UNIVERSITY OF COPENHAGEN

Geomagnetic Forecasting with Neural Networks

Machine Learning in Astrophysics

Ali Ahmad & Florent Imishti Mustafaj

Niels Bohr Institute

Contact Information:
Email (Ali): `xgh224@alumni.ku.dk`
Email (Florent): `wqb322@alumni.ku.dk`

Abstract

Geomagnetic storms, resulting from solar activity, pose significant risks to satellite operations, communication systems, and power grids. Accurate forecasting of these storms is crucial for mitigating their impacts [2]. As the final project in the course "Applied Machine Learning" at the University of Copenhagen, we explore the application of convolutional neural networks (CNNs) and recurrent neural networks (RNNs) to forecast geomagnetic storms using satellite image data from the Solar Dynamics Observatory (SDO). By leveraging solar images capturing phenomena such as solar flares and coronal mass ejections (CMEs) in the 171Å band, our neural network models are trained to identify patterns and temporal sequences indicative of the geomagnetic activity. Preliminary results demonstrate that the neural networks work well for geomagnetic forecasting on short timescales. Future work should focus on extending the models for predictions further into the future and perhaps also more specifically optimizing the models for geomagnetic storm prediction, if this is desired.

Introduction and background

The target variable for the model is, naturally, a measure of the geomagentic activity level. This is quantified by the Kp-index. The Kp-index ranges from 0 to 9, where higher values signify higher levels of geomagnetic activity. More specifically, the Kp-index represents the geomagnetic activity over a three-hour period, such that the data includes 8 measurements per day [1]. We consider all measurements between May 2010 and May 2024, giving a total of around 40 thousand measurements of the Kp-index. These measurements are paired with a roughly equal number of solar satellite images over the same period. An example of such a solar image can be seen in figure 3. Additionally, each Kp measurement and each solar image is paired with a timestamp for when the measurement was taken. These timestamps are also used in the machine learning model.

Data preprocessing

The data needs to be preprocessed before being passed to the model. This involved the following:

- Filtering out corrupted solar images by removing images with a total brightness under a certain threshold.
- Interpolating the Kp-index (using a cubic spline) to the timestamps of the solar images, since the Kp-indices are not measured at the same time as the solar images. Thus, each timestamp is associated with an image and an interpolated Kp-index.
- Encoding the periodicities of the physical system (11-year solar cycle, the sun's rotation around itself and the rotation of the Earth around the sun) as model inputs.
- Scaling the Kp-index such that the values range between -1 and 1.
- Scaling down the image resolution from 1024x1024 to 512x512 (for performance and memory efficiency). The images are further scaled such that all the pixel values are between -1 and 1, and each color channel has a mean and standard deviation of 0.5.
- Grouping the dataset into sequences for the RNN. We consider a sequence of 7 Kp-indices/solar images to predict the next value of the Kp-index.
- Splitting the data into test and train datasets (80% train, 20% test).

Model architecture

As mentioned, the model architecture consists of a convolutional neural network for processing the images, the output of which is fed into a recurrent neural network, which is responsible for the time series prediction. A more detailed overview of the model architecture can be found in figure 1.

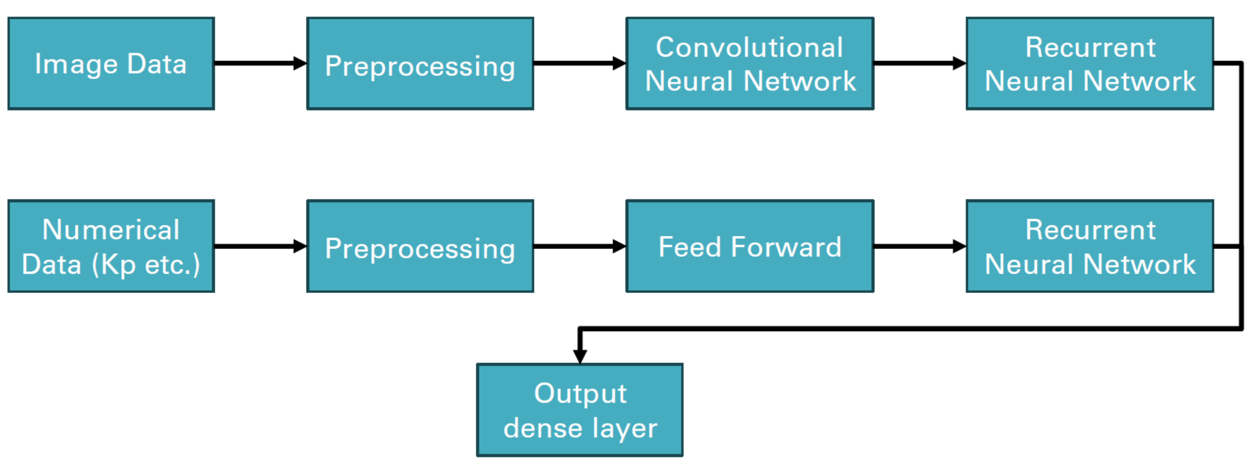

Figure 1: An overview of the model architecture. In the beginning, the image data and numerical features are treated separately; the image data is fed through a convolutional neural network, whereas the numerical data is fed through a feedforward neural network. Then, they are separately fed through long short-term memory RNNs, the outputs of which is combined and fed through a final dense layer in order to make the final prediction. The loss function is the mean squared error (MSE), the learning rate is 1.5×10^{-3} and the model is trained for 50 epochs.

