A -based space weather conditions machine learning classification model for GNSS PNT performance analysis

# Introduction

The Global Navigation Satellite System (GNSS) and its Positioning, Navigation, and Timing (PNT) service have matured to become an essential part of national infrastructure, public goods, and enablers of a vast number of emerging technology and socio-economic applications [1]. Maintaining the GNSS PNT quality is crucial for the sustainable development of modern economy and society [2]. Overcoming the shortcomings and vulnerabilities of GNSS PNT is a scientific challenge, and the need of a wide variety of scientists, developers, operators, regulators, and users of GNSS-based systems and services [3, 4]. The Earth’s ionosphere, a part of the Earth’s atmosphere stretching from to above the Earth’s surface and comprised of rare but mostly charged atoms and molecules, is the major natural cause of GNSS PNT degradation [5, 6]. This phenomenon especially affects most currently used GNSS receivers, which work as single-frequency receivers exposed to GNSS ionospheric effects [1, 7]. Driven by powerful and random flows of energy and particles from the Sun (space weather conditions), the ionospheric conditions define the properties of GNSS signal propagation through the Earth’s atmosphere and the resulting ionospheric delay [5, 8, 7]. The GNSS ionospheric delay causes errors in GNSS PNT algorithm deployment, designed to produce position, velocity, and time estimates for a GNSS receiver [1, 2]. The ionosphere affects GNSS satellite signals for position estimation by introducing signal propagation delay [1, 2]. The GNSS ionospheric delay is a stochastic variable, whose value is determined by complex physical processes of space weather [1, 2]. How space weather affects GNSS PNT performance was explained using the Space weather – GNSS PNT performance coupling model, as depicted in Figure [1](#fig:SpaceWeather).

![](data:application/pdf;base64,)

The Space weather – GNSS PNT performance coupling model.

The GNSS signal encounters a certain number of charged particles from the satellite aerial to a mobile unit’s (GNSS receiver’s) aerial [2, 7]. This encounter is quantified using the Total Electron Content () defined by Equation [[eqn:1]](#eqn:1) in [1, 2, 5], where denotes the height above the Earth’s mean sea level in , represents the vertical ionospheric profile, the volume density of charged particles at height , in . For reasons of convenience, may be expressed in Units (), with = . The dataset used in this study was derived from Receiver Independent Exchange Format (RINEX) observations using Global Positioning System (GPS) software by Seemala [9].

It should be noted that in the sense of Equation [[eqn:1]](#eqn:1), is defined as a result, a consequence, of the ionospheric conditions, and not their descriptor [1]. The GNSS ionospheric delay may be determined by derivation from the Appleton-Hartree equation, as given in Equation [[eqn:2]](#eqn:2), where denotes the GNSS ionospheric delay in , denotes the velocity of an electromagnetic wave in vacuum in , and denotes the carrier wave frequency of the satellite signal in [1, 2].

Combining Equation [[eqn:1]](#eqn:1) and Equation [[eqn:2]](#eqn:2), one can conclude the linear relation between and , as given in Equation [[eqn:3]](#eqn:3) [1].

The GNSS ionospheric delay has been identified as a source of GNSS PNT degradation since the dawn of GNSS. GNSS systems offer various standard GNSS ionospheric delay estimation (correction) models to mitigate the deteriorating effects on GNSS PNT, such as the Klobuchar model [1, 10]. The standard ionospheric correction models are global, and insufficiently flexible to update to mitigate GNSS ionospheric delay to satisfy rising demands on GNSS PNT performance [1, 11]. The development of regional and local models, such as the one presented in this paper, attempts to solve the problem of GNSS PNT sustainable performance in various ionospheric conditions. Some of the authors of this paper recently proposed the Ambient-Aware Application-Aligned ((AA)2) PNT to take into account the actual ionospheric and geomagnetic conditions near a mobile unit (a GNSS receiver) [12]. Direct measurements of the immediate geomagnetic and ionospheric condition variables may be supplied to a machine learning-based adapted GNSS ionospheric correction model, thus solving the single-frequency GNSS problem. Previous research has identified predictors and target variables (descriptors of geomagnetic, ionospheric, and GNSS PNT conditions) [13, 14]. Current space weather severity scales, such as the one provided by National Oceanic and Atmospheric Administration (NOAA), are based on global space weather and geomagnetic indices averaged over a certain period (for instance, hours for the global index). The current space weather severity scales do not directly address classifying scenarios of GNSS performance deterioration and have limited potential in deployment for a GNSS ionospheric correction model. The classification of different scenarios of GNSS ionospheric conditions with adverse effects on the GNSS PNT remained an unsolved precondition needed for the development of a machine learning-based GNSS ionospheric delay correction model to render the GNSS PNT algorithm ionospheric conditions-agnostic.

A methodology for a machine learning-based classification of ionospheric conditions based entirely on observations of geomagnetic indices is described in this study. The model is sufficiently simple to be applied on computationally capable platforms with suitable geomagnetic field sensors, such as smartphones and connected/autonomous vehicles. The research presented acquires ambient data and analyses its statistical properties. The dataset is split into training and test sets. Several candidates for the GNSS ionospheric delay model are developed based on Disturbance Storm-Time () data taken from the INTERMAGNET [15] dataset, and reformatted to match the format of data. The machine learning (ML) models include a Support Vector Machine (SVM) with a Polynomial Kernel, C5.0 Decision Tree (DT), Naive Bayes (NB), shallow Neural Network (NN), Partial Least Squares (PLS), Flexible Discriminant Analysis (FDA) and shallow Neural Network (NN) using Principal Component Analysis (PCA) of the input data. A tailored set of validation methods is used to assess their performance. The optimal GNSS ionospheric delay correction model is identified based on GNSS PNT-related objective criteria, and its performance is demonstrated in an independent case study.

