Fırat Üniversitesi Müh. Bil. Dergisi
36(1), 481-498, 2024

Araştırma Makalesi
https://doi.org/10.35234/fumbd.1425635

Anomaly Detection in Meteorological Data Using a Hierarchical Temporal Memory Model: A Study on the Case of Kazakhstan

Kürşat Mustafa KARAOĞLAN[1], Oğuz FINDIK[2], Erdal BAŞARAN[3*]
[1,2] Department of Computer Engineering, Faculty of Engineering, Karabük University, Karabuk, Türkiye
[3] Department of Computer Programming, Vocational School, Agri Ibrahim Cecen University, Agri, Türkiye

(Geliş/Received: 25/01/2024; Kabul/Accepted: 27/03/2024)

Abstract: In meteorology, which studies atmospheric events, data representing various properties such as temperature, rainfall, and wind speed are collected regularly over a certain period. Unexpected trends in the data may indicate that an abnormal situation is approaching. Therefore, time series (TS) data play an essential role in the early detection of potential meteorological risks. However, applying effective models by considering many complex parameters in performing accurate analysis and anomaly detection (AD) is an important criterion. In this study, machine learning-based AD is performed using a dataset containing meteorological data on different features collected between January 1, 2019, and June 30, 2023, for Kazakhstan, which has the ninth-largest surface area in the world. The Hierarchical Temporal Memory (HTM) model was used for AD, which can provide more accurate forecasts by modeling long-term dependencies and producing effective results in solving TS problems. Detected anomalies are reported at various levels depending on threshold values. In addition, to analyze the ADs more precisely, correlations are calculated using the Spearman model, which allows us to determine the strength and direction of the monotonic relationship between variables. The study's findings show that the HTM is an effective model for AD using TS data on meteorological features.

Keywords: Anomaly detection, time series, meteorological anomalies, machine learning, hierarchical temporal memory.

Hiyerarşik Zamansal Bellek Modeli ile Meteorolojik Verilerdeki Anomalilerin Tespiti: Kazakistan Örneği Üzerine Bir Çalışma

Öz: Atmosferik olayları inceleyen meteorolojide, sıcaklık, yağış ve rüzgar hızı gibi çeşitli özellikleri temsil eden veriler belirli bir süre boyunca düzenli olarak toplanmaktadır. Verilerdeki beklenmedik eğilimler anormal bir durumun yaklaşmakta olduğunu gösterebilmektedir. Bu nedenle, zaman serisi verileri potansiyel meteorolojik risklerin erken tespitinde önemli bir rol oynamaktadır. Ancak doğru ve güvenilir analizlerin gerçekleştirilmesinde ve anomali tespitinde karmaşık birçok parametreyi göz önünde bulundurarak etkin modelleri uygulamak önemli bir kriterdir. Bu çalışmada, dünyanın en büyük dokuzuncu yüzölçümüne sahip Kazakistan için 1 Ocak 2019 ile 30 Haziran 2023 tarihleri arasında toplanan farklı özelliklerdeki meteorolojik verileri içeren bir veri seti kullanılarak makine öğrenmesi tabanlı anomali tespiti gerçekleştirilmiştir. Anomali tespiti için uzun vadeli bağımlılıkları modelleyerek daha doğru tahminler sağlayabilen ve zaman serisi problemlerinin çözümünde etkin sonuçlar üreten Hiyerarşik Zamansal Bellek (HTM) modeli kullanılmıştır. Tespit edilen anomaliler eşik değerlerine bağlı olarak çeşitli seviyelerde raporlanmıştır. Ayrıca, anomali tespitlerini daha hassas bir şekilde analiz etmek için, değişkenler arasındaki monotonik ilişkinin gücünü ve yönünü belirlememizi sağlayan Spearman modeli kullanılarak korelasyonlar hesaplanmıştır. Çalışmanın bulguları, HTM modelinin meteorolojik özelliklere ilişkin zaman serisi verilerinin kullanıldığı AD problemlerinde etkin bir araç olduğunu göstermektedir.

Anahtar kelimeler: Anomali tespiti, zaman serileri, meteorolojik anomaliler, makine öğrenmesi, hiyerarşik zamansal bellek.

