

D. N. V. S. Vijaya Lakshmi¹,
 V.Rambabu²,
 Ch S K Chaitanya³,
 Bolla Eshwar⁴,
 Navya Padma Priya⁵

Machine Learning for Space Weather Prediction in Communication Networks



Abstract: - The Space Weather Forecasting Guide is a comprehensive overview of the literature detailing the scientific foundations for predicting space weather, specifically concerning the Sun and the heliosphere. It is always updated to include the latest advancements in solar physics and space weather.

This paper is a deliverable of the "SIDC Data Exploitation" project, financed by PRODEX. The space weather research conducted within this project enhances the forecast capacities of the ISES Regional Warning Center Belgium, managed by ROB/SIDC, and the Space Weather Forecasting Guide encapsulates the information gained via the project.

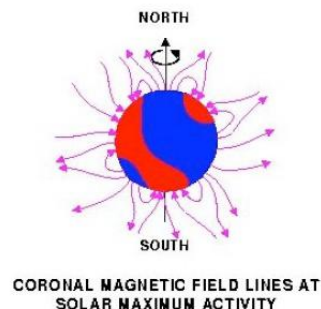
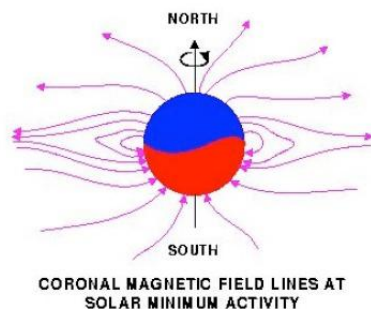
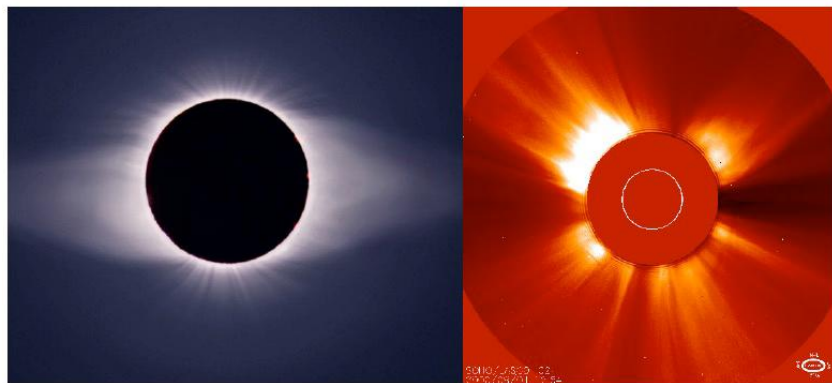
Index Terms: Machine Learning, Matlab, Neural Networks.

INTRODUCTION

Background solar wind

The corona is markedly non-uniform and shaped by the magnetic field that extends into interplanetary space, influencing interplanetary magnetic configurations (Figure 1). Charged particles are directed down the field line through the heliosphere, contributing to the solar wind.

Observations from the HELIOS spacecraft have shown that the solar wind has a bimodal character (Bame et al. 1977, Schwenn et al. 1978): fast and slow winds that vary not only in velocity but also in their attributes and origin inside the corona. The two-stream quasi-steady solar wind constitutes the so-called 'background solar wind.' The typical properties of these two stream types are encapsulated in Table 1.



^{1,2,3,4,5} International School of Technology and Sciences for Women

Table 1

Parameter	Slow Wind (<400 km/s)	Fast Wind (>600 km/s)
v_p [km/s]	327	702
n_p [cm ⁻³]	8.3	2.7
F_p [10 ⁸ cm ⁻² s ⁻¹]	2.7	1.9
n_a/n_p [%]	3.8 (highly variable)	4.8 (stationary)
T_p [10 ⁴ K]	3.4	23
Source	Helmet streamers, loops	Coronal holes
Signatures	Very variable	Stationary

The slow wind (speed under 400 km/s) emanates from areas where the magnetic configuration is characterized by complete loops, with both ends of the magnetic field lines anchored on the Sun, often at low latitudes during solar minimum and at any latitude at solar maximum. The high-velocity wind (exceeding 600 km/s) mostly emanates from coronal holes, areas characterized by rapidly growing magnetic fields that facilitate the escape of high-speed solar wind into the heliosphere (Cranmer et al. 1999).

