Long-Term X-Ray Solar Flare Intensity Prediction Using Machine Learning

Background

Solar flares are a space weather phenomenon which result from the sun shooting out electromagnetic waves, particularly in the forms of UV and X-ray radiation. When these flares reach the earth, they disrupt radio waves and signals created by satellites.[1] These disruptions can in turn lead to damages in many important systems, such as military communication, astronomical observations, weather forecasting, and global positioning systems (GPS). By creating prediction models that can forecast future severe solar flares, many industries will have the necessary time to prepare for these events.

Materials

1. PyCharm Integrated Development Environment
2. X-ray flare dataset from the National Oceanic and Atmospheric Administration (NOAA), specifically collected from Geostationary Operational Environmental Satellites (GOES)[2]
3. F10.7 Radio Emission dataset from NOAA, specifically from the Dominion Radio Astrophysical Observatory[3]
4. Python libraries such as Math, Matplotlib, and Keras
5. Ubuntu or any other file concatenation software

Procedure

1. Gather X-ray flare datasets from NOAA.[2]
2. Concatenate X-ray flare dataset files into one file using Ubuntu.
3. Convert X-ray flares into X-ray fluxes by translating data values from a letter notation system (A, B, C, M, X) into numerical fluxes in Watt per meter squared.
4. Normalize the X-ray flux dataset into decimals between 0 and 1. Prepare regression analysis by sorting the X-ray dataset into two sets, one being a set of arrays, and the other being a set of predictor values.
5. Construct a recurrent neural network utilizing the Keras library.
6. Forecast future X-ray flux values.
7. Project findings onto a graph by using the matplotlib library.

Engineering Goal

The goal of this project is to construct a recurrent neural network that adapts to X-ray flux timeseries derived from solar flares. In this way, machine learning can be used to generate a model that predicts the date and magnitude of future X-ray flares.