# Method and data

A Support Vector Machine (SVM) with a Polynomial Kernel, C5.0 Decision Tree (DT), Naive Bayes (NB), Neural Network (NN), Partial Least Squares (PLS), Flexible Discriminant Analysis (FDA) and shallow Neural Network (NN) using Principal Component Analysis (PCA) of the input data were tested based on their ability to classify a set of observations of the geomagnetic field in , and other predictors, into one of the scenarios of space weather conditions based on . Multiple -dependent classes were predefined using theoretical knowledge. Statistical analysis of the data confirmed that distributions of other variables change for different ranges, supporting the validity of the classification. The study assumes that the dependent output variable, the class, can be predicted based on the independent variables used as input. data was obtained using GPS software by Seemala to process RINEX observations [9]. The INTERMAGNET [15] dataset contains and data from 2014 for a measuring station maintained by Geoscience Australia in Kakadu, referred to as KDU in the database, at degrees of south latitude and degrees of east longitude near Darwin, Nothern Territory, Australia. The two datasets are merged based on location, year, month, day, and time of day in hours.

## Method

The methods were selected because they represent larger families of classification methods. SVM models are supervised maximum margin models. DT models also apply supervised learning. NB classifiers are probabilistic classifiers that can be parametric or non-parametric, but this study uses a non-parametric approach. PLS is a non-parametric linear regression model. FDA uses multiple non-parametric linear regression models to create a non-linear classification. PCA is a linear dimensionality reduction technique that extracts a predefined number of components for training an NN model. NN models imitate the brain using artificial neurons to produce outputs based on the input and the activation function. NN models require that the structure be predefined, and hyperparameters are tuned. All NN models were applied based on research by Kuhn for the *R* *caret* package [16, 17, 18].

### Support Vector Machine

In machine learning, a Support Vector Machine (SVM) or Support Vector Network (SVN) model is a supervised maximum margin model with associated learning algorithms used for classification. SVM models are also effective for non-linear classification using the hyperplane kernel trick [19]. Intuitively, a good separation is achieved by the hyperplane with the greatest distance to the nearest point in the training data belonging to any class [20]. Meyer, Leisch, and Hornik compared SVM models with other classifiers [21]. However, it is unclear whether SVM predictions perform better than other linear models, such as logistic, and linear regression. To keep the computational burden reasonable, a kernel probability density function is chosen to fit the problem [22].

### Decision Tree

Decision Tree (DT) models are used for supervised learning in statistics and machine learning. Classification trees use a discrete target variable. DT models are popular due to their comprehensibility and simplicity [23]. A tree is recursively partitioned by dividing the original set, or root node, into subsets that form descendants, or successors, using classification rules based on features [24]. C5.0, used in the *caret* package in *R*, has a similar approach and improves the ID3 and C4.5 algorithms.

### Naive Bayes

In statistics, Naive Bayes (NB) models, simple Bayes, or independent Bayes [25] classifiers are a family of linear "probabilistic classifiers" that assume that, given a target class, the features are conditionally independent. Maximum likelihood training for NB models evaluates a closed-form expression [26] in linear time instead of using iterative approximation. However, a comprehensive comparison in 2006 showed that NB models performed worse than boosted trees or Random Forest (RF) models [27]. An advantage of NB over other models is a smaller amount of required training data [28]. NB models assign probabilities to classes for an input vector with features [29], and use Bayes’ theorem in Equation [[eqn:4]](#eqn:4).

### Neural Networks

The neurons of human or animal brains provide the basis for a Neural Network (NN) or Artificial Neural Network (ANN) with connected units or nodes called artificial neurons in machine learning [30]. Shallow NN models typically contain only a few hidden layers for processing between the input layer that receives the data and the final layer that produces the output [31]. A network with at least two hidden layers [32] is a deep NN model. Gradient-based methods such as backpropagation estimate ANN parameters [33] to minimize the difference or empirical risk between the output and target labels, expressed in a loss function [34]. The hyperparameters may also be modified to suit the problem [35] during an extensive tuning process, like the one used in this study. Principal Component Analysis (PCA) [36] is a linear dimensionality reduction technique in exploratory data analysis, visualization [37], and preprocessing. The method in the *caret* package in *R* uses PCA in preprocessing [16, 17, 18].

### Partial Least Squares

Partial Least Squares (PLS) regression, or projection to latent structures, [38], is a linear regression statistical model that transforms the predicted and the observable variables to a new space. PLS methods are bilinear factor models because the and are projected to new spaces. In Partial Least Squares Discriminant Analysis (PLS-DA), is categorical [39]. Using paired observations . PLS finds the normalized direction that maximizes the covariance in the first step , shown in Equation [[eqn:5]](#eqn:5). Many versions of PLS exist for estimating the factor and loading matrices, such as the PLS1 algorithm [40].

### Flexible Discriminant Analysis

Flexible Discriminant Analysis (FDA) is a general methodology that creates the discriminant surface for a multigroup non-linear classification model [41] based on a mixture of non-parametric linear regression models, such as Multivariate Adaptive Regression Splines (MARS) and Linear Discriminant Analysis (LDA). Many predictors can be used in conjunction in FDA [42]. FDA is complex but execution time and computational load are adequate [43]. Feature normality and equal group covariances are assumed [44]. LDA, Normal Discriminant Analysis (NDA), or discriminant function analysis [45] is a generalization of Fisher’s linear discriminant defined in 1936 [41]. The results of LDA may be utilized directly for classification, as demonstrated in this experiment.

## Data description and analysis

Dynamic space weather conditions, such as solar activity and geomagnetic storms, can affect GNSS PNT performance and high-frequency GPS signals passing through the ionosphere, motivating work on error modeling [46]. Geomagnetic storms cause signal deterioration by affecting Global Electric Current (GEC) variability. The ionosphere may show changes related to location, geomagnetic and solar activity, sunspots, local time, seasonality, thunderstorms [47], nuclear experiments, earthquakes [48], and other phenomena. This study focuses on parameters describing disturbances of the Earth’s geomagnetic field, most importantly -indices derived from -indices, which are calculated using -indices, Total Electron Content (), standard deviation of Total Electron Content (), and Disturbance Storm-Time (). Incorporating parameters such as the -indices and -indices, which provide global measures of geomagnetic activity, alongside local and -index values, allows for a more detailed assessment of the space environment and its potential effects on GNSS signals. Values of , , , were used with the , , and components of the Earth’s magnetic field to train machine-learning models.

### Magnetic field indices

The Earth’s magnetic field has similarities to that of a bar magnet. However, plasma gushes from the solar corona and the domain of the Sun influence the interplanetary magnetic field [49]. The , , and vectors represent interplanetary magnetic field indices. and are parallel to the plane of orbits, and the third component is perpendicular. Widely available hand-held devices, such as Android smartphones [50], measure magnetic field indices in micro-Tesla (). The Android magnetometer reports accuracy through a status variable. Readings are calibrated using temperature compensation, factory (or online) soft-iron, and online hard-iron calibration.