The radial outflow of the solar wind causes rapid streams from equatorial coronal holes to encounter the slower wind in front, resulting in the formation of 'co-rotating interaction zones' or CIRs (Figure 2). This interaction creates a zone of compressed solar-wind plasma at the forefront of the rapid stream and an area of rarefied solar-wind plasma at its rear, which, as indicated by the nomenclature, co-rotates with the Sun (Hundhausen, 1972). Owing to the symmetry of the pressure augmentation induced by compression, shocks are propagated away from the pressure elevation in both directions (Gosling, 1996). This produces a pair of shocks: a forward shock (FS) that moves away from the Sun and a reverse shock (RS) that moves towards the Sun. It is important to observe that the RS often occur at distances beyond 1 AU.

An observer situated near the solar equatorial plane of the Sun will see alternating rapid and slow solar wind streams. Specifically, during the solar minimum, when coronal holes are positioned at the poles, alternating slow (low-latitude) and rapid (high-latitude) solar-wind streams will be seen sweeping past an observer owing to the inclination of the Sun's magnetic axis in relation to its rotational axis. The presence of latitudinal differences in solar wind speed has also been confirmed by in-situ observations.

Predicting the arrival of high speed streams (HSS) to the Earth

Upon identification of a coronal hole (CH) on the solar disc as it traverses the central meridian, the forecaster's task is to anticipate the arrival time of the high-speed solar wind stream (HSS) to Earth and the intensity of the resultant geomagnetic storm. The arrival of the High-Speed Stream (HSS) is contingent upon the velocity of the fast solar wind and is often anticipated around three days after the passage of an equatorial Coronal Hole (CH) over the Sun's central meridian (e.g., Temmer et al. 2007). It may be monitored in situ by satellites positioned near the L1 point.

The HSS will exhibit varying degrees of geoeffectiveness based on its velocity and the arrangement of the interplanetary magnetic field, particularly the presence of negative B_z . Empirical approaches and physics-based modeling are now being developed to assist forecasters in their estimations.

Nonetheless, as will be seen in the subsequent sub-sections, these models exhibit several deficiencies and are hence not entirely precise. A comprehensive understanding of the dynamical processes at all scales in the corona and their propagation throughout the heliosphere is essential, alongside a more precise knowledge of the three-dimensional magnetic configuration.

LITERATURE SURVEY

Nearly symmetric frontside complete halo coronal mass ejections are often oriented towards Earth. Schwenn et al. (2005) observed that around 7% of frontside complete halo coronal mass ejections (CMEs) did not reach Earth. A forecaster must additionally consider the location of the halo CME source zone on the solar disk.

Earth-directed coronal mass ejections (CMEs) have a distinct concentration of source areas in proximity to the solar disc center (distance to center < 0.4 solar radii) (Cane et al. 2000, Wang et al. 2002, Zhang et al. 2003, Manoharan et al. 2004, Srivastava and Venkatakrishnan 2004, Zhang et al. 2007, Dumbovic et al. 2015). Source locations situated farther from the disc center may be linked to partial halos. In the majority of these instances, only an interplanetary shock is detected (e.g. Manoharan et al. 2004) upon reaching Earth. The coronal mass ejection (CME) bypasses the Earth; nevertheless, the angular area of the shock exceeds that of the associated CME (Bothmer and Schwenn 1998). The storm may thereafter be generated by strong southern fields inside the sheath (Huttunen et al. 2002, Rodriguez et al. 2009). The greater the distance of the source area from the disc center, the higher the likelihood of encountering just a shock or no CME-associated structure at all. Twenty-five percent of frontside partial halo coronal mass ejections do not reach Earth (Schwenn et al. 2005). Wang et al. (2002) and Zhang et al. (2003) indicated that geoeffective CMEs have a marginal inclination to originate from the western solar hemisphere (see, however, Cane et al. 2000, Srivastava and Venkatakrishnan 2004). This phenomenon may be elucidated (Wang et al. 2004) by the observation that coronal mass ejections (CMEs) exceeding the velocity of the surrounding solar wind are redirected eastward by the magnetic influence of the ambient spiral interplanetary magnetic field (IMF), but a comprehensive dynamic model of this interaction yet to be established.

Certain ICMEs detected near Earth either lack associated CMEs (Cane and Richardson 2003) or seem to originate from a non-halo CME (Schwenn et al. 2005). This may result from inadequate LASCO sensitivity for halo CMEs that are predominantly situated far off of the plane of the sky. In many instances, CME signals in EIT data are readily identifiable, and a meticulous examination of the LASCO data may result in the discovery of the matching halo CME that was previously undetected (Zhukov 2005). Overall, LASCO sensitivity is sufficient to identify even feeble CMEs.