Results

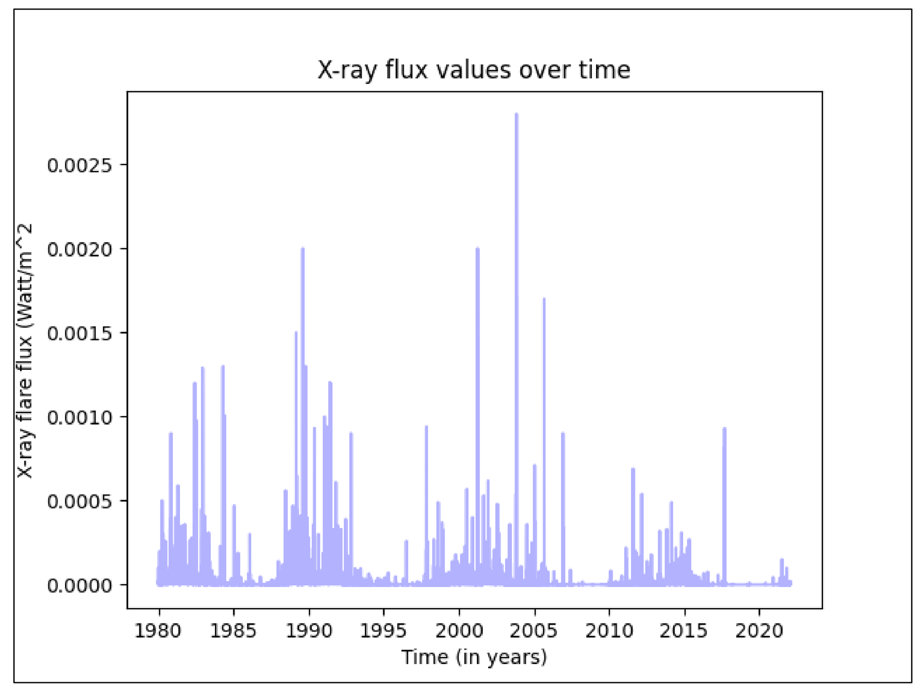

Figure 1. Graph of X-ray flux values over time

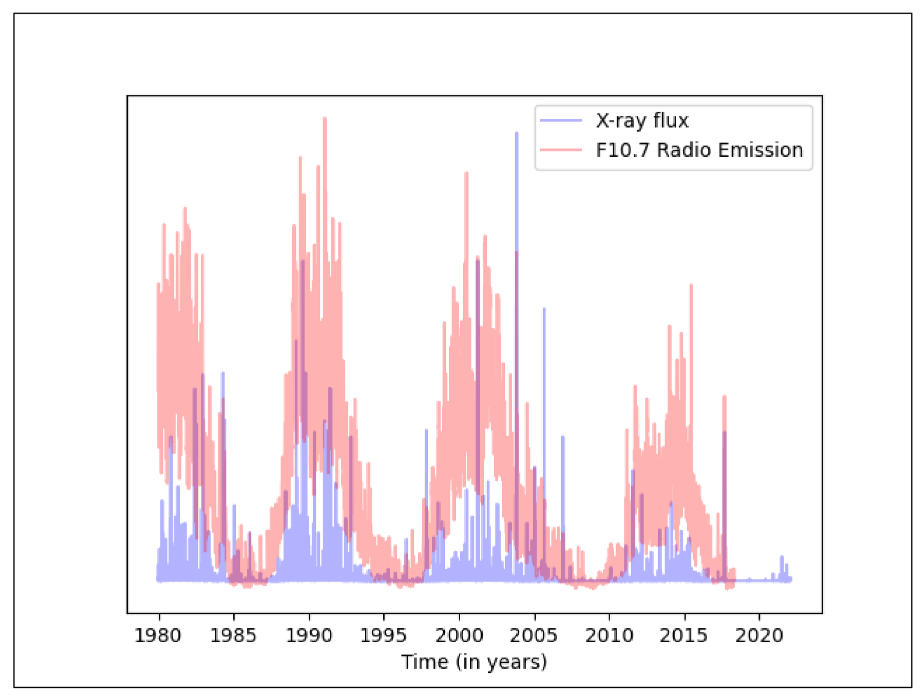

Figure 2. Graph of X-ray flux values overlayed with graph of F10.7 Radio Emission. F10.7 Radio Emission is a historic indicator of sun activity. These graphs showcase a noticeable correlation with each other. Furthermore, the X-ray flux graph clearly follows the 11-year sun cycle[4] (Data not to scale).

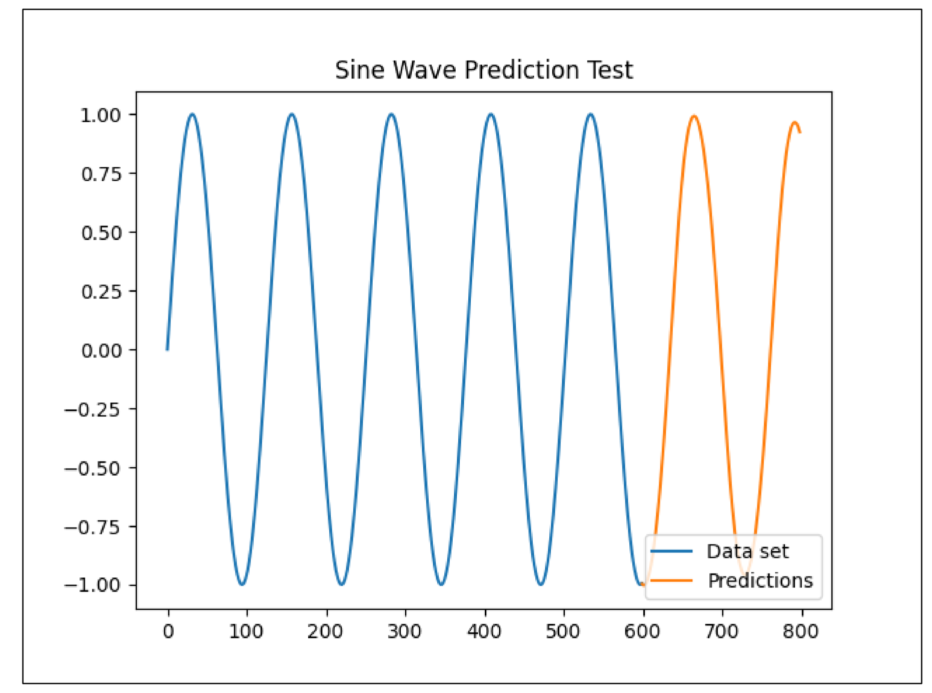

Figure 3. Sine waves used to test the reliability of the recurrent neural network