### Geomagnetic storm indices

The geomagnetic storm -index is an integer from to measuring disturbances in global geomagnetic activity. The maximum positive and negative fluctuations of the horizontal components of the Earth’s magnetic field, and , during hours, relative to a quiet day, are added to determine the total maximum fluctuation. Each observatory uses different threshold values to convert the maximum (nano-Tesla) fluctuation to a -index value. The thresholds for each observatory are adjusted so that the historical rate of occurrence for each -index value is similar across all observatories. Observatories with a lower geomagnetic latitude use a lower fluctuation in and to achieve each -index value.

### Planetary geomagnetic storm indices

The planetary geomagnetic storm -index is derived from -hour-based -indices from magnetometer stations between and degrees of north and south latitude. Announcements and warnings of geomagnetic changes and disturbances in the Earth’s magnetic field are based on the -index. Hourly geomagnetic storm index data are available on the NASA Goddard SPDF web pages. The scale values of the -index are determined by the change of the geomagnetic field and the geomagnetic storm effect in . The official planetary -index is a weighted average of -indices from multiple observatories. When -index data is not available in real-time, operators such as The National Oceanic and Atmospheric Administration (NOAA) Space Weather Prediction Center (SWPC) calculate near real-time estimates of the -index [51]. The -index is related to geomagnetic storm descriptions and warnings using the NOAA G scale. For a -index value of , , , (including ), and , a NOAA Space Weather Scale Geomagnetic Storm Level of G1, G2, G3, G4, and G5 is assigned, respectively. For a -index , the G0 designation means no warning is issued. In March 2021, was assigned a Digital Object Identifier (DOI) with a dataset [52] and a scientific publication [53] for reference.

### Equivalent three hourly range geomagnetic storm indices

The -index represents a daily average level of magnetic activity. Because the relationship between the -index and magnetometer fluctuations is not linear, the -index values are not directly used for calculating average values. Each -index or -index is converted into the "equivalent three hourly range" -index or -index that uses a linear scale. An average of -indices (lowercase) is used as the daily -index (uppercase).

### Disturbance Storm-Time

Disturbance Storm-Time (), also known as the geomagnetic activity -index, depicts an averaged measure of the geomagnetic storm intensity in the Earth’s sub-equatorial region. The -index is obtained by post-processing, with the final version often published months after experimental observations were collected. It clearly shows developments of various levels of geomagnetic disturbances, and consequently serves as invaluable input considering GNSS PNT performance degradation due to space weather events and disturbances. The -index is a geomagnetic indicator of magnetic flux changes derived from measurements taken by a network of ground-based magnetometer stations near the magnetic equator, which continuously monitor and , the horizontal components of Earth’s magnetic field [46]. The -index describes ring currents forming above the sub-equatorial region and affecting the ionospheric regions in mid-latitudes. To calculate the -index, variations in and , the horizontal magnetic field, are obtained from multiple stations and averaged. The average is subtracted from a baseline value representing the quiet-time magnetic field. The resulting value in measures the intensity of geomagnetic disturbances, with increasingly negative values indicating stronger geomagnetic storms. The -indice measurements as an hourly average were evaluated and published on a web interface by the NASA Goddard SPDF, and the Geomagnetism and Space Magnetism Data Analysis Center of the Institute of Science, Kyoto University in Japan. Loewe and Prölss [54] classified magnetic activity -indices into storm classes in 1997. Gonzalez et al. [55] used groups for the same data in 1994, similar to Kamide et al. in 1998 [56], Rozhnoi et al. in 2004 [57], and Contadakis et al. in 2012 [58].

### Training and testing dataset

The original set of observations has been split into training and testing subsets, considering the results of exploratory statistical, and outlier analysis, and sustaining proportions of data relating to classes of geomagnetic disturbances (quiet/normal geomagnetic conditions, the positive phase of a geomagnetic storm, deep negative depression of a geomagnetic storm, and the negative recovery phase of a geomagnetic storm). Classes are based on and differential time series. Out of original samples, were marked as outliers in exploratory statistical analysis, as the is larger than or equal to . values larger than usually appear due to measurement errors or errors in the estimation process. The remaining samples are used for training and testing. The samples are divided into training and testing datasets as close as possible to a ratio of for training and for testing. The division was stratified so that an approximately equal ratio of classes was present in both the training and testing data, which is a feature of the *createDataPartition* function from the *caret* *R* library that was used [16, 17]. Samples were split into class ranges, based on values derived from theoretical knowledge of different storm phases, similar to Loewe and Prölss [54]. Table [1](#tab:Dstranges) lists the class ranges used in this study, the total number of samples in each class, and the number of samples used for testing, and training.

-based classification rules used in this study, the total number of samples in each class, and the number of samples used for testing, and training. The upper range limits are excluded, while the lower ones are included.

|  | Storm phase classification | Total samples | Test samples | Train samples |
| --- | --- | --- | --- | --- |
|  | positive phase (P) |  |  |  |
|  | normal (N) |  |  |  |
|  | recovery phase (R) |  |  |  |
|  | through (T) |  |  |  |
|  | extreme (E) |  |  |  |
| Any | Any |  |  |  |

It is evident from Table [1](#tab:Dstranges) that the normal (N), and recovery (R) classes are more common than those with very high or low values, impacting model performance.

### Data preprocessing

Data preprocessing can increase classification accuracy [59]. There are many ways to standardize data, such as minimum-maximum, normalization by decimal scaling, and Z-score [60]. Subtracting the mean and dividing by the variance for each feature are commonly used for SVM models [61] and other models tested in this study, so this approach was chosen. The values *scale* and *center* were used in this study in the *preProcess* parameter for the *train* function from the *R* *caret* package. The option *center* subtracts the mean of each feature while *scale* divides by the standard deviation.

### Distribution analysis for predictors

Table [2](#tab:minmax) provides the minimum, quartile, median, arithmetic mean, quartile, and maximum values for all variables when the is less than , suggesting that variables are not normally distributed.

The minimum, quartile, median, arithmetic mean, quartile, and maximum values for all variables when the is less than .