1. Introduction

Time is a non-modifiable intrinsic factor that plays an important role in the human brain's learning process [1]. This intangible factor can be found in many different contexts, ranging from measurements of natural phenomena such as meteorology to artificial systems such as stock markets and robotics, and it can be found in real-world problems. The evaluation and effective modeling of phenomena with their temporal components [2] help to achieve more accurate predictions and advances in various fields of study and practical applications. In this context, temporal data allow in-depth data analyses thanks to the opportunities to capture the dynamics and evolving patterns inherent in complex phenomena [3]. Time series (TS), on the other hand, represent an ordered and temporally indexed collection of data points, usually obtained at regular intervals [4] [5]. TS are essential in various academic disciplines as they are numerical, continuous, and high dimensional. Each time point can often be associated with more than one variable or feature and needs to be constantly updated with new observations. In

[*] Sorumlu yazar: erdalbasaran085@gmail.com. Yazarların ORCID Numarası: [1]0000-0001-9830-7622, [2] 0000-0001-5069-6470, [3] 0000-0001-8569-2998

this way, TS allows researchers to analyze and discover emerging trends, non-specific patterns, and existing relationships [6].

Time series analysis (TSA) has been an active research topic for researchers for many years [7] [8]. Researchers have conducted a great deal of research in areas covering various aspects such as TSA [9], searching for subsequences, segmentation metrologies [10], and dimensionality reduction techniques [11]. Early work in the literature attracted a great deal of attention from the database and pattern recognition communities. TSA was observed to contribute critical insights in a variety of domains while at the same time contributing to a better understanding of the broad scope of TSA and the development of techniques, thereby increasing the usability of these techniques in different contexts [6] [12]. Currently, TSA is widely used in fields ranging from cybersecurity [13] to finance [14], economics [14], and environmental sciences [13] [14] to perform subtasks such as revealing valuable features, making reliable predictions, and guiding decision-making processes. Furthermore, researchers have gained the ability to perform anomaly detection (AD) [15], outcome prediction, examine relationships between variables, analyze dependencies, patterns, and trends in TS, and evaluate the effects of various factors using statistical or Machine learning (ML)-based methods such as Hierarchical Temporal Memory (HTM). In this context, researchers can successfully perform basic tasks such as classification [16] [17], and clustering [18] using these techniques [2].

Anomalies refer to deviated, discordant, or outlier values that do not conform to the expected normal [19]. AD is one of the challenging tasks of Data Mining [19]. It is the process of detecting or identifying data points, anomalies, events, and components that deviate from the expected pattern [20]. Three main types of anomalies are usually identified in the literature [21]. The first type is the point anomaly, which results from the significant deviation of a data sample from the overall dataset [22]. The second type is contextual anomalies, which exhibit anomalous behavior within a specific context. The third type is collective anomalies, which refer to similar data samples that exhibit anomalous behavior concerning the entire dataset [23]. Classical methods for AD, such as linear models [24], distance-based methods [25], and support vector machine [26], are still valid. However, the presence of multidimensional data may impose limitations, and difficulties may arise when dealing with large and complex systems due to the need for labeled anomalies. Furthermore, TSA can present additional challenges as it considers sequencing and cause-effect relationships along the time axis. Researchers have developed various methods to overcome these challenges [27].

Accurate detection of anomalies fulfills essential roles in decision-making, risk management, predictive maintenance, and early warning systems [28]. AD in TS problems is of great interest due to its practical applications and potential impact on various industries. Such AD helps organizations proactively approach critical problems, minimize losses, and improve operational efficiency. In the era of big data and increasing digitalization, effectively detecting anomalies in TS data is crucial [29]. Early detection of outliers in TS data can play a critical role in preventing costly failures and mitigating potential risks [30]. Identifying and flagging anomalous patterns in TS is critical in detecting fraudulent activities, identity or passport documents, insurance claims, and healthcare fraud [31], predicting system failures, and ensuring the smooth operation of complex systems [32].

AD in TS data is categorized as univariate and multivariate based on the input data [33] [34]. Univariate AD focuses on identifying deviations of individual variables from long-term patterns and identifies outliers according to the general distribution or discriminant model. On the other hand, multivariate AD problems involve multiple variables observed at each timestamp or time instant, and this type of AD considers the relationships and interactions between variables, which is a more complex process than univariate analyses [35]. Anomaly studies aim to detect point, contextual, collective, or pattern anomalies [36].