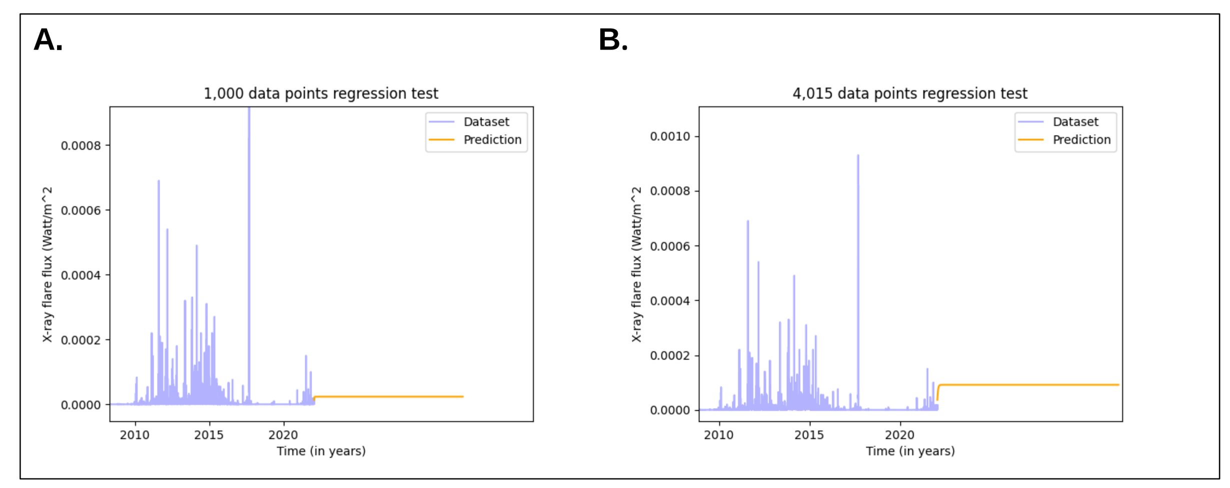

Figure 4. X-ray flux prediction tests. Two trials were recorded, one that used 1,000 datapoints for regression analysis (**A**), and another which used 4,015 datapoints (**B**). Both trials showcased unreliable results where the predictions that the recurrent neural network generated would flatline. Other factors in the algorithm did not showcase significant effect on the flatlining. The dataset's high variance might be causing this to occur.

Conclusion

Machine learning was unable to be utilized here due to the data's high variance among other factors, so a reliable prediction model was not able to be constructed. However, analysis of the data has still led to some interesting trends. The X-ray flux data corresponds highly to that of solar cycles through its correlation with the F10.7 Radio Emission graph. This gives the data an understandable pattern. In the future, this information may be used to further advance space weather predictions.

References

[1]Jyothi, S. (2021). Solar Superstorms: Planning for an Internet Apocalypse. *SIGCOMM '21*: 693.
[2]https://www.ngdc.noaa.gov/stp/space-weather/solar-data/solar-features/solar-flares/x-rays/goes/xrs/
[3]https://www.ngdc.noaa.gov/stp/space-weather/solar-data/solar-features/solar-radio/noontime-flux/penticton/penticton_observed/listings/listing_drao_noontime-flux-observed_daily.txt
[4]Bhowmik, P. and Nandy, D. (2018). Prediction of the strength and timing of sunspot cycle 25 reveal decadal-scale space environmental conditions. *Nature Communications*: 2.
All graphs and images created by me.