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Min. |  |  |  |  |  |  |  |
| Qu. |  |  |  |  |  |  |  |
| Median |  |  |  |  |  |  |  |
| Mean |  |  |  |  |  |  |  |
| Qu. |  |  |  |  |  |  |  |
| Max. |  |  |  |  |  |  |  |

The Kolmogorov-Smirnov and Shapiro-Wilk normality tests, using *R* functions *ks.test* and *shapiro.test*, did not yield a -value larger than the selected -value of for any variable, further strengthening the claim based on Table [2](#tab:minmax) that variables do not follow a normal (Gaussian) distribution.

### Correlation analysis for predictors

Observations of statistical variables were assessed for their mutual association/correlation to identify the classification model structure, potential predictors, and targets. Figure [2](#fig:correlation) contains a heat map of the correlation between all variables used in the study.

![](data:application/pdf;base64,)

A heat map of the correlation between all variables used in this study, when the is less than . Red represents a high positive correlation, blue represents a high negative correlation, and white represents a low correlation. Variables are fully correlated with themselves, so values on the secondary diagonal equal . The matrix is symmetrical concerning the secondary diagonal because the same combination of correlated variables is achieved when swapping the row and column indices.

The and variables exhibit the largest correlation coefficient in Figure [2](#fig:correlation), equaling . The association between the two variables seems logical since the index is defined as a result of geomagnetic field conditions described by geomagnetic field components. The second largest correlation coefficient depicts a significant correlation between and and equals , thus confirming that the variable can be used as a predictor, with the variable as a target, and an opposite trend.

The box plots of all variables for different ranges of values in Figure [3](#fig:iono3boxplot) are used to support the correlations shown in Figure [2](#fig:correlation) by exhibiting the trend of each variable.

![](data:application/pdf;base64,)

Box plots of all variables, when the is less than , for different ranges of values defining the class label used in this study.

Figure [3](#fig:iono3boxplot) shows the minimum, maximum, and arithmetic mean of decreasing for larger values. The opposite is true for , as indicated by a high correlation in Figure [2](#fig:correlation). exhibits a reverse trend compared to , but it is less prominent. is the most stable among geomagnetic indices, with the smallest changes related to .

## Confusion Matrix

The metrics and terminology defined in the *R* function *confusionMatrix* in the *caret* library [16, 17] were used to evaluate classifier performance. The default approach to a confusion matrix uses only two groups (Yes and No, positive and negative). For multiple classes, results are calculated by a "one versus all" approach, viewing each class as positive and all others as negative.

### McNemar-Bowker test

McNemar’s test for correlated proportions used on paired categorical data is based on the chi-squared distribution and was originally designed for methods that differentiate between two classes. The null hypothesis of marginal homogeneity states that marginal probabilities for each outcome are equal, which is more indicative of model difference than directly comparing the sensitivity and specificity of two candidate models.

However, for groups, the classification can be annotated in a contingency table with the number of samples classified using the first method in separate rows by class, divided into columns depending on the class assigned by the opposing method. The McNemar-Bowker test is applied to this table, formulated by Fagerland et al. [62], and Chow et al. [63].

If any element in the matrix is smaller than , as shown using class sizes in Table [1](#tab:Dstranges), the distribution is not well-approximated by the chi-squared distribution. Edwards [64] developed an approximation of the binomial exact -value for continuity-correction, given in Equation [[eqn:6]](#eqn:6) for two groups. In Equation [[eqn:6]](#eqn:6), is the number of samples classified in the first group by the first test, and the second group by the second test, while is the number of samples for which the opposite is true.

# Research results

Candidate models were assessed to determine the optimal method and set of predictors to be used in the final model for application in real settings. The accuracy and the execution time in seconds () utilizing the *R* *system.time* function for all candidate models are displayed in Table [3](#tab:acc:time). Research results are presented with the following initial set of six predictors: Total Electron Content (), standard deviation of Total Electron Content (), , , , and . was not used as a predictor since the initial classification is derived from values. Candidate models have also been developed using: (1.): all predictors except , and (, , , and ), (2.): Geomagnetic indices (, , and ), (3.): , , and , (4.): , , and , and (5.): , , and . The reasoning behind this approach was the assumption that reducing the set of predictors would reduce model complexity and computation time. Additionally, it is theoretically established that geomagnetic indices , , and affect , and in turn . This is why geomagnetic indices (, , and ), and were used as predictors in the final model. The highest accuracy in Table [3](#tab:acc:time) was achieved using , , , and as predictors, supporting the hypothesis that , and should be removed from the set of predictors. The Naive Bayes (NB) method yielded the model with the highest accuracy, leading to the decision that it should be used in the final model.

The accuracy (top) and the execution time in seconds () (bottom) for each candidate model developed using different methods and sets of predictors.

| Candidate | Accuracy | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| 2-7 model | Predictors | | | | | |
| Method | All | , , , | , , |  |  |  |
| SVM Poly |  |  |  |  |  |  |
| C5.0 DT |  |  |  |  |  |  |
| NB |  |  |  |  |  |  |
| NN |  |  |  |  |  |  |
| PLS |  |  |  |  |  |  |
| FDA |  |  |  |  |  |  |
| PCA NN |  |  |  |  |  |  |
| Candidate | Execution time in seconds () | | | | | |
| 2-7 model | Predictors | | | | | |
| Method | All | , , , | , , |  |  |  |
| SVM Poly |  |  |  |  |  |  |
| C5.0 DT |  |  |  |  |  |  |
| NB |  |  |  |  |  |  |
| NN |  |  |  |  |  |  |
| PLS |  |  |  |  |  |  |
| FDA |  |  |  |  |  |  |
| PCA NN |  |  |  |  |  |  |

The experiment was run on *Windows* 11 using *R Studio* version 2024.04.2+764 and *R* version 4.4.1, the AMD Radeon RX 6600 Graphics Processing Unit (GPU), GB of Random Access Memory (RAM), and the AMD Ryzen 5 PRO 4650G Central Processing Unit (CPU) with cores. Execution time is significant because built-in systems for mobile devices using GNSS PNT have low computational capabilities. The candidate models using the Partial Least Squares (PLS) method in Table [3](#tab:acc:time) have the lowest execution time.

