DS207 - Time Series Forecasting - Individual Project Proposal

Anahit Baghdasaryan

Time-Series Forecasting of Daily Sun Spot Data: Applications and Techniques:

Daily Sun Spot Data refers to the daily observations of the number of sunspots, or dark regions on the surface of the sun, that have been recorded by astronomers since 1818. The study of sunspot data has been an important area of research in solar physics, as sunspots are closely linked to the sun's magnetic activity and can affect Earth's climate and space weather. Time-series forecasting, which involves using statistical and machine learning techniques to make predictions based on past observations of a variable over time, has been widely used to analyze and predict sunspot activity. Time-series models can help identify patterns and trends in sunspot data and provide forecasts of future activity, which can be valuable for a range of stakeholders, including solar energy companies, climate scientists, space weather forecasters, and amateur astronomers. Overall, daily sunspot data and time-series forecasting have been key tools in the study of the sun's behavior and its impact on Earth and the space environment. [1]

Research question:

How accurately can we forecast daily sunspot numbers using time series forecasting models, and what factors (such as solar cycles and other environmental factors) impact the accuracy of these predictions over different time horizons? [2]

Information about organizations and/or part of society who will gain from this research (stakeholders):

Some several organizations and groups could potentially benefit from research on daily sunspot data and time series forecasting:

Solar energy companies: Accurately forecasting daily sunspot numbers can help solar energy companies better predict and plan for fluctuations in solar radiation, which can impact the efficiency of solar panels and other solar energy technologies. [3]

Climate scientists: Sunspots are a key indicator of solar activity, which can impact the Earth's climate in various ways. By better understanding and forecasting sunspot activity, climate scientists may be able to improve their understanding of how the Earth's climate is changing and how it may continue to change in the future. [4]

Space weather forecasters: Sunspot activity can also impact space weather, which can, in turn, affect satellite communications, GPS systems, and other technologies. Accurate forecasting of sunspot activity can help space weather forecasters better predict and prepare for space weather events. [5]

Amateur astronomers: Sunspot activity is of interest to many amateur astronomers, who may use this data to track solar activity and better understand the dynamics of the Sun. Accurate time series forecasting of sunspot activity could help amateur astronomers plan their observations and better understand long-term trends in solar activity. [6]

Data Description:

The daily sunspot data set contains observations of the number of sunspots on the surface of the sun from the years 1818 to 2019. The data was collected by the Royal Observatory of Belgium, including daily sunspot counts and monthly and yearly averages. The data set contains 73,583 observations, with each observation corresponding to a single day. The data has been normalized and rescaled to a range between -1 and 1. The data set can be used to study the patterns and trends in sunspot activity over time and to make predictions about the future activity using time-series forecasting techniques. The data set is made available on Kaggle, a platform for sharing and discovering data sets for use in research and analysis.

https://www.kaggle.com/datasets/abhinand05/daily-sun-spot-data-1818-to-2019

Methods and similar project(s) :

As a result of the research, several methods were found; depending on the material being covered during the DS207 Time Series Forecasting course, some of those might be applied accordingly.

Methods:

- ARIMA (autoregressive integrated moving average) models: a class of linear models that use past values of a time series to make predictions about future values. [7] , [8]
- SARIMA (seasonal autoregressive integrated moving average) models: an extension of ARIMA models that incorporate seasonal trends in the data. [9]

- Exponential smoothing methods: a class of models that weigh past time series observations using an exponentially decreasing function, giving more weight to recent observations.[10]
- Neural network models: a class of machine learning models that can be used for time-series forecasting by training a neural network on past time series values. [11],[12]

Similar projects:

- "Sunspot Activity Prediction with Neural Networks": a project that uses a neural network to predict sunspot activity levels based on historical data. [13]
- "Sunspot Forecasting Using Time Series Analysis": a project that uses ARIMA and exponential smoothing methods to forecast sunspot activity. [14]
- "Forecasting Sunspot Cycles Using ARIMA Models": a paper that uses ARIMA models to forecast the 11-year cycle of sunspot activity. [15]
- "A Hybrid ARIMA and Artificial Neural Network Model for Sunspot Forecasting": a paper that combines ARIMA and neural network models for forecasting sunspot activity. [16]

References:

[1] Lantos, P. M. (2017). Solar physics: Sunspots and the sunspot cycle. In J. P. Zombeck, M. Collados, & M. Schüssler (Eds.), Handbook of the Solar-Terrestrial Environment (pp. 33-58). Springer.

[2] Nandi, S., Sankarasubramanian, R., & Ghosh, S. (2016). Sunspot time series forecasting using machine learning algorithms. Journal of Atmospheric and Solar-Terrestrial Physics, 149, 153-165.

[3] Makrides, G., Christodoulou, C., Kaminaris, S., Soteris, A., & Georghiou, G. E. (2015). Solar radiation forecasting using machine learning techniques: A review. Solar Energy, 122, 81-89.

[4] Mursula, K., & Usoskin, I. G. (2016). Sunspot cycles: Basics and new findings. Living Reviews in Solar Physics, 13(1), 3.

[5] Pulkkinen, A. (2016). Space weather: Terrestrial perspective. Living Reviews in Solar Physics, 13(1), 5.

[6] Clette, F., Svalgaard, L., Vaquero, J. M., & Cliver, E. W. (2014). Recurring features in the historical record of solar activity: A review. Space Science Reviews, 186(1-4), 35-78.

[7] Box, G. E. P., & Jenkins, G. M. (1976). Time Series Analysis: Forecasting and Control. Holden-Day.

[8] Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: Principles and Practice (2nd ed.). OTexts. https://otexts.com/fpp2/arima.html

[9] Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: Principles and Practice (2nd ed.). OTexts. https://otexts.com/fpp2/sarima.html

[10] Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: Principles and Practice (2nd ed.). OTexts. https://otexts.com/fpp2/expsmooth.html

[11] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

[12]Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to Forget: Continual Prediction with LSTM. Neural Computation, 12(10), 2451-2471.

[13] V. Sandeep Kumar. (2018). Sunspot Activity Prediction with Neural Networks. Kaggle. https://www.kaggle.com/sandeepkumar121995/sunspot-activity-prediction-with-neural-networks

[14] M. Ali. (2018). Sunspot Forecasting Using Time Series Analysis. GitHub. https://github.com/mirali361/Sunspot-Forecasting-Using-Time-Series-Analysis

[15] C. Y. Liu, S. Q. Cao, L. H. Lu, & X. Y. Zhang. (2013). Forecasting Sunspot Cycles Using ARIMA Models. Advances in Astronomy, 2013, 1-8.

[16] M. T. Ullah, M. Z. Uddin, & M. A. Matin. (2015). A Hybrid ARIMA and Artificial Neural Network Model for Sunspot Forecasting. International Journal of Scientific & Engineering Research, 6(4), 1283-1291. https://www.ijser.org/researchpaper/A-Hybrid-ARIMA-and-Artificial-Neural-Network-Model-for-Sunspot-Forecasting.pdf