Time series analysis using MLP for solar flares based on GOES satellite data

Chris Ahouzi
Student ID: 1008074058

Brayden Wilson
Student ID: 1006905093

Mojtaba Khavaninzadeh
Student ID: 1006533324

Date: Nov 29th, 2024

Abstract

Solar flares

1 Introduction

The solar flares are classified based on their peak flux (in units of W/m^2) into 5 classes: X, M, C, B, and A, where the X-class have the the largest intensity and the intensity decreases by an order of 10 respectively. Furthermore each class is categorized into a numerical suffix from 1 - 9, where the suffix determines the factor of the event in the class [3]. M-class and X-class flares are one of the main reasons that heliophysicist study solar flares as they have direct impacts on Earth. M-class flares can cause brief radio blackouts at the earth poles and minor radiation storms that might endanger astronauts aboard the International Space Station (ISS). However the X-class flares can cause more severe problems such as creating long lasting radiation storms which can cause world-wide blackouts, interference in the global transmissions, harm satellites, and even give passengers flying near the poles small radiation doses [2]. This is why identifying the peak solar intensities and forecasting when they will happen in the future is of importance.

2 Data Assimilation

Our data as used throughout the paper is gathered from the Interactive Multi-Instrument Data of Solar Flares [1]. The settings we used was:

- Dates: 2002/01/01 0h 0m 0s to 2024/11/27 0h 0m 0s
- Position (in arcsec): X (-1100, 1100), Y (-1100, 1100)
- Flux duration: from detection of flare's rise to its peak to the end of it's life cycle
- Flux intensity: the peak intensity of the flare during it's life cycle
- Hiding instrument based filters and only using GOES Flare catalog, including A0 to X99 to account for all the intensities captured.

The training data is from 2002/01/01 0h 0m 0s to 2016/01/01. And the validation data is from 2016/01/02 to 2019/05/16. This is shown in Fig. 1.

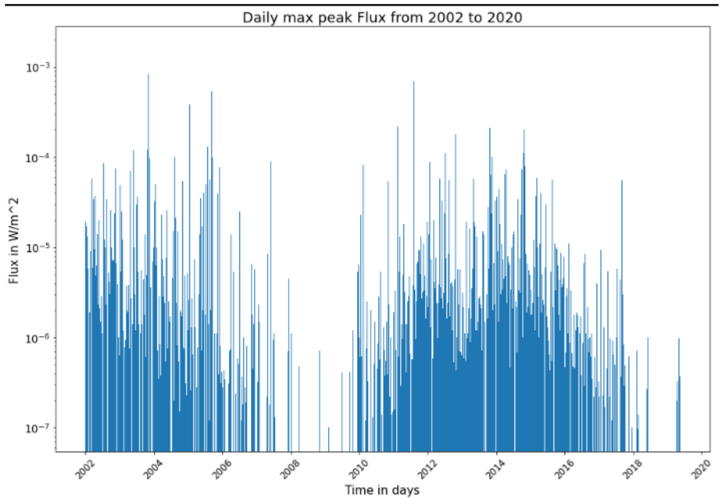

Figure 1: Daily max peak for the observed solar flares as seen by the GOES satellite as extracted by [1].

3 Methodology

3.1 Forecasting Solar Flux Approaches

The statistical relation between the flux of the solar flares and the measured solar flares has been a subject of study since the 1930's. More recently with the development of Neural Networks (NN) and applications of Artificial Intelligence (AI) as well as the launch of the network of satellites such as the Geostationary Operational Environmental Satellites (GOES), Solar and Heliospheric Observatory (SOHO) and Solar Dynamic Observatory (SDO), there have been many Machine Learning (ML) approaches that employed the data collected from the satellites to make a solar flare forecasting. Huang et al. employed a Convolutional Neural Network, a type of deep learning approach, and used the line-of-sight magnetogram of active solar regions [4] as features and the solar flares as response variables. In another work, Jiao et al. employed a Long Short Term Memory (LSTM) approach (another type of deep learning model) and also used the same features and variables (although focusing on different active solar regions and time-frames) [5].

In this work we will be focusing on using the max peak daily flux of the past seven days (obtained from the GOES satellite) and use that information to forecast the max peak flux of the next day. We will be implementing our neural network using the multilayer perceptron (MLP), another type of deep neural network.

