using flare intensity threshold criteria, while observation window, latency, and prediction window concepts were utilized for custom slicing and labeling.

Ahmadzadeh and colleagues[1] discussed the challenges posed by the SWAN-SF dataset introduced by [22]. They highlighted the extreme class imbalance within the data, as well as the temporal coherence. To tackle these issues, the researchers began by extracting statistical features from the time series, thereby reducing the dimensionality of the dataset. They employed an SVM machine learning classifier to conduct their experiments and addressed class imbalance through a combina-

tion of undersampling and oversampling techniques at the data level. Additionally, they fine-tuned the misclassification weighting parameter at the classifier level to minimize false positives and false negatives. To account for temporal coherence, they employed testing and training data from different partitions to prevent overlapping of data points.

### 3. Dataset

The benchmark dataset Space Weather ANalytics for Solar Flares (SWAN-SF) by [22] serves as an illustrative example of a multivariate time series, aiming to achieve unbiased flare forecasting and classification. This dataset encompasses five distinct flare classes, ranging from the most powerful X-class and M-class flares to the smaller B-class and C-class flares. Additionally, it includes a non-flaring class denoted as the F class. In this paper, we refer to the flaring (M and X class flares) and non-flaring (F, B, and C class flares) as the positive and negative classes, respectively. To ensure temporal segmentation, the dataset has been divided into five partitions, each containing approximately equal proportions of X- and M-class flares (Table 1).

**Table 1.** Event type statistics of each partition of the SWAN-SF dataset.

Flare Type	Partitions				
	P1	P2	P3	P4	P5
Q	60,130	73,368	34,762	43,294	62,688
B	5,692	4,978	685	846	5,924
C	6,416	8,810	5,639	5,956	5,763
M	1,089	1,329	1,288	1,012	971
X	165	72	136	153	19
<b>sum</b>	<b>73,492</b>	<b>88,557</b>	<b>42,510</b>	<b>51,261</b>	<b>75,365</b>

The dataset comprises time series features derived from solar photospheric magnetograms, alongside NOAA's active region flare history. The Solar Dynamics Observatory's ([23]) HMI Active Region Patches (HARP) data product provides magnetograms ([24]). While the magnetic field parameters are initially derived from the Space weather HMI Active Region Patches (SHARP) data product ([4]), they were recalculated and augmented with additional parameters for validation purposes, including parameters not found in SHARPs (refer to Table 1 in [22]). The dataset consists of sliding time series slices, with each instance representing 24 physical magnetic field parameters (see Table 2). These time series instances are logged at 12-minute intervals over a total of 12 hours (60-time steps).

Each solar active region exhibits a different range of flare classes or maintains a region of tranquil activity within a prediction window. These changes are reflected in the representation of solar event  $i$  as  $mvts_i$ , a multivariate time series instance, along with its corresponding class label,  $y_i$ . The term  $y_i$  characterizes the various flare classes. Comprising  $N$  magnetic field parameters, the multivariate time series instance  $mvts_i \in R^{T \times N}$  encompasses multiple time series with periodic observations over an interval of  $T$ . The  $t$ -th timestamp value is denoted as  $x^{<t>} \in R^N$ , while the  $j$ -th parameter time series is denoted as  $P_j \in R^T$ . The event is classified based on the active region's state after the observation time  $T$  and the subsequent prediction interval  $L$ . NOAA records of flare events are utilized to determine the state of a given timestamp.

When the population of one or more data classes is significantly smaller than the majority classes, the dataset is considered as class imbalanced. The minority classes consist of data points from the smaller group, while the data points from the other group are referred to as the majority classes. Table 1 illustrates the substantial class imbalance present in the SWAN-SF benchmark dataset. Traditional machine learning classifiers tend to favor the majority class, as highlighted by [29]. This becomes especially concerning in solar flare classification, where the focus lies on a minority of cases. Class imbalance significantly impacts various performance metrics, including accuracy, precision, and the F1 score. This is primarily due to the metrics disregarding the number of misclassifications. For instance,

a traditional model that assigns all instances to the majority class may achieve high accuracy while learning very little about the minority class. In the following sections, we discuss TSS and HSS2 evaluation metrics, which are commonly employed in such class imbalance scenarios to measure the model performance.