The performance assessment of candidate models revealed the success of several methods, such as the Neural Network (NN) and Naive Bayes (NB) classifiers. Accuracy results in Table [3](#tab:acc:time) may lead to an incorrect conclusion as the dataset is unbalanced, as shown in Table [1](#tab:Dstranges). Further assessment is needed to determine if a candidate model differs significantly from others. The results of McNemar’s test are presented only for candidate models using either the Naive Bayes (NB) method or , , , and as predictors variables, as this combination presented as the most accurate among tested candidate models and has a lower computational load than competing models.

## Results of McNemar’s test

McNemar’s test can help assess if a candidate model performs significantly worse or better than others. Figure [4](#fig:pvalueplot) contains -values of McNemar’s test when comparing candidate models using the Naive Bayes (NB) method and various predictor variable sets, or , , , and as predictor variables combined with various classification methods. A lower -value indicates that the classifications are unequal for a pair of candidate models. A higher -value suggests that the classifications are equal for a pair of candidate models.

![](data:application/pdf;base64,)

Candidate model comparison using -values of McNemar’s test, and , , , and as predictor variables (excluding Total Electron Content (), and standard deviation of Total Electron Content ()) combined with various classification methods (left), or the Naive Bayes (NB) method combined with various predictor variable sets (right). Black represents a low -value near , white represents a high -value near , and red represents a -value near between the two extremes. Each candidate model is equal to itself, so values on the primary diagonal equal . The matrix is symmetrical concerning the primary diagonal because the same result is achieved when swapping the order of the first and second compared candidate models.

The classifications are not different when comparing the candidate models using the NB method, or the candidate models using , , , and as predictor variables, as shown in Figure [4](#fig:pvalueplot). The -values of McNemar’s test larger than indicate that any of the compared models might yield similar results and could be used to implement a geomagnetic/ionospheric storm warning system, even though the NB method and , , , and predictor variables yield the best accuracy.

# Discussion

The Support Vector Machine (SVM) method with a Polynomial Kernel, C5.0 Decision Tree (DT), Neural Network (NN), and NN method with Principal Component Analysis (PCA) applied in preprocessing have the highest execution time, above , so they were excluded from further analysis and application in the final model.

The Neural Network (NN) method has a high execution time for any set of predictors due to its complexity and extensive training, evident from the data in Table [3](#tab:acc:time). However, it achieved a accuracy, also shown in Table [3](#tab:acc:time). The Naive Bayes (NB) method has the same accuracy when using , , , and as predictors, and the training time for the NN model is more than twice as long, so the NB method is preferred as the approach in the final model.

An analysis in 2004 showed reasonable theoretical reasons for the seemingly incredible performance of NB classifiers [65]. The expression used for the NB method is a linear time algorithm if time complexity is expressed as a function of the size of the input and observing asymptotic behavior, explaining the reduced execution time [26]. Despite their simplicity, NB classifier models have performed well in real-world situations [66], even with a relative lack of data compared to other approaches [28, 67].

All candidate models developed using methods other than NN and NB fail to achieve accuracy over . This indicates that they are less suitable for this particular application. The SVM method with a Polynomial Kernel is consistently the worst-performing for any set of predictors, never achieving an accuracy over . All models achieved an accuracy over , suggesting they could indicate Global Navigation Satellite System (GNSS) Positioning, Navigation, and Timing (PNT) performance.

# Conclusion

The presented study aims to classify ambient conditions of space weather events for sub-equatorial regions. Global Navigation Satellite System (GNSS) Positioning, Navigation, and Timing (PNT) performance is significantly affected by such events. It would be beneficial to warn users of a geomagnetic/ionospheric storm.

Classification models using machine learning were applied to descriptions of the geomagnetic field expressed in Total Electron Content (), standard deviation of Total Electron Content (), , , (geomagnetic field indices), and . It was assumed observations contained independent variables to generate the dependent variable representing the Disturbance Storm-Time () class.

Statistical analysis confirmed that other variables change distribution based on , not . Continuous values in different ranges were converted into discrete classes based on statistics, previous theories, and research.

An Support Vector Machine (SVM) with a Polynomial Kernel, C5.0 Decision Tree (DT), Naive Bayes (NB), Neural Network (NN), Partial Least Squares (PLS), Flexible Discriminant Analysis (FDA), and Principal Component Analysis (PCA) NN model created a -based classification from multiple combinations of input variables.

The NB method using , , , and as predictors achieved perfect accuracy on the test set. The total execution time is at least two times shorter than for the NN method, making it more suitable for compact, low-performance, and low-cost portable devices such as smartphones. The exclusion of and is supported by theory since they contain redundant information already presented by , , and which impact , and in turn .

# Acronyms

*(AA)2* Ambient-Aware Application-Aligned

*ANN* Artificial Neural Network

*BA* Balanced Accuracy

*CCA* Canonical Correlation Analysis

*CI* Confidence Interval

*CNN* Convolutional Neural Network

*CPU* Central Processing Unit

*DBSCAN* Density-Based Spatial Clustering

*DOI* Digital Object Identifier

*DP* Detection Prevalence

*DR* Detection Rate

*DT* Decision Tree

*Dst* Disturbance Storm-Time

*EOF* Empirical Orthogonal Functions

*EPB* Equatorial Plasma Bubbles

*EUV* Extreme ultraviolet

*EVD* Eigenvalue Decomposition

*FDA* Flexible Discriminant Analysis

*FN* False Negative

*FP* False Positive

*GEC* Global Electric Current

*GFZ* German Research Centre for Geosciences

*GNSS* Global Navigation Satellite System

*GPS* Global Positioning System

*GPU* Graphics Processing Unit

*ION* The Institute of Navigation

*KLT* Karhunen–Loève Theorem

*LDA* Linear Discriminant Analysis

*MANOVA* Multivariate Analysis Of Variance

*MARS* Multivariate Adaptive Regression Splines

*ML* machine learning

*NAS* Neural Architecture Search

*NASA* National Aeronautics and Space Administration

*NB* Naive Bayes

*NDA* Normal Discriminant Analysis

*NIR* No Information Rate

*NN* Neural Network

*NOAA* National Oceanic and Atmospheric Administration

*NPV* Negative Predictive Value

*NRCAN* Natural Resources Canada

*PCA* Principal Component Analysis

*PLS* Partial Least Squares

*PLS-DA* Partial Least Squares Discriminant Analysis

*PNT* Positioning, Navigation, and Timing

*POD* Proper Orthogonal Decomposition

*PPV* Positive Predictive Value

*RAM* Random Access Memory

*RF* Random Forest

*RIN* The Royal Institute of Navigation

*RINEX* Receiver Independent Exchange Format

*SAR* Synthetic Aperture Radar

*SPDF* Space Physics Data Facility

*SVD* Singular Value Decomposition

*SVM* Support Vector Machine

*SVN* Support Vector Network

*SWPC* Space Weather Prediction Center

*TDIDT* Top-Down Induction of Decision Trees

*TEC* Total Electron Content

*TID* Traveling Ionospheric Disturbance

*TN* True Negative

*TNR* True Negative Rate

*TP* True Positive

*TPR* True Positive Rate

*URSI* Union Radio-Scientifique Internationale

*dTEC* standard deviation of Total Electron Content

# Declarations

## Availability of data and materials

The datasets used during the current study are available from the corresponding author upon reasonable request.