3.2 Multilayer Perceptron Setup

A simple example of an MLP setup is shown in Fig. 2. The MLP includes a layer of inputs, hidden layers, and a layer of outputs. The use of MLP is common in the case of time-series forecasting as it is a feed forwarding algorithm. The input into the algorithm consists of events up to a certain point p (i.e. $[x_t, x_{t-1}, ..., x_{t-p}]$). Here we will be inputting the data for the last 7 days into the algorithm as follows

$$F_{t_i}, F_{t_i+1}, ..., F_{t_i+6} \tag{1}$$

where the F represents the solar flux. Since the solar flux was not directly available from the database, we converted it using the code below.

```python
#constants
m = {"X" : 1e-4, "M": 1e-5, "C": 1e-6, "B": 1e-7, "A": 1e-8} #classification for each solar flare GOES class

def goesclasstoflux(goesclass):
    class_letter = goesclass[0]#first element of GOES class
    intensity = float(goesclass[1:])#to convert all numbers to float

    flux = intensity*m[class_letter] #flux is the classification letter times the intensity for example C7.2 = 1e-6*7.2
    return f"{flux:.2e}" #to write in scientific noration

df["Flux (W/m^2)"] = df["GOES Class"].apply(goesclasstoflux) #applying the function to the GOES class column

print(df)
```

Next, we decided to proceed with 2 hidden layers and took as ansatz a size of 32 neurons for the first hidden layer and 16 neurons for the second hidden layer.

3.3

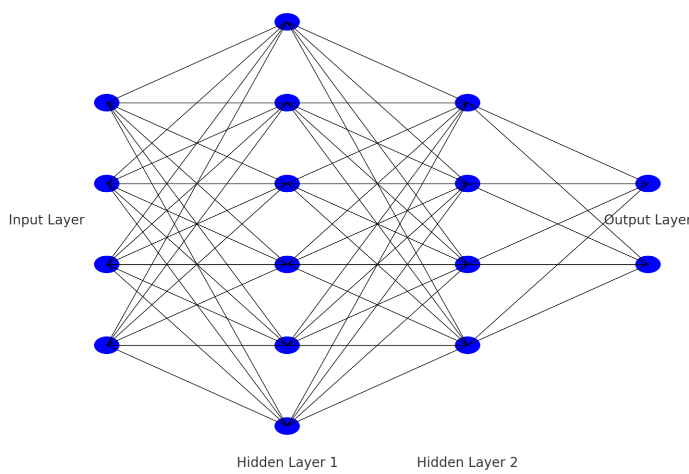

Figure 2: A general multilayer perceptron model with an input layer, 2 hidden layers, and an output layer. The neurons are shown in blue and the mapping of their relation (weights) is shown by a line connecting them.

4 Result

To set the data to be fed into the machine learning model. First, we must import the modules to be used, as the code below did.

pandas for managing the data, numpy for the arrays

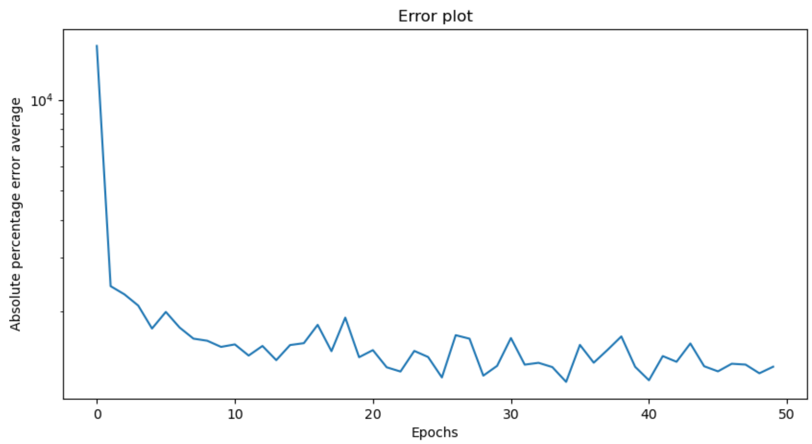

Figure 3: Absolute percentage error

5 Discussion

As mentioned in the Sec.3, the work by [4] and [5] use the solar flares as response variables and employ the magnetometers for active solar regions as predictors of the model. Given the time constraints for this project, by employing the use of MLP, we decided that while our response variable is indeed the solar flares, to instead find the corelation between the past solar flares in the 168 hrs cycle (7 day period). As can be seen in Fig. **??**

References

[1] National Aeronautics and Space Administration. *Heliophysics Data Portal: Solar Flares.* `https://data.nas.nasa.gov/helio/portals/solarflares`. Accessed Nov. 27th, 2024.

[2] National Aeronautics and Space Administration. *X-Class: A Guide to Solar Flares.* `https://svs.gsfc.nasa.gov/10109/`. Accessed: Nov. 27, 2024.

[3] European Space Agency. *What are solar flares?* `https://www.esa.int/Science_Exploration/Space_Science/What_are_solar_flares`. Accessed: Nov. 27, 2024.

[4] Huang et al. "Deep Learning Based Solar Flare Forecasting Model. I. Results for Line-of-sight Magnetograms". In: *The Astrophysical Journal* 856 (2018). Accessed: Nov. 28, 2024. DOI: `10.3847/1538-4357/aaae00`. URL: `https://iopscience.iop.org/article/10.3847/1538-4357/aaae00`.

[5] Jiao et al. "Solar Flare Intensity Prediction with Machine Learning Models". In: *arXiv preprint arXiv:1912.06120* (2019). Accessed: Nov. 27, 2024. URL: `https://arxiv.org/abs/1912.06120`.