**Table 2.** List of AR magnetic field parameters.

Abbreviation	Description	Formula
ABSJZH [14]	Absolute value of the net current helicity	$H_{c,abs} \propto  \sum B_z \cdot J_z $
EPSX [25]	Sum of x-component of normalized Lorentz force	$\delta F_x \propto \frac{\sum B_x B_z}{\sum B^2}$
EPSY [25]	Sum of y-component of normalized Lorentz force	$\delta F_y \propto \frac{-\sum B_y B_z}{\sum B^2}$
EPSZ [25]	Sum of z-component of normalized Lorentz force	$\delta F_z \propto \frac{\sum (B_x^2 + B_y^2 - B_z^2)}{\sum B^2}$
MEANALP [26]	Mean characteristic twist parameter, $\alpha$	$\alpha_{total} \propto \frac{\sum J_z \cdot B_z}{\sum B^2}$
MEANGAM [14]	Mean angle of field from radial	$\bar{\gamma} = \frac{1}{N} \sum \arctan\left(\frac{B_h}{B_z}\right)$
MEANGBH [14]	Mean gradient of horizontal field	$ \nabla B_h  = \frac{1}{N} \sum \sqrt{\left(\frac{\partial B_h}{\partial x}\right)^2 + \left(\frac{\partial B_h}{\partial y}\right)^2}$
MEANGBT [14]	Mean gradient of total field	$ \nabla B_{tot}  = \frac{1}{N} \sum \sqrt{\left(\frac{\partial B}{\partial x}\right)^2 + \left(\frac{\partial B}{\partial y}\right)^2}$
MEANGBZ [14]	Mean gradient of vertical field	$ \nabla B_z  = \frac{1}{N} \sum \sqrt{\left(\frac{\partial B_z}{\partial x}\right)^2 + \left(\frac{\partial B_z}{\partial y}\right)^2}$
MEANJZD [14]	Mean vertical current density	$\bar{J}_z \propto \frac{1}{N} \sum \left(\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y}\right)$
MEANJZH [14]	Mean current helicity ( $B_z$ contribution)	$\bar{H}_c \propto \frac{1}{N} \sum B_z \cdot J_z$
MEANPOT [27]	Mean photospheric magnetic free energy	$\bar{\rho} \propto \frac{1}{N} \sum (B^{Obs} - B^{Pot})^2$
MEANSHR [27]	Mean shear angle	$\bar{\Gamma} = \frac{1}{N} \sum \arccos\left(\frac{B^{Obs} \cdot B^{Pot}}{ B^{Obs}   B^{Pot} }\right)$
R_VALUE [28]	Sum of flux near polarity inversion line	$\Phi = \sum  B_{LoS}  dA$ (within R mask)
SAVNCPP [14]	Sum of the modulus of the net current per polarity	$J_{z, sum} \propto  \sum B_z^+ J_z dA  +  \sum B_z^- J_z dA $
SHRGT45 [14]	Fraction of Area with shear > 45°	Area with shear > 45° / total area
TOTBSQ [25]	Total magnitude of Lorentz force	$F \propto \sum B^2$
TOTFX [25]	Sum of x-component of Lorentz force	$F_x \propto -\sum B_x B_z dA$
TOTFY [25]	Sum of y-component of Lorentz force	$F_y \propto \sum B_y B_z dA$
TOTFZ [25]	Sum of z-component of Lorentz force	$F_z \propto \sum (B_x^2 + B_y^2 - B_z^2) dA$
TOTPOT [14]	Total photospheric magnetic free energy density	$\rho_{tot} \propto \sum (B^{Obs} - B^{Pot})^2 dA$
TOTUSJH [14]	Total unsigned current helicity	$H_{c, total} \propto \sum B_z \cdot J_z$
TOTUSJZ [14]	Total unsigned vertical current	$J_{z, total} = \sum  J_z  dA$
USFLUX [14]	Total unsigned flux	$\Phi = \sum  B_z  dA$

#### 4. Methodology

While machine learning and deep learning-based classifiers for time series classification have achieved impressive levels of accuracy, they often suffer from high computational complexity. This drawback becomes particularly problematic for larger datasets, resulting in longer training times and rendering them impractical. Moreover, many existing techniques focus on specific aspects of the data, such as shape or frequency, neglecting the broader picture. To address these challenges, [7] introduced the RandOm Convolutional KERNels Transform (ROCKET) method. This novel approach leverages the success of convolutional neural networks in time series classification by utilizing random convolutional kernels to extract informative features. These features are then used to train a linear classifier. To further enhance efficiency, [6] proposed a modified version called the MINImally RandOm Convolutional KERNels Transform (MINIROCKET), which achieves faster execution and nearly deterministic behavior.

The ROCKET method transforms time series data by convolving each series with a set of random convolutional kernels. These kernels, similar to those in convolutional neural networks, possess random characteristics such as length, weights, bias, dilation, and padding. They capture a wide range of information and patterns at various frequencies and scales. The output of each kernel undergoes two pooling techniques: global max pooling and percentage of positive values (PPV) pooling. Global

max pooling selects the maximum feature value, while PPV pooling evaluates the prevalence of a pattern captured by the kernel.  $ppv = 1/n \sum_{i=0}^{n-1} [z_i > 0]$ , where  $z_i$  is the output of the convolution operation. The fraction of positive values derived from PPV pooling plays a vital role in assessing the significance of the captured patterns, contributing to the method's high precision. Each kernel produces two features, resulting in a total of 20,000 features per input time series when employing 10,000 random convolutional kernels. These extracted features are then utilized for training a linear classifier.

Both the MINIROCKET and ROCKET methods rely on PPV pooling to assess the convolution values. MINIROCKET further optimizes computational efficiency by employing a fixed set of kernels with specific hyper-parameter settings, refer to Table 3. Key modifications include fixing the kernel length at 9, limiting the weight hyper-parameter to a fixed range, adapting the bias hyper-parameter to random convolutional output values, restricting the dilation hyper-parameter, and employing only PPV pooling instead of global max pooling and PPV. These optimizations enable MINIROCKET to generate half as many features as ROCKET while maintaining equivalent precision. MINIROCKET achieves remarkable computational efficiency through a combination of the aforementioned optimizations. It utilizes the mathematical properties of fixed kernels and PPV pooling to compute PPV for both positive and negative weights simultaneously, effectively doubling the number of applied kernels without increasing computations. It also maximizes the reuse of convolution output and avoids multiplications by employing additive operations. Additionally, MINIROCKET computes all kernels for each dilation at once, further optimizing computation and output reuse. These optimizations significantly improve computational efficiency while preserving the accuracy level of the ROCKET classifier.

**Table 3.** Difference between ROCKET and MINIROCKET Kernel's hyper-parameters.

Hyper-parameters	ROCKET	MINIROCKET
Length	{7, 8, 11}	9
Weight	N(0,1)	(-1,2)
Bias	U(-1,1)	From convolution output
Dilation	Random	Fixed
Padding	Random	Fixed

In our experiments with the SWAN-SF dataset, MINIROCKET demonstrated superior performance compared to other classifiers. Its computational efficiency and accuracy make it an excellent choice for time series classification tasks. In the following section, we will discuss the outcomes of our experimentation and provide a concise overview of the remaining classifiers evaluated.

## 5. Experiments

In this section, we present an overview of the various models that we investigated and compare them to the state-of-the-art MINIROCKET. The study was conducted by evaluating the performance of each model under different data configurations and comparing the results. To ensure the validity of the results, we adopted a 5-fold cross-validation approach. In this approach, one partition was used for training, and the remaining four partitions were utilized for testing. For instance, we employed partition 1 to train the model, while partitions 2, 3, 4, and 5 were used individually for testing. This strategic approach yielded a total of 20 distinct partition pairs. This setting was chosen to align with the methodology used by [1] to prevent data overlap and remedy temporal coherence. The performance of the models was evaluated using the True Skill Statistic (TSS) score and Heidke Skill Score (HSS2), the two most frequently used metrics for flare prediction in class imbalance data.