## Competing interests

The authors declare no conflict of interest.

## Funding

The authors have no funding sources to declare.

## Authors’ contributions

LŽ contributed to conceptualization, methodology, software, validation, formal analysis, investigation, data curation, original draft writing, text review and editing, and visualization. DK contributed to conceptualization, methodology, investigation, original draft writing, and text review and editing. TI contributed to validation, formal analysis, investigation, text review and editing, supervision, and project administration. RF contributed to conceptualization, methodology, investigation, resources, original draft writing, text review and editing, and supervision. All authors read and approved the final manuscript.

## Acknowledgements

Not applicable

## Authors’ information

LŽ is a doctoral student at the Faculty of Engineering, University of Rijeka, currently employed as an assistant at the Department of Computer Science. Her research interests include applied machine learning in biology, chemistry, medicine, and transportation. Maritime transportation is also the subject of her doctoral research.

DK is a doctoral student at the Faculty of Engineering, University of Rijeka, currently employed as an assistant at the Department of Computer Science. His research interests include applied machine learning in object recognition and image segmentation. Image classification and labeling are also the subject of his doctoral research.

TI earned his MSc and PhD in telecommunication engineering from the University of Ruse “Angel Kanchev,” Ruse, Bulgaria, in 1999 and 2007, respectively. He is a full professor at the Department of Telecommunication, University of Ruse “Angel Kanchev.” His research interests include mobile communications networks, signal processing, wireless technologies, and satellite navigation. Professor Iliev is a Member of IEEE.

RF is an external Professor of Electronics Engineering with the Department for Computer Science, Faculty of Engineering, and Center for Artificial Intelligence and Cybersecurity, both at the University of Rijeka, Croatia, and Head of the Laboratory for Spatial Intelligence at Hrvatsko Zagorje Krapina University of Applied Science, Krapina, Croatia. He holds a BSc, MSc, and PhD in electrical engineering, obtained in 1987, 1994, and 2007, respectively, from the Faculty of Electrical Engineering and Computing, University of Zagreb, Croatia. His professional interests include Ambient-Aware Application-Aligned ((AA)2) Global Navigation Satellite System (GNSS) Positioning, Navigation, and Timing (PNT), spatial uncertainty quantification, spatial statistical learning, predictive modeling, statistical signal processing, trajectory analysis and prediction, and occupancy modeling. Professor Filjar is a Fellow of The Royal Institute of Navigation (RIN) (London, United Kingdom), a Member of The Institute of Navigation (ION) (Manassas, Virginia), a Senior Member of Union Radio-Scientifique Internationale (URSI) (Ghent, Belgium) and a Member of The Society for Industrial and Applied Mathematics (Philadelphia, Philadelphia).

# Bibliography

[1] J. J. Spilker Jr, P. Axelrad, B. W. Parkinson, and P. Enge, Global positioning system: theory and applications, volume I. Reston, Virginia: American Institute of Aeronautics, 1996.

[2] S. Schaer, Mapping and predicting the Earth’s ionosphere using the Global Positioning System, vol. 59. Zürich: Institut für Geodäsie und Photogrammetrie, Eidg. Technische Hochschule, 1999.

[3] M. Durmaz and M. O. Karslioglu, “Regional vertical total electron content (VTEC) modeling together with satellite and receiver differential code biases (DCBs) using semi-parametric multivariate adaptive regression B-splines (SP-BMARS),” Journal of Geodesy, vol. 89, pp. 347–360, 2015, doi: 10.1007/s00190-014-0779-8.

[4] R. Jin, S. Jin, and G. Feng, “M\_DCB: Matlab code for estimating GNSS satellite and receiver differential code biases,” GPS solutions, vol. 16, no. 4, pp. 541–548, 2012, doi: 10.1007/s10291-012-0279-3.

[5] K. Davies, Ionospheric Radio. Futures Place, Stevenage: Institution of Engineering & Technology, 1990. [Online] . Available: https://books.google.hr/books?id=qdWUKSj5PCcC

[6] J. Y. Liu et al., “Seismoionospheric GPS total electron content anomalies observed before the 12 May 2008 Mw7. 9 Wenchuan earthquake,” Journal of Geophysical Research: Space Physics, vol. 114, no. A4, 2009, doi: 10.1029/2008JA013698.

[7] G. Prölss, Physics of the Earth’s space environment: an introduction. Berlin/Heidelberg, Germany: Springer Science & Business Media, 2012.

[8] A. Oxley, Uncertainties in GPS Positioning: A mathematical discourse. Cambridge, Massachusetts: Academic Press, 2017.

[9] G. K. Seemala, “Chapter 4 - Estimation of ionospheric total electron content (TEC) from GNSS observations,” in Atmospheric Remote Sensing, A. K. Singh and S. Tiwari, Eds. Amsterdam, Netherlands: Elsevier, 2023, pp. 63–84. doi: 10.1016/B978-0-323-99262-6.00022-5.

[10] J. A. Klobuchar, “Ionospheric time-delay algorithm for single-frequency GPS users,” IEEE Transactions on aerospace and electronic systems, pp. 325–331, 1987, doi: 10.1109/TAES.1987.310829.

[11] P. K. Enge, “The global positioning system: Signals, measurements, and performance,” International Journal of Wireless Information Networks, vol. 1, pp. 83–105, 1994, doi: 10.1007/BF02106512.