Results

An important thing to note about the current model architecture is that we can only predict a single "step" (meaning 3 hours) into the future. This limitation arises from the fact that the input and output of the model are not the same; we input Kp-indices, images and various transformations of the timestamps and output a single prediction of the Kp-index. Thus, it is not trivially possible to feed the prediction back into the model to generate further predictions, as is usually done with simpler time series models.
Due to this limitation, we predict each future measurement in the test dataset by looking at the 7 previous measurements. This can be repeated to generate predictions for all measurements in the test dataset. The results of this process is shown for the first 100 data points in the left part of figure 2. Here, it can be seen that the model can consistently make relatively accurate predictions of the Kp-index 3 hours into the future. The plot of the left of figure 2 shows that the accuracy of the predictions is still mostly retained on data that is years into the future from the latest parts of the training sample. Although the model does struggle slightly with the large solar storms in May 2024.

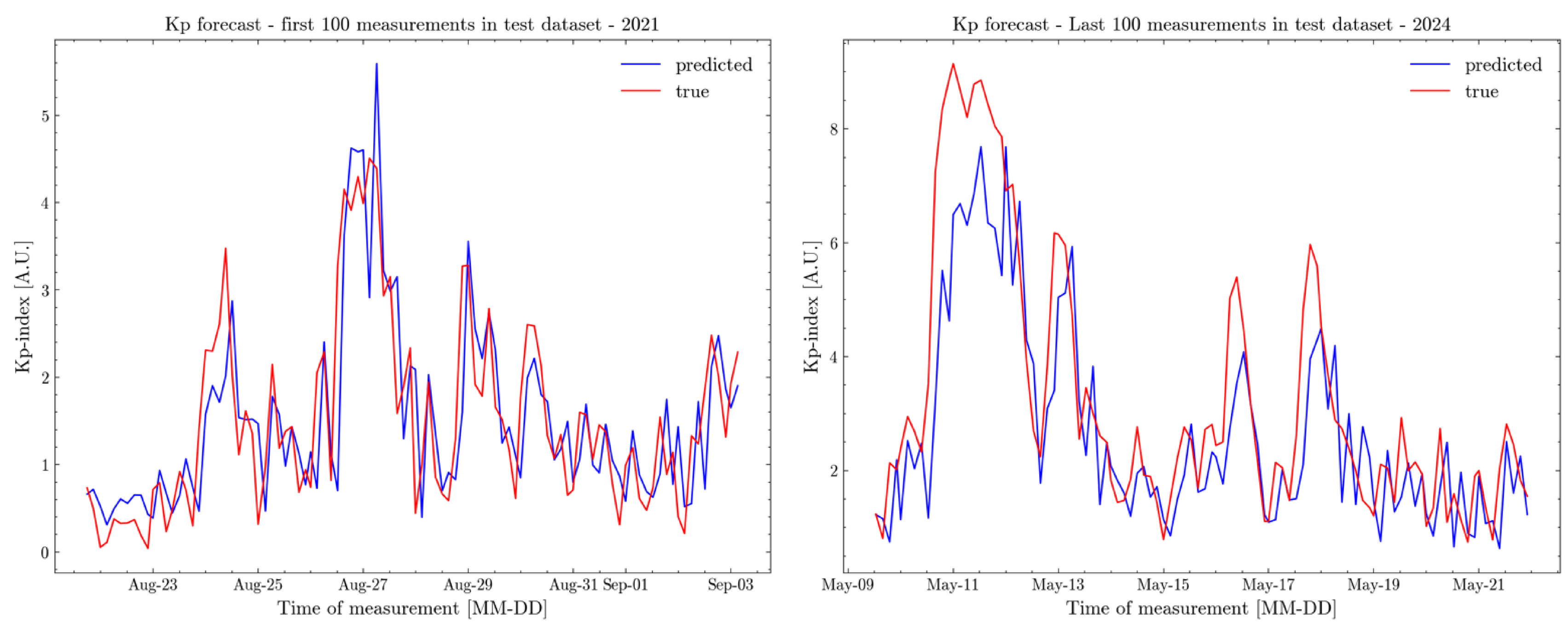

Figure 2: Geomagnetic forecasting of the first 100 measurements of the validation data (left), as well as forecasting of the last 100 measurements of the validation data (right). Note in particular the forecasted values for the May 2024 solar storms where $Kp > 5$ for the predicted values. Note that the Kp-index in this case is the interpolated Kp-index.

