Sequential Time Analyses to Predict Space Weather

Zak Simpson - 230118073
Computer Science
University of Northern British Columbia
Prince George, Canada
zsimpson@unbc

Spencer Peterson - 230157543
Computer Science
University of Northern British Columbia
Prince George, Canada
speterson@unbc.ca

Ryan Van Aalst - 230152928
Computer Science
University of Northern British Columbia
Prince George, Canada
rvan@unbc

***Abstract*—Space weather is a description of the conditions in space around our solar system. These conditions can cause interference and damage to technology, both in space and on Earth's surface. Our research explores methods to mine patterns from space weather data and conduct time series analysis to forecast future weather. Space weather forecasts can be used to predict when periods of high solar activity occur, such that preventative measures can be taken to protect our technology. We explored three models for finding patterns in space weather data, and the results of these models' predictions were then compared to determine a most appropriate model for forecasting space weather..**

Keywords—Pattern Mining, Time Series, solar wind, Space Weather, Forecasting

I. Introduction

The sun is constantly sending charged particles into the solar system at speeds from 500 to 800 km/s. This is solar wind. When those particles interact with Earth's magnetic field, they can induce beautiful aurorae in the upper atmosphere. However, those same interactions, when strong enough, can cause geomagnetically induced currents in our electrical grids which cause physical damage. The total cost to repair and replace damaged equipment after the Quebec blackout of March 13, 1989 amounted to about $6.5 million. Oughton et al. [1] estimate a sufficiently severe solar storm can cost billions of dollars per day.

It is therefore of utmost importance to understand the sun and its behaviour. In this work we apply machine learning methods to data collected by the Deep Space Climate Observatory (DSCOVR) of the National Oceanic and Atmospheric Administration and predict solar wind density values.

The remainder of this paper is organized as follows. Section II consists of a brief review of related works and motivations. Section III contains the results of applying various machine learning models to DSCOVR data to predict the next values the probe will measure. We report on only the density of solar wind predictions.

II. Literature Review

In 2016 Oughton et al. [1] published a research article through the American Geophysical Union examining daily economic loss due to failure in the electrical transmission system of the U.S.A. In their paper, they researched how severe of an impact a Coronal Mass Ejection (CME) could have on America's economy if one were to hit. They discussed a split in how long of an outage a CME could cause. Some believe the outage would be a temporary one that will not last too long, while others believe it could be more prolonged. Either way though, with how many economies rely heavily on electronic devices, a CME could cause disruptions in supply chains and economic growth, resulting in potentially billions of dollars in loss of GDP per day.

The authors [1] looked at various scenarios for where blackouts could hit each region of the U.S.A., and from their analysis of all these scenarios, they found that a daily loss of GDP could range from as low as 6.2 billion USD to as high as 42 billion USD. However, for blackout zones affected by CMEs in the United States, they only represent about 49% of total economic cost on a larger scale. With these results in mind, the authors of the paper recommend that not only should their be investment in better space weather forecasting and mitigation systems, but that these systems must also consider the loss of indirect supply chains, as the increased use of certain manufacturing strategies increased the vulnerability to indirect supply chain loss in the event of a blackout.

In 2025, Guo et al. [2] published a research article in which they presented a 3D simulation of the formation of a CME. This allows them to better represent the behaviour of magnetic flux ropes, magnetic loops on the surface of the sun which can "snap" and cause a CME. The simulation better models flux ropes than the 2D simulations it was compared to, as it can model behaviour only possible in 3D space, such as rotation.

In 2024 He & Zhu [3] published a research article in which they analyzed the patterns of solar wind in solar cycles 20 through 24. A solar cycle lasts approximately 11 years, resulting in approximately 55 years of data being analyzed. This data was used to train a long short-term memory model, which then predicted the solar wind activity in solar cycle 25, which we are currently in. Their analysis also found a relationship between solar wind activity and the number of sunspots, a known indicator of solar activity.