### 5.1. Performance Metrics: TSS Score and HSS2 Score

An effective method of evaluating the efficacy of a classifier is by comparing its performance against a known standard classifier, which is called a benchmark. To make this comparison, a skill score is calculated. The skill score is calculated by taking the difference between the score value of the classifier's predictions and the score value of the standard forecast. This difference is then divided by the difference between a perfect score and the standard forecast. This entire calculation helps us understand how well the classifier is performing compared to the best possible scenario (perfect score) and the standard prediction. In the case of solar flare prediction, the development of such a skill score is logical, considering that the number of non-flaring regions is considerably greater than that of flaring ones. To assess the performance of various classifiers on the SWAN-SF dataset for flare prediction, we relied on the use of forecast verification metrics, with a focus on the True Skill Statistic (TSS) and Heidke Skill Score (HSS2) [30].

TSS and HSS2 are calculated based on the model's confusion matrix, which depicts the frequencies of predicted and actual values. The term "TN" represents True Negatives, reflecting the accurate classification of negative examples. Correspondingly, "TP" signifies True Positives, indicating the correct classification of positive examples. On the other hand, "FP" denotes False Positives, representing the number of actual negative examples mistakenly classified as positive. Finally, "FN" represents False Negatives, indicating the number of actual positive examples erroneously classified as negative. An example of a confusion matrix for binary classification is shown in Table 4.

**Table 4.** Confusion matrix for binary classification.

	Actual Positive	Actual Negative
Predicted Positive	True Positive	False Positive
Predicted Negative	False Negative	True Negative

[31] utilize a definition of the Heidke Skill Score (HSS) proposed by the Space Weather Prediction Center, denoted as HSS2, which quantifies the enhancement of the forecast compared to a random forecast. HSS2 is calculated using the following formula:

$$HSS_2 = \frac{TP + TN - E}{P + N - E} \quad (1)$$

Where E represents the anticipated number of correct predictions that can be corresponded to chance:

$$E = \frac{(TP + FP) * (TP + FN) + (FP + TN) * (FN + TN)}{P + N} \quad (2)$$

Alternatively, HSS2 can be derived from the true positive (TP), true negative (TN), false negative (FN), and false positive (FP) classification outcomes, in addition to the total number of positive (P) and negative (N) instances:

$$HSS_2 = \frac{2 * [(TP * TN) - (FN * FP)]}{P * (FN + TN) + (TP + FP) * N} \quad (3)$$

While HSS2 may be influenced by the class-imbalance ratio of the testing set, TSS has been recommended by [9] as a more suitable metric in these cases, as it is known to be unbiased with respect to the class-imbalance ratio and is considered to be more equitable. The TSS is defined as follows.

$$TSS = \frac{TP * TN - FP * FN}{P * N} = \frac{TP}{TP + FN} - \frac{FP}{FP + TN} \quad (4)$$

also known as the Hansen-Kuipers skill score or Peirce skill score ([32], measures the difference between the recall and false alarm rate, and ranges from -1 to 1, with a score of 1 indicating a perfect forecast, a score of 0 representing a random or constant forecast, and a score of -1 indicating a forecast that is always incorrect. The TSS is considered a desirable metric for comparing the performance of various classifiers for solar flare forecasting, as it takes into account both false negatives and false positives in a balanced manner, and is not affected by the imbalance of the testing set.

A potential limitation of the True Skill Statistic (TSS) is that it considers false positive (FP) and false negative (FN) predictions as having equal weight, despite the fact that the consequences of these misclassifications can vary. In the context of forecasting solar flares, the cost of a false negative (not predicting a flare that occurs) can be higher than the cost of a false positive (predicting a flare that does not occur), such as in the scenario of a satellite that needs to be rotated to protect against an increase in energetic particles. The costs of false positives and false negatives are not symmetrical. The TSS is insensitive to the imbalance ratio of the testing set, whereas the Heidke Skill Score (HSS2) can be influenced by this ratio and converge to zero as the ratio increases.

## 5.2. Comparing Different Classes of Classifiers

We evaluated the effectiveness of several time series classifiers, including LSTM, SVM, Mr-SEQ, and CIF, based on their TSS and HSS2 scores. By comparing the performance of these models, our analysis reveals that MINIROCKET achieves the highest TSS and outperforms the aforementioned classifiers in binary classification and all-class classification. This study highlights the potential of MINIROCKET as a powerful tool for flare classification. In the following sections, we provide a brief overview of each classifier and later compare the results obtained.

### 5.2.1. Long Short-Term Memory (LSTM)

We utilized Long Short-Term Memory (LSTM) networks in this research to learn representations of Multivariate Time Series (MVTs) instances without relying on hand-engineered statistical characteristics. The LSTM network was trained by sequentially feeding magnetic field parameter vectors into LSTM cells. Cell weights were optimized using gradient descent and backpropagation. The model effectively identified underlying patterns in the data and produced reliable predictions for flare occurrences through automated feature learning [33]. LSTM networks excel in processing and categorizing time-series data due to their ability to capture order dependence and long-term dependencies that regular RNNs cannot. Additionally, deep LSTM networks, created by stacking multiple LSTM layers, can learn even more complex patterns in sequential data. The usage of LSTM networks in this study showcases their usefulness in learning time series data representations and their potential for various domains.

### 5.2.2. Support Vector Machine (SVM)

The support vector machine (SVM) classifier aims to identify a hyperplane in N-dimensional space that can accurately classify input points. The search for an ideal hyperplane involves finding a plane with the greatest margin, which represents the maximum distance between data instances of different classes. This margin is crucial as it enables effective generalization and improves prediction accuracy. Hyperplanes act as decision boundaries, separating data points, and their size is determined by the number of features in the data. Support vectors, which are data instances closest to the hyperplane, greatly influence its orientation and placement. They play a critical role in optimizing the classifier's margin. By utilizing these support vectors, the SVM classifier finds the best hyperplane and achieves high prediction accuracy [34].