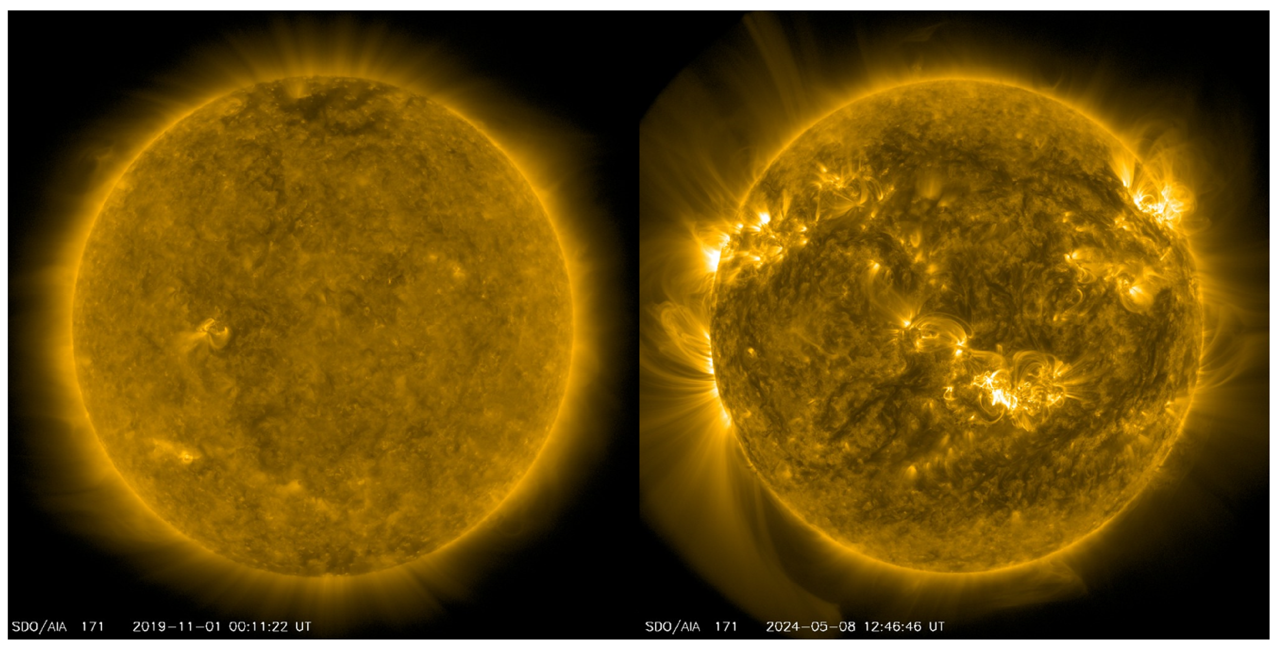

Figure 3: Example of two 171Å bandwidth image frame captured by the Atmospheric Imaging Assembly (AIA) instrument onboard the SDO, with low solar activity (left) & high solar activity before the May 2024 solar storms (right) (Credits: SDO/NASA)

Conclusions

We have shown that, at least on short timescales, neural networks are a suitable tool for forecasting the Kp-index using solar satellite images. Naturally, for practical applications, it will be of interest to make predictions further into the future than 3 hours. A possible way of doing this would be to use a concept familiar from natural language processing; sequence-to-sequence models [3].
If the practical application, that is specifically of interest, is storm prediction, it might be more sensible to frame the problem as a classification problem, specifically focused on predicting storm events. A quick and crude implementation of this was attempted. Although this showed some promise, it was clear that more work was required to make this practically useful.
All in all, given the limited time we had to work on it, our implementation serves as a proof of concept, demonstrating that the discussed methods are viable for predicting geomagnetic activity. However, more work is needed to make this practically useful.

References

[1] J. Bartels. *The geomagnetic measures for the time-variations of solar corpuscular radiation, described for use in correlation studies in other geophysical fields*. Ann. Intern. Geophys., 1957. Year **4**, p. 227-236.

[2] D. Conde, F. L. Castillo, C. Escobar, C. García, J. E. García, V. Sanz, B. Zaldívar, J. J. Curto, S. Marsal, and J. M. Torta. Forecasting geomagnetic storm disturbances and their uncertainties using deep learning. *Space Weather*, 21(11):e2023SW003474, 2023. e2023SW003474 2023SW003474.

[3] Chao Yang, Zhongwen Guo, and Lintao Xian. Time series data prediction based on sequence to sequence model. *IOP Conference Series: Materials Science and Engineering*, 692(1):012047, nov 2019.

Acknowledgements

We would like to acknowledge and give our deepest thanks to our former professor in the course "Applied Machine Learning" Troels C. Petersen from the Niels Bohr Institute for giving us excellent guidances and teachings in the topics of machine learning, that would prove invaluable to the succession & execution of this project.

We would also like to give a special thanks to Morten Holm for his greatly useful assistance & advice in developing key aspects of the solar model.