He & Zhu [3] used a long short-term memory (LSTM) model with re-prediction. In contrast with traditional training methods for an LSTM, the input data for a prediction is the output of the latest forecast in the previous prediction. They claim this method better simulates how future real-world values are unknown and trains the model to make predictions using patterns in the data instead of comparing current trends to historical data. Using this technique, their model predicted trends for solar cycle 25 that align with predictions made in other studies.

III. Training & Evaluation

DSCOVR records a measurement of various aspects of solar wind every five minutes and publishes them to the NOAA website. In this study we consider a weeks' measurements and apply machine learning methods to mine the time series for indicators of when drastic shifts are about to occur. Figure 1 shows a plot of solar wind density, speed, and temperature. We were able to apply our desired methods to the density data only.

The analyses using Python and Tensorflow start at three random entries and proceed to make predictions. The purpose of these plots is to show how each learning model handles different variations in the data.

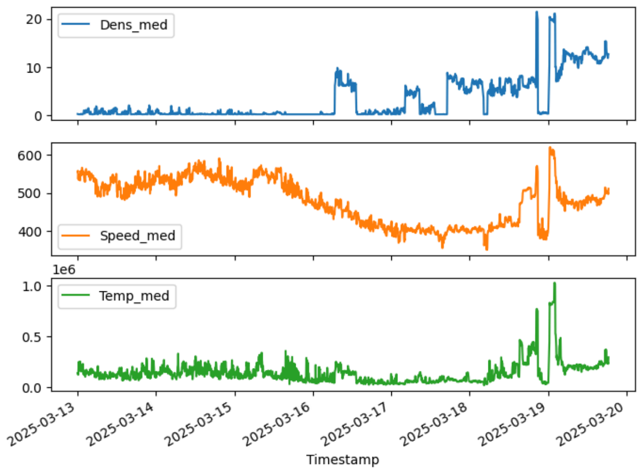

Figure 1: Seven days' worth of solar wind measurements

A. Baseline

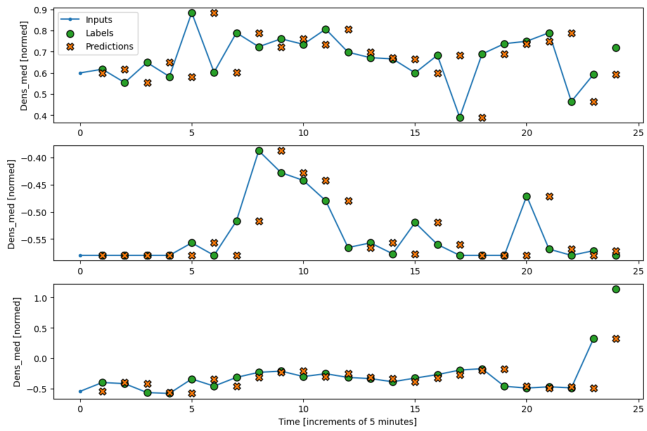

Figure 2: Baseline model density predictions and actual

The Baseline model simply subclasses the Model class from the Keras framework. Three random timestamps at least 125 minutes from the last measurement recorded are given to Baseline. It is then tasked with predicting each value including one it has not seen yet.

B. Dense

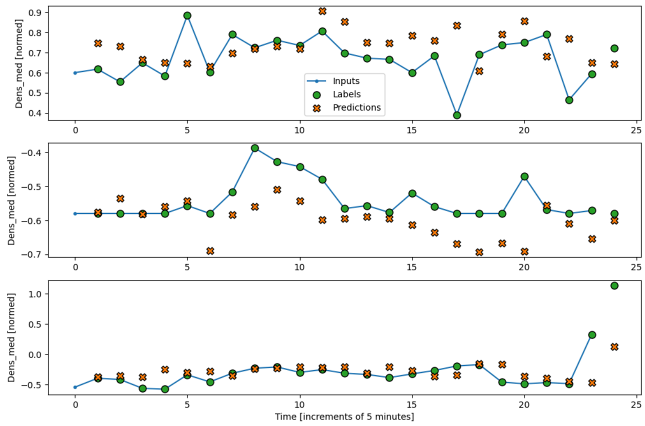

Figure 3: Dense model density predictions and actual

Dense is a simple neural network with 64 nodes in 2 hidden layers activated by the ReLu function. Repeated tests suggest that this model works well when there is little variation in the data–unsuitable for anticipating erratic activity by the sun.