A straightforward method for recreating CMEs with a single viewpoint is to develop a geometric representation of the CME's outer circular contour, disregarding its interior composition. An illustrative instance is the CME cone model, which, with varying levels of complexity, has been used to represent halo CMEs as circles (Zhao et al. 2002, Xie et al. 2004, Xue et al. 2005, Michalek 2006) or ellipses (Zhao 2005, Cremades and Bothmer 2005, Michalek 2006, Zhao 2008). Based on the model parameters, the fitting may provide estimates of the source region's location, CME breadth, propagation direction, and velocity.

Cremades and Bothmer (2004) extracted the geometrical characteristics of structured (three-part) CMEs from a collection of 124 flux-rope CMEs recorded in LASCO-C2 data. Thernisien, Howard, and Vourlidas (2006) devised a forward-modeling approach for flux-rope-like coronal mass ejections (CMEs) to replicate their morphology based on this work.

The recent data from the Solar Terrestrial Relations Observatory (STEREO) (Kaiser et al. 2008), established in October 2006, provide stereoscopic pictures of the Sun's atmosphere.

Utilizing triangulation on coronal mass ejections (CMEs) seen by the Sun Earth Connection Coronal and Heliospheric Investigation (SECCHI) instrument suite (Howard et al. 2008) onboard STEREO enables 3-D calculations of the structural and kinematic properties of CMEs.

Subsequently, we delineate each of the aforementioned ways together with the corresponding online applications, if applicable. A comparative analysis of CME reconstruction before to and during the STEREO period is presented in Thernisien et al. 2011.

PREDICTING FLARES

The magnetic field is the principal catalyst of solar activity throughout periods from millennia to seconds. Eruptive phenomena transpire on the Sun when the local magnetic field arrangement becomes unstable. The accumulated magnetic energy is then unleashed to propel substantial quantities of chromospheric or coronal material (such as filament eruptions and coronal mass ejections) or to accelerate particles (including electrons, protons, or ions) inside the solar environment.

In a simplified manner, these particles dissipate their energy via collisions with the surrounding plasma, which subsequently becomes heated and emits soft X-ray (SXR) and extreme ultraviolet (EUV) radiation. Flares serve as the observational indicators of this activity. In actuality, emissions are not confined to the SXR or EUV domains but are distributed over the whole electromagnetic spectrum, ranging from low-frequency radio waves to gamma rays for the most intense phenomena. Flares arise throughout solar eruptions; however, this forecaster guidance

only addresses flares occurring in active zones. Within the space weather community, the soft X-ray (SXR) characteristic of flares, as recorded by the NOAA Geostationary Operational Environmental Satellites (GOES) since 1976, has established itself as the definitive standard for flare research and discourse, which will be used in this book. The various flare classifications based on the GOES scale are delineated. Flares, particularly intense ones, arise in magnetically intricate active areas; hence, flare predictions often associate the observable characteristics of active regions (AR) at photospheric levels (magnetograms, white light pictures) with the likelihood of a flare event. The rationale is that these data types have been accessible for several years (spanning several decades or even centuries for sunspot observations) and that various observatories continue to provide these observations regularly. Analyzing these photospheric characteristics in conjunction with the soft X-ray flare history documented by NOAA is fundamental to any research focused on flare prediction. Throughout the years, terrestrial observers have compiled libraries of data and classifications designed to elucidate the intricate nature of sunspot groups and magnetic configurations. This chapter presents statistical techniques developed from these catalogs and the flare history of active locations, including all the necessary information for precise flare prediction. With the emergence of enhanced computer capabilities and the advancement of artificial intelligence methodologies, many teams have been engaged in extracting relevant information directly from the original data, considering the rapid temporal evolution.

This section first examines the instruments used to evaluate the success of various forecasting approaches, followed by a presentation of both conventional and innovative techniques.

PERFORMANCE METRIC

The flare forecasting issue may be articulated as the development of a regression model that correlates variable measurements X with the observed reaction Y . Measurements X may pertain to historical flaring activity or attributes of an active area derived from continuum and white light imagery.

Response Y indicates the flare production of the associated active zone in the imminent future, often within the next 24 to 48 hours. Y may represent an ordered collection of variables (such as cumulative flare strength), in which case the link function will denote the conditional expectation of Y given X and may be analyzed using a regression model. In flare forecasting applications, Y is often specified as class membership (the class of AR with at least one flare $\geq C1$; $\geq M1$, etc.); in the context of a discrimination model, the link function represents the likelihood of belonging to a certain class given the value of X .