[12] R. Filjar, I. Hedji, J. Prpić-Oršić, and T. Iliev, “An Ambient Adaptive Global Navigation Satellite System Total Electron Content Predictive Model for Short-Term Rapid Geomagnetic Storm Events,” Remote Sensing, vol. 16, no. 16, p. 3051, 2024, doi: 10.3390/rs16163051.

[13] R. Natras, B. Soja, and M. Schmidt, “Ensemble machine learning of random forest, AdaBoost and XGBoost for vertical total electron content forecasting,” Remote Sensing, vol. 14, no. 15, p. 3547, 2022, doi: 10.3390/rs14153547.

[14] R. Natras et al., “Regional ionosphere delay models based on CORS data and machine learning,” NAVIGATION: Journal of the Institute of Navigation, vol. 70, no. 3, 2023, doi: 10.33012/navi.577.

[15] INTERMAGNET and others, “Intermagnet reference data set (IRDS) 2019 – definitive magnetic observatory data,” GFZ Data Services, 2022, doi: 10.5880/INTERMAGNET.1991.2019.

[16] M. Kuhn, “caret: Classification and Regression Training,” CRAN: Contributed Packages, Oct. 2007, doi: 10.32614/cran.package.caret.

[17] M. Kuhn, “Building predictive models in R using the caret package,” Journal of statistical software, vol. 28, pp. 1–26, 2008, doi: 10.18637/jss.v028.i05.

[18] M. Kuhn, Applied predictive modeling. Berlin/Heidelberg, Germany: Springer, 2013.

[19] B. E. Boser, I. M. Guyon, and V. N. Vapnik, “A training algorithm for optimal margin classifiers,” in Proceedings of the fifth annual workshop on Computational learning theory, Jul. 1992, pp. 144–152. doi: 10.1145/130385.130401.

[20] T. Hastie, S. Rosset, J. Zhu, and H. Zou, “Multi-class AdaBoost,” Stat. Interface, vol. 2, no. 3, pp. 349–360, 2009, doi: 10.4310/SII.2009.v2.n3.a8.

[21] D. Meyer, F. Leisch, and K. Hornik, “The support vector machine under test,” Neurocomputing, vol. 55, no. 1–2, pp. 169–186, Sep. 2003, doi: 10.1016/S0925-2312(03)00431-4.

[22] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, Numerical recipes 3rd edition, Third. Cambridge, England: Cambridge University Press, 2007.

[23] X. Wu et al., “Top 10 algorithms in data mining,” Knowl. Inf. Syst., vol. 14, no. 1, pp. 1–37, Jan. 2008, doi: 10.1007/s10115-007-0114-2.

[24] S. Shalev-Shwartz and S. Ben-David, “Decision Trees,” in Understanding Machine Learning: From Theory to Algorithms, Cambridge, England: Cambridge University Press, 2014, pp. 212–218. doi: 10.1017/CBO9781107298019.019.

[25] D. J. Hand and K. Yu, “Idiot’s Bayes: Not So Stupid after All?,” Int. Stat. Rev., vol. 69, no. 3, p. 385, Dec. 2001, doi: 10.1111/j.1751-5823.2001.tb00465.x.

[26] S. J. Russell and P. Norvig, Artificial intelligence: a modern approach. London, England: Pearson, 2016.

[27] R. Caruana and A. Niculescu-Mizil, “An Empirical Comparison of Supervised Learning Algorithms,” Proceedings of the 23rd international conference on Machine learning - ICML ’06, vol. 2006, pp. 161–168, Jun. 2006, doi: 10.1145/1143844.1143865.

[28] G. H. John and P. Langley, “Estimating Continuous Distributions in Bayesian Classifiers,” arXiv preprint arXiv:1302.4964, 2013, doi: 10.48550/arXiv.1302.4964.

[29] M. N. Murty and V. S. Devi, Pattern recognition: An algorithmic approach. Berlin/Heidelberg, Germany: Springer Science & Business Media, 2011.

[30] A. Brahme, Comprehensive biomedical physics. 8-11 Southampton Street, London: Newnes, 2014.

[31] J. D. Olden and D. A. Jackson, “Illuminating the ‘black box’: a randomization approach for understanding variable contributions in artificial neural networks,” Ecological modelling, vol. 154, no. 1–2, pp. 135–150, 2002, doi: 10.1016/S0304-3800(02)00064-9.

[32] C. Bishop, Pattern Recognition and Machine Learning. Berlin/Heidelberg, Germany: Springer, 2006.

[33] V. Vapnik, The nature of statistical learning theory. Berlin/Heidelberg, Germany: Springer science & business media, 2013.

[34] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, Massachusetts: MIT Press, 2016.

[35] P. Probst, A. L. Boulesteix, and B. Bischl, “Tunability: Importance of hyperparameters of machine learning algorithms,” Journal of Machine Learning Research, vol. 20, no. 53, pp. 1–32, 2019, doi: 10.48550/arXiv.1802.09596.

[36] F. M. Stewart, Introduction to linear algebra. Mineola, New York, United States of America: Courier Dover Publications, 2019.

[37] I. T. Jolliffe and J. Cadima, “Principal component analysis: a review and recent developments,” Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences, vol. 374, no. 2065, p. 20150202, 2016, doi: 10.1098/rsta.2015.0202.

[38] H. Abdi, “Partial least squares regression and projection on latent structure regression (PLS Regression),” Wiley interdisciplinary reviews: computational statistics, vol. 2, no. 1, pp. 97–106, 2010, doi: 10.1002/wics.51.

[39] S. Sæbø, T. Almøy, A. Flatberg, A. H. Aastveit, and H. Martens, “LPLS-regression: a method for prediction and classification under the influence of background information on predictor variables,” Chemometrics and Intelligent Laboratory Systems, vol. 91, no. 2, pp. 121–132, 2008, doi: 10.1016/j.chemolab.2007.10.006.

[40] V. González, R. Giraldo, and V. Leiva, “PLS1-MD: A partial least squares regression algorithm for solving missing data problems,” Chemometrics and Intelligent Laboratory Systems, vol. 240, p. 104876, 2023, doi: 10.1016/j.chemolab.2023.104876.

[41] G. J. McLachlan, Discriminant analysis and statistical pattern recognition. Hoboken, New Jersey, United States of America: John Wiley & Sons, 2005. doi: 10.1002/0471725293.