C. Long Short-term Memory

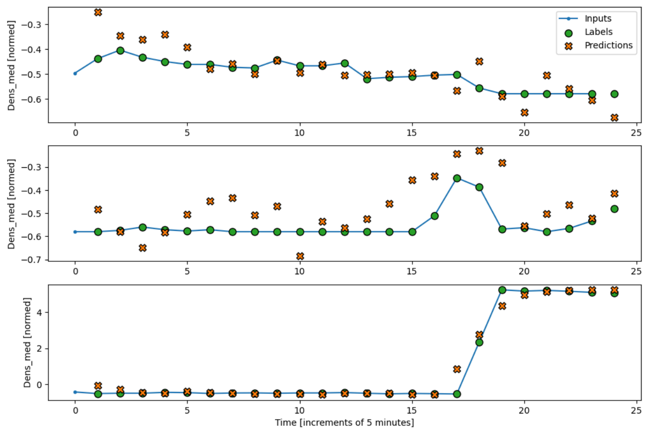

Figure 4: LSTM model density predictions and actual

The last model we applied is a LSTM Recurrent Neural Network. This model aims to solve the vanishing gradient problem, which can hinder the learning of a neural network model, by giving it a short-term memory that it can keep for a long time. The model learns to selectively remember information and forget what is no longer necessary.

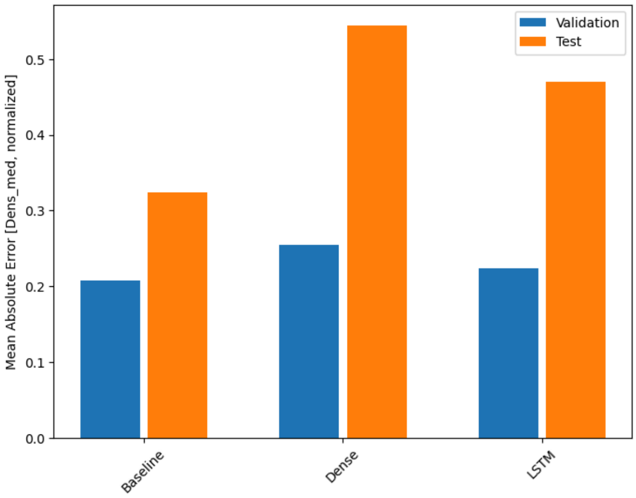

Figure 5: Mean Absolute Error of each model

These measurements of mean absolute error indicate that the more complicated models performed worse than the Baseline model in forecasting. This can be seen visually by comparing Figures 2 to 3 and 4; the prediction marks in the former are on average closer to the labels than in the latter two.

IV. Conclusion

Time series forecasting is a promising method of analyzing patterns in solar wind data, just as it is in forecasting weather on the surface of the Earth. We expect that more refined techniques in model creation, feature selection, performance measurement, and parameter-tuning will yield yet lower loss measurements. Such models will undoubtedly identify the patterns we sought but could not observe.

Acknowledgment

We would like to thank Dr. Fan Jiang and Colton Aarts for their suggestions in pursuing this topic.

References

[1] Oughton, E. J., A. Skelton, R. B. Horne,A. W. P. Thomson, and C. T. Gaunt (2017), Quantifying the daily economic impact of extreme space weather due to failure in electricity transmission infrastructure, Space Weather, 15, 65–83, doi:10.1002/2016SW001491.

[2] Guo, J., W.Ni, Y., Schmieder, B., Guo, Y., Xia, C., Devi, P., Chandra, R., Poedts, S., Joshi, R., Zhou, Y.H., Li, H.T., & Chen, P.F. (2025). The Birth of a Major Coronal Mass Ejection with Intricate Magnetic Structure from Multiple Active Regions.

[3] He, M., Zhu, H. Statistical analysis and forecasting of solar wind parameters across solar cycles. Sci Rep 14, 19529 (2024), doi:10.1038/s41598-024-70564.