A "flare/no-flare" discriminating threshold must be implemented to these probabilities to ascertain the quantity of "predicted flare" and "predicted no-flare" instances. Subsequently, it is feasible to juxtapose these anticipated values with the "ground truth" supplied by online catalogs of active locations and flares. The observed and predicted values of "flare/no-flare" states facilitate the construction of a contingency (or confusion) matrix, as illustrated in Table 3. This matrix comprises the elements TP (true positives, where "flare" is both predicted and observed), FN (false negatives, where "no flare" is predicted but a flare is observed), FP (false positives, where "flare" is predicted but none is observed), and TN (true negatives, where "no flare" is predicted and none is observed).

Problem formulation and data acquisition

Tim Mitchell [18] defined a well-posed machine learning issue as follows: "A computer program is considered to learn from experience E concerning task T and performance measure P if its performance on T , evaluated by P , enhances with experience E ." This research defines the machine learning issue as the job of VTEC forecasting, using training data as experiences and selecting root mean square error (RMSE) and correlation coefficient as performance metrics. The objective of anticipating space weather phenomena in the ionosphere is established by the forecasting of VTEC in that region. The issue is articulated in a manner conducive to resolution via guided learning. Supervised learning may be seen as a function approximation or predictive learning challenge [19]. The objective is to derive an approximation of the function that accurately characterizes the connection between the input (predictors, features, or independent variables) and output (response or dependent variable) vectors, using a training sample. This is accomplished by minimizing a specific loss function over the joint distribution of all values. This work utilizes data on solar activity, solar wind, and the geomagnetic field obtained from NASA/GSFC via OMNIWeb (<https://omniweb.gsfc.nasa.gov/form/dx1.html>) [20]. The VTEC values are obtained for high, mid, and low latitude locations from the Global Ionosphere Map (GIM) supplied by CODE.

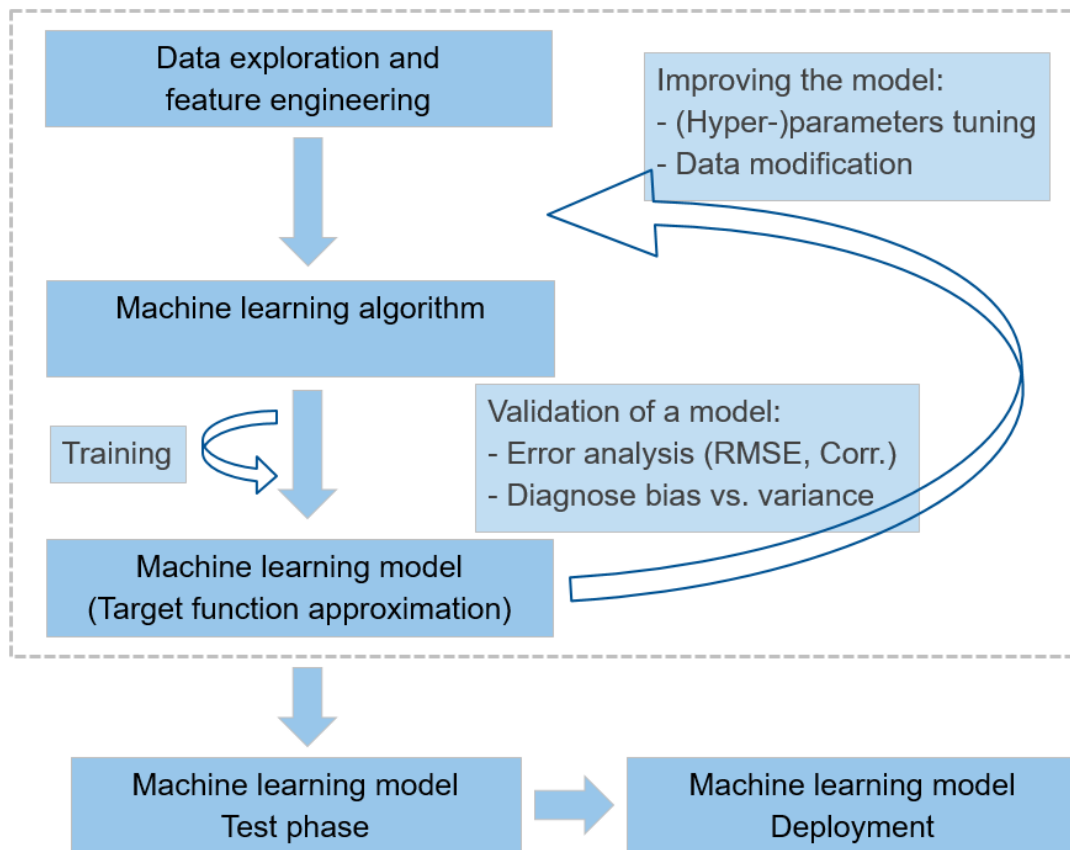
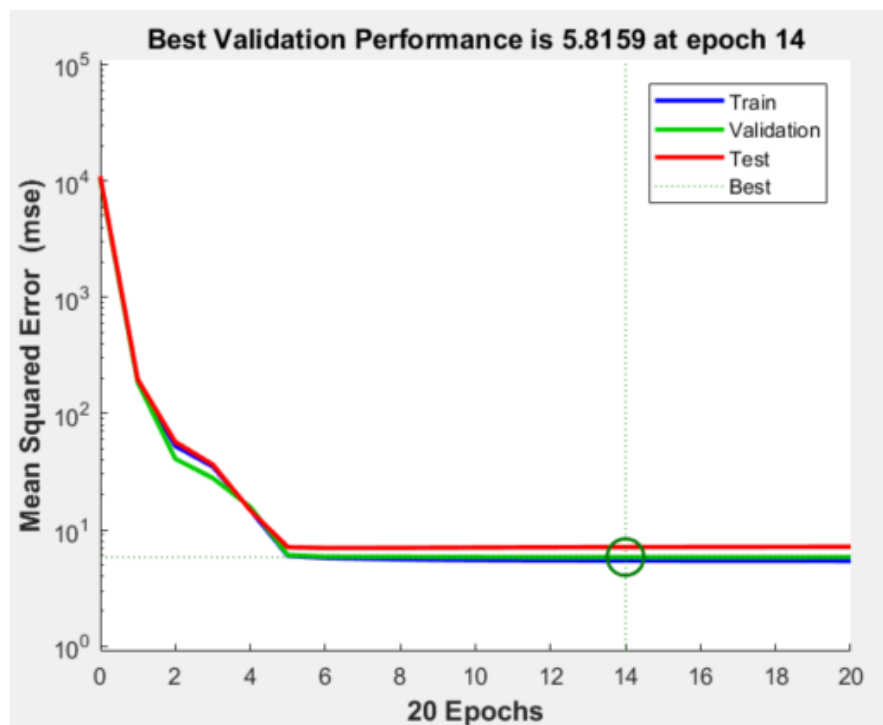