In the case of class imbalance in the flare dataset, the optimal hyperplane that intersects the decision boundary is pushed further toward the domain of the minority class. This adjustment aims to minimize the overall number of incorrect classifications, leading to an increase in true negatives (i.e., accurate classification of CBF-class flares) and a decrease in true positives (i.e., accurate classification of XM-class flares). In a class-imbalanced setting, models tend to exhibit a bias towards the majority class, which becomes concerning as the focus of flare-forecasting research lies on minority instances rather than the majority. The SVM classifier has gained popularity due to its ability to efficiently learn nonlinear decision surfaces, facilitated by support vectors and transformation functions (kernels). Various kernels can be utilized to enhance the transformation of data into new feature spaces, allowing

for a more accurate separation of instances. Kernels require the specification of one or more variables in advance, similar to any other function.

### 5.2.3. Canonical Interval Forest (CIF)

The time series forest (TSF) classifier, known for its high performance, quick training, and prediction, is commonly regarded as a powerful interval method. However, it has fallen behind in recent advancements in alternative techniques. TSF initially summarized intervals using only three basic summary statistics. In recent developments, the 'catch22' feature set ([35]) was engineered as a concise and useful collection of 22-time series features to facilitate extensive time series analysis. Building upon these advancements, the Canonical Interval Forest (CIF) classifier, proposed by [36], combines the capabilities of both TSF and catch22. The CIF classifier aims to enhance performance and accuracy in time series analysis by leveraging the unique strengths of both techniques.

### 5.2.4. Multiple Representations SEquence Learner (Mr-SEQL)

Mr-SEQL, proposed by [37], is a robust univariate time series classifier that trains on features derived from multiple symbolic representations of time series. These representations include Symbol Aggregation Approximation (SAX) and Symbol Fourier Approximation (SFA), which are used with linear classification models (logistic regression). Mr-SEQL utilizes SEQL ([38]) to extract features based on three key ideas. Firstly, instead of relying on a single fixed representation, Mr-SEQL combines multiple symbolic representations obtained from various parameters, such as multiple SAX representations. Secondly, it incorporates numerous domain representations in time (such as SAX) and frequency (such as SFA), making it resilient across a wide range of problems. Finally, Mr-SEQL extends a symbolic sequence classifier (SEQL) to effectively explore the significant symbolic-words space, employing an efficient greedy feature selection technique to find optimal features for each representation. Mr-SEQL is a highly effective time series classifier with important qualities that make it well-suited for a wide range of applications.

## 5.3. Binary Classification

In the preliminary experiments, we executed a transformation of the original data labels into binary labels with the aim of simplifying the classification process. The positive class, denoted as flaring, encompasses M and X class flares, while the negative class, referred to as non-flaring, encompasses F, B, and C class flares.

To evaluate the effectiveness of the classification models, we trained five different models, namely MINIROCKET, CLF, Mr-SEQL, LSTM, and SVM, and compared their performances in terms of TSS and HSS2 scores. We present the results of our experiments in the line plots displayed in Figures 1 and 2. These plots highlight the obtained scores for TSS and HSS2, respectively.

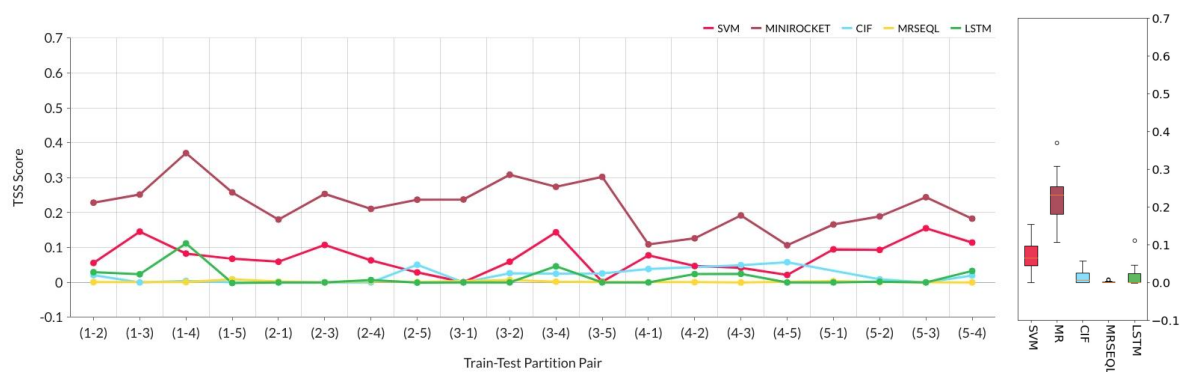
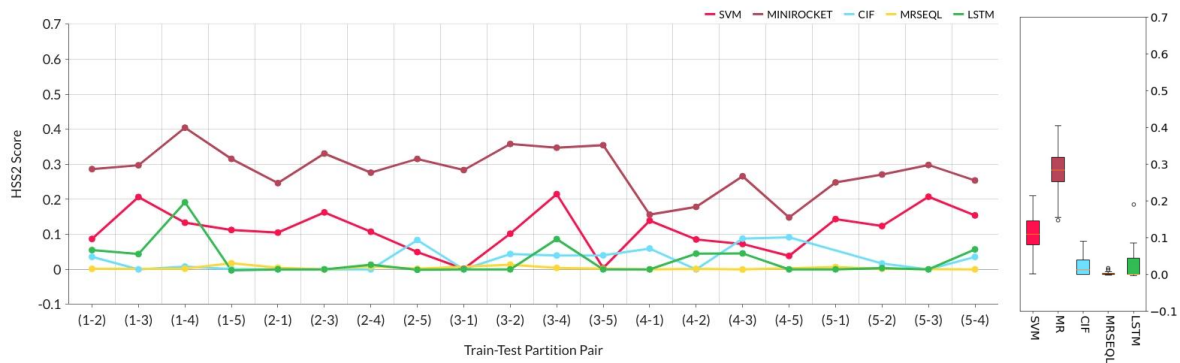


Figure 1. TSS score comparison of 5 different models.