[42] T. Hastie, R. Tibshirani, J. H. Friedman, and J. H. Friedman, The elements of statistical learning: data mining, inference, and prediction, vol. 2. Berlin/Heidelberg, Germany: Springer, 2009.

[43] C. Reynès, R. Sabatier, and N. Molinari, “Choice of B-splines with free parameters in the flexible discriminant analysis context,” Computational statistics & data analysis, vol. 51, no. 3, pp. 1765–1778, 2006, doi: 10.1016/j.csda.2005.11.018.

[44] D. Wetcher-Hendricks, Analyzing quantitative data: An introduction for social researchers. Hoboken, New Jersey, United States of America: John Wiley & Sons, 2011.

[45] J. Cohen, P. Cohen, S. G. West, and L. S. Aiken, Applied multiple regression/correlation analysis for the behavioral sciences. Milton Park, Abingdon-on-Thames, Oxfordshire, England, United Kingdom: Routledge, 2013.

[46] B. Zolesi and Lj. R. Cander, Ionospheric prediction and forecasting. Berlin/Heidelberg, Germany: Springer, 2014.

[47] P. I. Vellinov, C. W. Spassov, and S. I. Kolev, “Ionospheric effects of lightning during the increasing part of solar cycle 22,” Journal of atmospheric and terrestrial physics, vol. 54, no. 10, pp. 1347–1353, 1992, doi: 10.1016/0021-9169(92)90044-L.

[48] M. Liu, G. Luo, and H. Wang, “The 2013 Lushan earthquake in China tests hazard assessments,” Seismological Research Letters, vol. 85, no. 1, pp. 40–43, 2014, doi: 10.1785/0220130117.

[49] R. Schwenn, “Solar wind: Global properties,” in Encyclopedia of Astronomy & Astrophysics, Boca Raton, Florida, United States of America: CRC Press, 2001, pp. 1–9. doi: 10.1888/0333750888/2301.

[50] H. Bojinov, Y. Michalevsky, G. Nakibly, and D. Boneh, “Mobile Device Identification via Sensor Fingerprinting,” arXiv preprint arXiv:1408.1416, 2014, doi: 10.48550/arxiv.1408.1416.

[51] L. M. M. Myint, K. Hozumi, S. Saito, and P. Supnithi, “Analysis of local geomagnetic index under the influence of equatorial electrojet (EEJ) at the equatorial Phuket geomagnetic station in Thailand,” Advances in Space Research, vol. 70, no. 5, pp. 1429–1440, 2022, doi: 10.1016/j.asr.2022.06.024.

[52] J. Matzka, O. Bronkalla, K. Tornow, K. Elger, and C. Stolle, “Geomagnetic Kp index,” GFZ Data Services, 2021, doi: 10.5880/Kp.0001.

[53] J. Matzka, C. Stolle, Y. Yamazaki, O. Bronkalla, and A. Morschhauser, “The geomagnetic KP index and derived indices of geomagnetic activity,” Space Weather, vol. 19, no. 5, May 2021, doi: 10.1029/2020SW002641.

[54] C. A. Loewe and G. W. Prölss, “Classification and mean behavior of magnetic storms,” Journal of Geophysical Research: Space Physics, vol. 102, no. A7, pp. 14209–14213, 1997, doi: 10.1029/96JA04020.

[55] W. D. Gonzalez et al., “What is a geomagnetic storm?,” Journal of Geophysical Research: Space Physics, vol. 99, no. A4, pp. 5771–5792, 1994, doi: 10.1029/93JA02867.

[56] Y. Kamide et al., “Two-step development of geomagnetic storms,” Journal of Geophysical Research: Space Physics, vol. 103, no. A4, pp. 6917–6921, 1998, doi: 10.1029/97JA03337.

[57] A. Rozhnoi, M. S. Solovieva, O. A. Molchanov, and M. Hayakawa, “Middle latitude LF (40 kHz) phase variations associated with earthquakes for quiet and disturbed geomagnetic conditions,” Physics and Chemistry of the Earth, Parts A/B/C, vol. 29, no. 4–9, pp. 589–598, 2004, doi: 10.1016/j.pce.2003.08.061.

[58] M. E. Contadakis, D. N. Arabelos, C. Pikridas, and S. D. Spatalas, “Total electron content variations over southern Europe before and during the M 6.3 Abruzzo earthquake of April 6, 2009,” Annals of geophysics, vol. 55, no. 1, 2012, doi: 10.4401/ag-5322.

[59] R. E. Fan, K. W. Chang, C. J. Hsieh, X. R. Wang, and C. J. Lin, “LIBLINEAR: a library for large linear classification,” Journal of Machine Learning Research, vol. 9, pp. 1871–1874, Aug. 2008, doi: 10.1145/1390681.1442794.

[60] I. B. Mohamad and D. Usman, “Standardization and its effects on K-means clustering algorithm,” Research Journal of Applied Sciences, Engineering and Technology, vol. 6, no. 17, pp. 3299–3303, Sep. 2013, doi: 10.19026/rjaset.6.3638.

[61] P. G. Fennell, Z. Zuo, and K. Lerman, “Predicting and explaining behavioral data with structured feature space decomposition,” EPJ Data Sci., vol. 8, no. 1, Dec. 2019, doi: 10.48550/arXiv.1810.09841.

[62] M. Fagerland, S. Lydersen, and P. Laake, Statistical Analysis of Contingency Tables. Boca Raton, Florida: Chapman, 2017.

[63] S. C. Chow, J. Shao, H. Wang, and Y. Lokhnygina, Sample Size Calculations in Clinical Research. Boca Raton, Florida: Chapman, 2018.

[64] A. L. Edwards, “Note on the ‘correction for continuity’ in testing the significance of the difference between correlated proportions,” Psychometrika, vol. 13, no. 3, pp. 185–187, 1948, doi: 10.1007/BF02289261.

[65] H. Zhang, “The Optimality of Naive Bayes,” in Proceedings of the Seventeenth International Florida Artificial Intelligence Research Society Conference, FLAIRS 2004, Jan. 2004, vol. 2.

[66] V. Metsis, I. Androutsopoulos, and G. Paliouras, “Spam Filtering with Naive Bayes - Which Naive Bayes?,” Jan. 2006.

[67] A. Mccallum and K. Nigam, “A Comparison of Event Models for Naive Bayes Text Classification,” Work Learn Text Categ, vol. 752, May 2001.