Figure 2: Machine learning model development workflow for the space weather forecast as an iterative process.

RESULTS



The performance graph shows that we attained a best validation performance of 5.8159 at epoch 14.

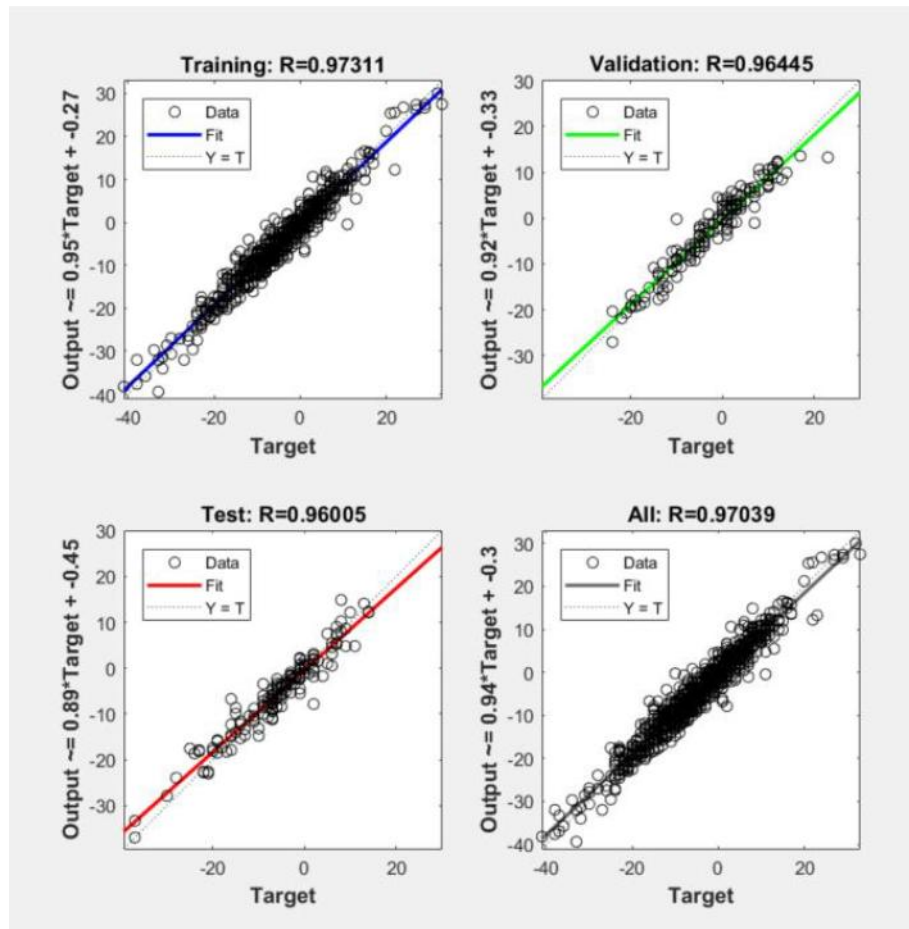


Figure 3: Regression plot when 4 features are given as input It is observed that the R value of the test when we have 4 features is 0.96005. Regression R Values measure the correlation between outputs and targets. An R value of 1 means a close relationship, 0 a random relationship. Thus, our model exhibits a very strong correlation between predicted and actual values.

CONCLUSIONS

We provided a summary of the prevailing understanding of primary space weather processes occurring on the Sun and inside the heliosphere, particularly concerning Earth. This review aims to assist forecasters of solar and heliospheric weather in comprehending the relevant phenomena, including solar flares, coronal mass ejections (CMEs) and their interplanetary equivalents, high-speed solar wind streams, and solar energetic particle (SEP) occurrences, as well as their impacts on Earth.

This Space Weather Forecasting Guide encapsulates the latest advancements in the field of space weather research. This document is expected to be updated often. In the future, we want to convert the scientific insights articulated in the text into practical methodologies to enhance operational space weather forecasting. Certain methodologies have been converted into tools used daily by the space weather forecasters at the Solar Influences Data Analysis Center at the Royal Observatory of Belgium.

REFERENCES

- [1] D. J. Knipp, J. L. Gannon, The 2019 national space weather strategy and action plan and beyond, Space Weather 17 (2019) 794–795. doi:10.1029/2019SW002254.
- [2] A. D. Richmond, Ionosphere, Springer Netherlands, Dordrecht, 2007, pp. 452–454. doi:10.1007/978-1-4020-4423-6_159.

- [3] B. Hofmann-Wellenhof, H. Lichtenegger, J. Collins, *Global Positioning System: Theory and Practice*, Springer, Berlin, 2001. doi:10.1007/978-3-7091-6199-9.
- [4] M. Poniowski, G. Nykiel, Degradation of kinematic ppp of gnss stations in central Europe caused by medium-scale traveling ionospheric disturbances during the st. patrick's day 2015 geomagnetic storm, *Remote Sensing* 12 (2020). URL: <https://www.mdpi.com/2072-4292/12/21/3582>. doi:10.3390/rs12213582.
- [5] I. Zakharenkova, I. Cherniak, Effects of storminduced equatorial plasma bubbles on gps-based kinematic positioning at equatorial and middle latitudes during the september 7–8, 2017, geomagnetic storm, *GPS Solutions* 25 (2021). doi:10.1007/s10291-021-01166-3.
- [6] L. Mallika I, D. V. Ratnam, S. Raman, G. Sivavaraprasad, Machine learning algorithm to forecast ionospheric time delays using global navigation satellite system observations, *Acta Astronautica* 173 (2020) 221–231.
- [7] T. Hastie, R. Tibshirani, J. Friedman, *The elements of statistical learning: data mining, inference and prediction*, 2 ed., Springer, New York, NY, 2009. doi:10.1007/978-0-387-84858-7_3.
- [8] J. H. King, N. E. Papitashvili, Solar wind spatial scales in and comparisons of hourly wind and ace plasma and magnetic field data, *Journal of Geophysical Research: Space Physics* 110 (2005). doi:10.1029/2004JA010649.
- [9] R. Dach, S. Schaer, D. Arnold, E. Orliac, L. Prange, A. Susnik, A. Villiger, A. Jaeggi, Code final product series for the igs (2016).
- [10] [22] A. Zheng, A. Casari, *Feature engineering for machine learning: principles and techniques for data scientists*, O'Reilly Media, Inc., 2018.