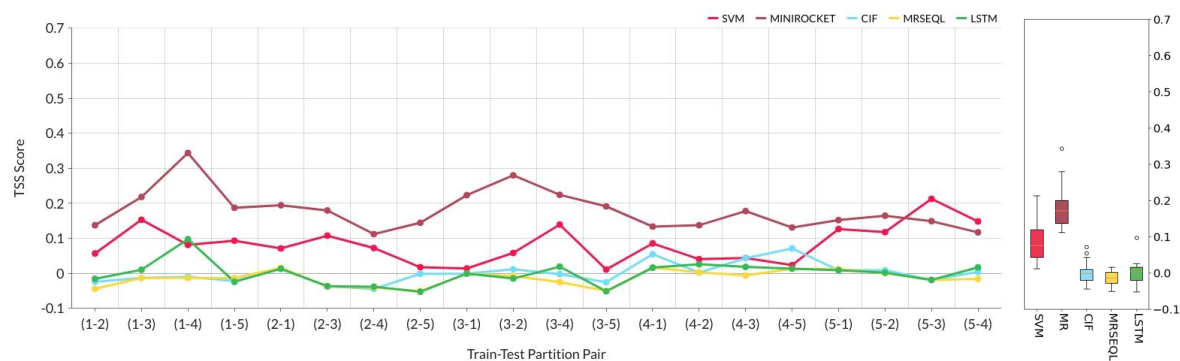


**Figure 2.** HSS2 score comparison of 5 different models

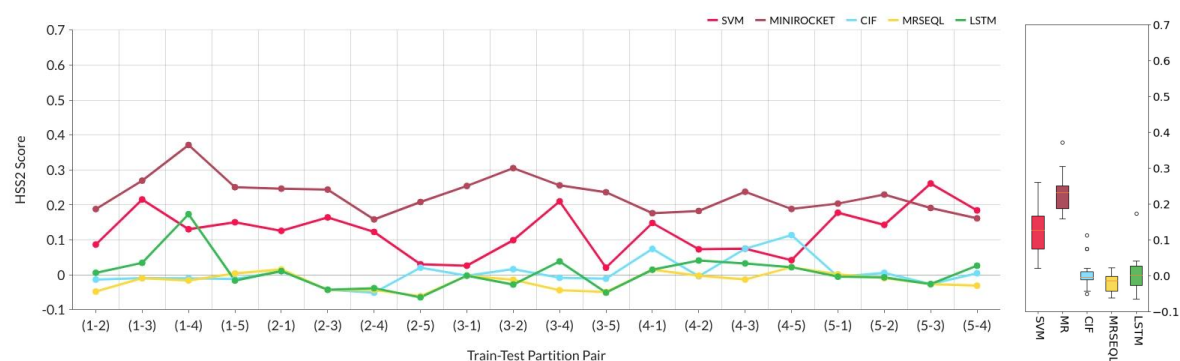
Our analysis has demonstrated that the MINIROCKET classifier outperformed aforementioned classifiers on the SWAN-SF dataset, with an average improvement of 19.4% and 23.9% in the TSS and HSS2 scores, respectively. Furthermore, the box plots illustrate the distribution of TSS and HSS2 score data from various classifiers across 20 distinct partition pairs. These plots offer insights into the variability and distribution of the data. A longer length in the box plots signifies increased variability in the data, as observed with the SVM and MINIROCKET classifiers. Notably, the MINIROCKET classifier exhibited the best performance, followed by SVM, LSTM, CIF, and Mr-SEQL models.

#### 5.4. Multi-class: All Class Classification

At this stage, we are engaged in the task of classifying the five distinct categories, namely F, B, C, M, and X. The experimental settings remain constant, wherein the training and testing are performed on 20 unique partition pairs, and the performance of the selected models, MINIROCKET, and SVM, are compared based on the TSS and HSS2 scores. Refer to the line plots displayed in Figures 3 and 4 for TSS and HSS2 score comparison.



**Figure 3.** TSS score comparison for all class classification.



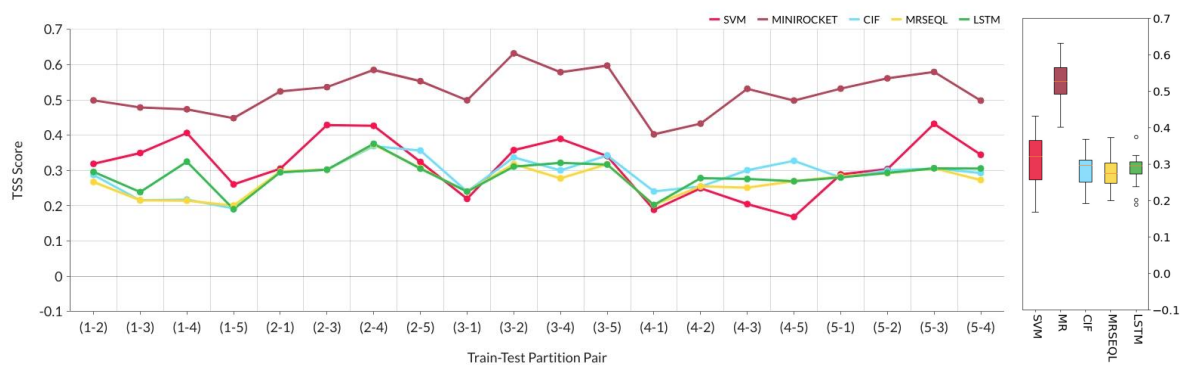
**Figure 4.** HSS2 score comparison for all class classification.

Our analysis of the all-class classification showed that the MINIROCKET classifier outperformed the aforementioned classifiers with a 9.61% higher TSS score and 10.36% higher HSS2 score. Upon analyzing the box plots depicting TSS and HSS2 scores for multi-class classification, it was evident that the MINIROCKET model demonstrated superior performance once more, with SVM, LSTM, CIF, and Mr-SEQL following suit in that order. The SVM model's box plot also exhibited the highest variability.

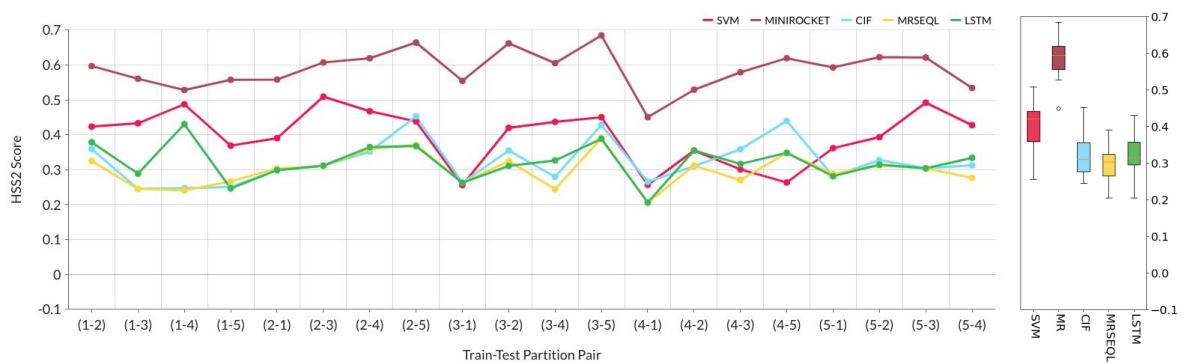
### 5.5. Analysis with the Exclusion of B and C Class Flares

In this phase of the experiment, B- and C-class flares would be excluded. This decision was made on the basis of the research conducted by [9], which suggested that the inclusion of C-class flares may have a negative impact on performance metrics. In our analysis, we observed an improvement in the TSS score for all models after the removal of B and C-class flares. This underscores the significance of this exclusion in achieving optimal model performance.

Following the exclusion of B and C-class flares, the experiment was further divided into two categories: binary class classification (Figures 5 and 6) and all-class classification (Figures 7 and 8).

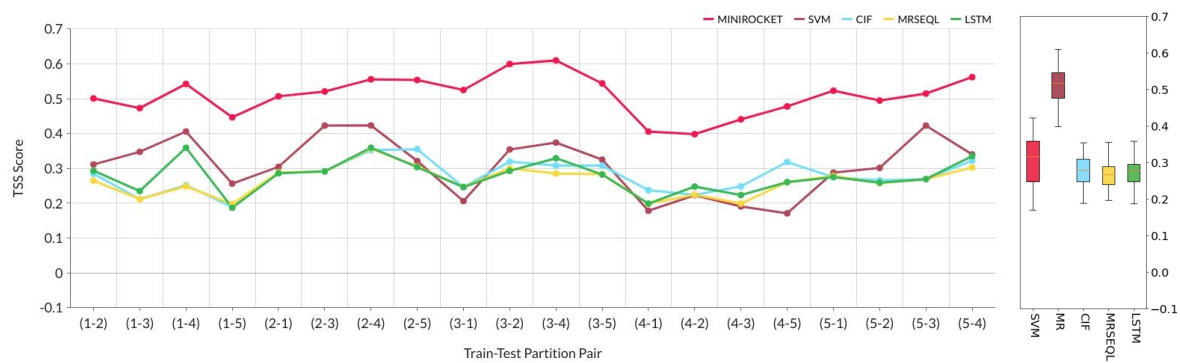


**Figure 5.** TSS score comparison of binary class classification after removing B and C class flares.

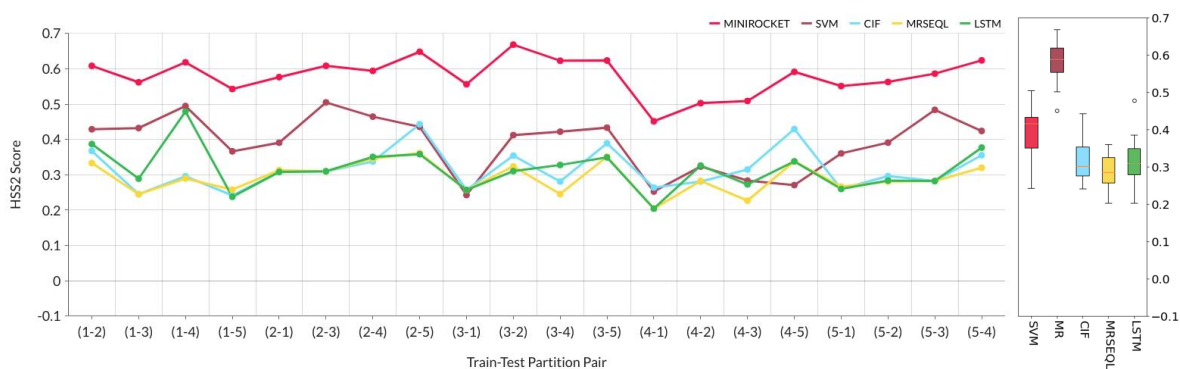


**Figure 6.** HSS2 score comparison of binary class classification after removing B and C class flares.

After removing the B and C class flares, our analysis for binary classification showed that MINIROCKET achieved a remarkable 30.06% increase in the TSS score and a 30.55% increase in the HSS2 score compared to aforementioned classifiers.



**Figure 7.** TSS score comparison of all class classification after removing B and C class flares.



**Figure 8.** HSS2 score comparison of all class classification after removing B and C class flares.

After removing the B and C class flares, we also conducted an analysis of the all-class classifications and observed that MINIROCKET again outperformed aforementioned classifiers by 20.13% in terms of TSS score and 18.94% in terms of HSS2 score. A notable observation arises from the analysis of the box plots in the conducted experiments. It has been deduced that the TSS and HSS2 scores for all classifiers experienced a surge when B and C class flares were excluded, in contrast to the scenario where these classes were included. This discovery greatly reinforces the findings of [9]. Moreover, the recurring observation showcased the superior performance of the MINIROCKET classifier over other classifiers. This trend was followed by the SVM, LSTM, CIF, and Mr-SEQL models.

## 6. Conclusions

In this study, we introduced the utilization of the MINIROCKET classifier for classifying the SWAN-SF dataset. We compared our model with alternative classifiers such as LSTM, Mr-SEQL, SVM, and CIF. To evaluate the performance, we utilized the True Skill Statistic (TSS) score and Heidke Skill Score (HSS2), widely employed metrics for evaluating flare prediction models dealing with class-imbalanced data. Our findings indicated that MINIROCKET outperformed other classifiers on the SWAN-SF dataset, demonstrating a consistent average improvement of 19.8% in TSS score and 20.92% in HSS2 score across all experimental settings. We also found that after removing B and C class flares, the trained models exhibited a significant improvement, resulting in a substantial increase in TSS and HSS2 scores. The removal of B and C class flares for maximizing flare prediction performance was also suggested by the experimental findings of multiple previous studies [20,21].

These results highlight the potential of our approach in enhancing the accuracy of solar physics and space weather forecasting. The effectiveness of MINIROCKET in handling MVTS data complexities is evident. It can significantly advance the vision of real-time solar flare classification. These contributions hold the promise for improving space weather forecasting.

As a direction for future research, we propose exploring a Transformers/Attention-based model integrated with the SWAN-SF dataset. This integration could address long-range dependencies and enable a comparative analysis against the benchmark MINIROCKET classifier, further advancing our understanding and capabilities in solar physics and space weather prediction.

**Funding:** This project has been supported in part by funding from the Division of Atmospheric and Geospace Sciences within the Directorate for Geosciences, under NSF awards #2301397, #2204363, and #2240022, and by funding from the Office of Advanced Cyberinfrastructure within the Directorate for Computer and Information Science and Engineering, under NSF award #2305781. The authors also acknowledge all those involved with the GOES missions as well as the SDO mission.

**Data Availability Statement:** The SWAN-SF dataset, as referenced in this manuscript [39], is accessible for download at the following URL: <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/EBCFKM>. The experimental results were generated using Python, which can be found at: <https://github.com/1420kartik/Solar-Flare-Classification.git>. Our code makes use of several libraries, including sktime [40], scikit-learn [41], matplotlib [42], and numpy [43].

## References

1. Ahmazdazadeh, A.; Aydin, B.; Georgoulis, M.K.; Kempton, D.J.; Mahajan, S.S.; Angryk, R.A. How to train your flare prediction model: Revisiting robust sampling of rare events. *The Astrophysical Journal Supplement Series* **2021**, *254*, 23.
2. Larsen, E. Predicting Solar Flares with Remote Sensing and Machine Learning. *CoRR* **2021**, *abs/2110.07658*, [2110.07658].
3. Abd Elrahman, S.M.; Abraham, A. A review of class imbalance problem. *Journal of Network and Innovative Computing* **2013**, *1*, 332–340.
4. Bobra, M.G.; Sun, X.; Hoeksema, J.T.; Turmon, M.; Liu, Y.; Hayashi, K.; Barnes, G.; Leka, K. The Helioseismic and Magnetic Imager (HMI) vector magnetic field pipeline: SHARPs–space-weather HMI active region patches. *Solar Physics* **2014**, *289*, 3549–3578.
5. Kusano, K.; Iju, T.; Bamba, Y.; Inoue, S. A physics-based method that can predict imminent large solar flares. *Science* **2020**, *369*, 587–591.
6. Dempster, A.; Schmidt, D.F.; Webb, G.I. Minirocket: A very fast (almost) deterministic transform for time series classification. In Proceedings of the Proceedings of the 27th ACM SIGKDD conference on knowledge discovery & data mining, 2021, pp. 248–257.
7. Dempster, A.; Petitjean, F.; Webb, G.I. ROCKET: exceptionally fast and accurate time series classification using random convolutional kernels. *Data Mining and Knowledge Discovery* **2020**, *34*, 1454–1495.
8. Middlehurst, M.; Large, J.; Bagnall, A. The canonical interval forest (CIF) classifier for time series classification. In Proceedings of the 2020 IEEE international conference on big data (big data). IEEE, 2020, pp. 188–195.
9. Bloomfield, D.S.; Higgins, P.A.; McAteer, R.J.; Gallagher, P.T. Toward reliable benchmarking of solar flare forecasting methods. *The Astrophysical Journal Letters* **2012**, *747*, L41.
10. McIntosh, P.S. The classification of sunspot groups. *Solar Physics* **1990**, *125*, 251–267.
11. Boubrahimi, S.F.; Aydin, B.; Kempton, D.; Angryk, R. Spatio-temporal interpolation methods for solar events metadata. In Proceedings of the 2016 IEEE International Conference on Big Data (Big Data). IEEE, 2016, pp. 3149–3157.
12. Cui, Y.; Li, R.; Zhang, L.; He, Y.; Wang, H. Correlation between solar flare productivity and photospheric magnetic field properties: 1. Maximum horizontal gradient, length of neutral line, number of singular points. *Solar Physics* **2006**, *237*, 45–59.
13. Jing, J.; Song, H.; Abramenko, V.; Tan, C.; Wang, H. The statistical relationship between the photospheric magnetic parameters and the flare productivity of active regions. *The Astrophysical Journal* **2006**, *644*, 1273.
14. Leka, K.; Barnes, G. Photospheric magnetic field properties of flaring versus flare-quiet active regions. II. Discriminant analysis. *The Astrophysical Journal* **2003**, *595*, 1296.
15. Yu, D.; Huang, X.; Wang, H.; Cui, Y. Short-term solar flare prediction using a sequential supervised learning method. *Solar Physics* **2009**, *255*, 91–105.
16. Song, H.; Tan, C.; Jing, J.; Wang, H.; Yurchyshyn, V.; Abramenko, V. Statistical assessment of photospheric magnetic features in imminent solar flare predictions. *Solar Physics* **2009**, *254*, 101–125.

17. Ahmed, O.W.; Qahwaji, R.; Colak, T.; Higgins, P.A.; Gallagher, P.T.; Bloomfield, D.S. Solar flare prediction using advanced feature extraction, machine learning, and feature selection. *Solar Physics* **2013**, *283*, 157–175.
18. Al-Ghraibah, A.; Boucheron, L.; McAteer, R. An automated classification approach to ranking photospheric proxies of magnetic energy build-up. *Astronomy & Astrophysics* **2015**, *579*, A64.
19. Nishizuka, N.; Sugiura, K.; Kubo, Y.; Den, M.; Watari, S.; Ishii, M. Solar flare prediction model with three machine-learning algorithms using ultraviolet brightening and vector magnetograms. *The Astrophysical Journal* **2017**, *835*, 156.
20. Bobra, M.G.; Couvidat, S. Solar flare prediction using SDO/HMI vector magnetic field data with a machine-learning algorithm. *The Astrophysical Journal* **2015**, *798*, 135.
21. Hamdi, S.M.; Kempton, D.; Ma, R.; Boubrahimi, S.F.; Angryk, R.A. A time series classification-based approach for solar flare prediction. In Proceedings of the 2017 IEEE International Conference on Big Data (Big Data). IEEE, 2017, pp. 2543–2551.
22. Angryk, R.A.; Martens, P.C.; Aydin, B.; Kempton, D.; Mahajan, S.S.; Basodi, S.; Ahmadzadeh, A.; Cai, X.; Filali Boubrahimi, S.; Hamdi, S.M.; et al. Multivariate time series dataset for space weather data analytics. *Scientific data* **2020**, *7*, 227.
23. Pesnell, W.D.; Thompson, B.J.; Chamberlin, P.C., The Solar Dynamics Observatory (SDO). In *The Solar Dynamics Observatory*; Chamberlin, P.; Pesnell, W.D.; Thompson, B., Eds.; Springer US: New York, NY, 2012; pp. 3–15. [https://doi.org/10.1007/978-1-4614-3673-7\\_2](https://doi.org/10.1007/978-1-4614-3673-7_2).
24. Hoeksema, J.T.; Liu, Y.; Hayashi, K.; Sun, X.; Schou, J.; Couvidat, S.; Norton, A.; Bobra, M.; Centeno, R.; Leka, K.; et al. The Helioseismic and Magnetic Imager (HMI) vector magnetic field pipeline: overview and performance. *Solar Physics* **2014**, *289*, 3483–3530.
25. Fisher, G.H.; Bercik, D.J.; Welsch, B.T.; Hudson, H.S. Global forces in eruptive solar flares: the Lorentz force acting on the solar atmosphere and the solar interior. *Solar Physics* **2012**, *277*, 59–76.
26. Leka, K.; Skumanich, A. On the value of  $\alpha$ AR' from vector magnetograph data. *Solar Physics* **1999**, *188*, 3–19.
27. Wang, J.; Shi, Z.; Wang, H.; Lue, Y. Flares and the magnetic nonpotentiality. *The Astrophysical Journal* **1996**, *456*, 861.
28. Schrijver, C.J. A characteristic magnetic field pattern associated with all major solar flares and its use in flare forecasting. *The Astrophysical Journal* **2007**, *655*, L117.
29. Guo, X.; Yin, Y.; Dong, C.; Yang, G.; Zhou, G. On the class imbalance problem. In Proceedings of the 2008 Fourth international conference on natural computation. IEEE, 2008, Vol. 4, pp. 192–201.
30. Allouche, O.; Tsoar, A.; Kadmon, R. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology* **2006**, *43*, 1223–1232. <https://doi.org/10.1111/j.1365-2664.2006.01214.x>.
31. Mason, J.P.; Hoeksema, J. Testing automated solar flare forecasting with 13 years of Michelson Doppler Imager magnetograms. *The Astrophysical Journal* **2010**, *723*, 634.
32. Woodcock, F. The evaluation of yes/no forecasts for scientific and administrative purposes. *Monthly Weather Review* **1976**, *104*, 1209–1214.
33. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Computation* **1997**, *9*, 1735–1780, [<https://direct.mit.edu/neco/article-pdf/9/8/1735/813796/neco.1997.9.8.1735.pdf>]. <https://doi.org/10.1162/neco.1997.9.8.1735>.
34. Cortes, C.; Vapnik, V. Support-vector networks. *Machine learning* **1995**, *20*, 273–297.
35. Lubba, C.H.; Sethi, S.S.; Knaute, P.; Schultz, S.R.; Fulcher, B.D.; Jones, N.S. catch22: CAnonical Time-series CHaracteristics. *CoRR* **2019**, *abs/1901.10200*, [[1901.10200](https://arxiv.org/abs/1901.10200)].
36. Middlehurst, M.; Large, J.; Bagnall, A.J. The Canonical Interval Forest (CIF) Classifier for Time Series Classification. *CoRR* **2020**, *abs/2008.09172*, [[2008.09172](https://arxiv.org/abs/2008.09172)].
37. Nguyen, T.L.; Gsponer, S.; Ilie, I.; O'Reilly, M.; Ifrim, G. Interpretable Time Series Classification using Linear Models and Multi-resolution Multi-domain Symbolic Representations. *CoRR* **2020**, *abs/2006.01667*, [[2006.01667](https://arxiv.org/abs/2006.01667)].
38. Nguyen, T.L.; Gsponer, S.; Ifrim, G. Time Series Classification by Sequence Learning in All-Subsequence Space. In Proceedings of the 2017 IEEE 33rd International Conference on Data Engineering (ICDE), 2017, pp. 947–958. <https://doi.org/10.1109/ICDE.2017.142>.
39. Angryk, R.; Martens, P.; Aydin, B.; Kempton, D.; Mahajan, S.; Basodi, S.; Ahmadzadeh, A.; Cai, X.; Filali Boubrahimi, S.; Hamdi, S.M.; et al. SWAN-SF **2020**. <https://doi.org/10.7910/DVN/EBCFKM>.

40. Löning, M.; Király, F.; Bagnall, T.; Middlehurst, M.; Ganesh, S.; Oastler, G.; Lines, J.; Walter, M.; ViktorKaz.; Mentel, L.; et al. sktime/sktime: v0.13.4 **2022**. <https://doi.org/10.5281/zenodo.7117735>.
41. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **2011**, *12*, 2825–2830.
42. Hunter, J.D. Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering* **2007**, *9*, 90–95.
43. Oliphant, T.E. A Guide to NumPy. *Trelgol Publishing* **2006**.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.