Traditional approaches in AD research usually include statistical, classification-based, clustering-based, and information-theoretic methods, such as PCA, Support Vector Machines, k-nearest Neighbor algorithm, and various correlation analysis techniques [28] [37]. However, these approaches can be needed to handle system dynamics that change over time, making it challenging to characterize anomalous contexts accurately. In recent years, successful results have been achieved in detecting complex contextual anomalies with dynamic and time-varying features, such as RNN-based architectures. Moreover, some works in the literature have provided an overview of semi-supervised, unsupervised, and hybrid DF techniques and discussed their advantages, disadvantages, and computational complexities [38]. In particular, other research focuses on deep learning-based AD approaches such as Boltzmann machines, Autoencoders, and RNNs [39]. In addition, AD has an essential role in many fields; there are various studies in the literature in this field, for example, in cyber security [40] [41], fraud detection [42], medical field, industrial damages, images, textual data, sensor networks, geochemical data processing [43], and various ML and non-ML techniques as well as statistical and spectral sensing methods for AD in discrete/symbolic sequences [28] [44] [45] in the work of Nassif et al., several ML techniques used in AD as independent models in various applications are presented hierarchically under the categories: classification, ensemble analysis, optimization, rule-based systems, clustering, and regression. These methods can be used as independent models [44].

DL approaches differ from surface ML algorithms in that they can perform feature extraction and classification tasks simultaneously. Furthermore, these approaches differ regarding integrated functionalities and resource requirements, such as data dependency and hardware dependency [46]. In the field of AD, many DL-based approaches have been proposed, especially in intrusion detection in cyber security [39] [47] [48]. Various aspects of using DL techniques have been discussed in detail. Multivariate AD involves multiple variables observed at the same timestamp or instant [34]. In such problems, multiple data points correspond to different variables at each time step. Such analyses present additional challenges compared to univariate analyses, as the relationships and interactions between variables need to be considered to develop an effective AD and modeling approach. Although various methods are presented in the literature for detecting anomalies in TS data, most of these methods primarily focus on univariate TS approaches [49]. Another essential approach in AD studies is HTM. HTM is a biologically based ML technique that mimics the architecture and functioning of the neocortex [19]. Notable features of HTM include sparsity, hierarchy, and modularity. The method uses three main components, encoder, spatial pooling (SP), and temporal memory (TM), to efficiently discover patterns and contexts in TS data. HTM can learn from continuously evolving data without the need for labeled data, which increases the usability of HTM in scenarios where manual labeling is difficult or costly. The continuous learning capability allows the model to adapt to changing data structures and effectively identify new anomalies [50] [51]. Furthermore, HTM can store, remember, and predict patterns in an aggregated manner, which provides the ability to anticipate the next pattern occurrence.

AD can produce important outputs in various fields, including meteorology, where it plays an essential role in identifying irregular patterns, outliers, and anomalies in TS data. In particular, meteorologists are keenly interested in identifying anomalies in weather data to predict future outcomes [52]. Accurately detecting anomalies in meteorological data can contribute to critical disaster management, agricultural planning, and environmental monitoring decisions. For this reason, accurate and reliable analysis of meteorological data is of great importance for weather forecasts and climate change studies. Detection of anomalies in these analyses is critical in warning of extreme weather events and other unexpected situations. In addition, integrating AD methods into meteorological research offers many benefits, such as improving weather forecast accuracy, monitoring climate change dynamics, and assisting in disaster management.

This study aims to develop a practical and sensitive approach for AD on TS data of meteorological events occurring from 1 January 2019 to 30 June 2023 in Kazakhstan, the ninth largest country in the world in terms of area. In this study, an AD approach based on HTM, which is among the state-of-the-art approaches known for its effectiveness in AD tasks [53] performed on TS problems, is implemented. AD performances are evaluated by considering the correlations of features. HTM, which can be applied unsupervised by eliminating the need for labeled data, has shown effective performance on TS data since it is designed in a hierarchical structure to capture complex temporal patterns. The publicly available dataset used in the study includes comprehensive meteorological data, including temperature, precipitation, cloud cover, atmospheric pressure, and wind [54]. In addition, this study contributes to developing methods for the sensitive, accurate, and reliable detection of unusual anomalies in meteorological TS data. The study focuses on developing an advanced and region-specific AD system in this context. It has the potential to assist in monitoring climate change dynamics and effective disaster management.

The contributions of the paper presented below are summarized:

- An efficient and sensitive approach for AD in meteorological TS data of Kazakhstan, covering the period from January 2019 to June 2023, is developed.
- The applicability, performance, and effectiveness of HTM, a state-of-the-art unsupervised ML technique for AD, are presented.
- Univariate AD was performed on meteorological features and output performances were evaluated.
- Using a comprehensive data set, various meteorological parameters such as temperature, precipitation, cloud cover, atmospheric pressure, and wind properties are evaluated in terms of anomaly.
- A sensitive, accurate, and robust AD system is proposed, which contributes to monitoring climate change dynamics and effective disaster management.
- Thresholds are used to define the level of anomalies, and the Spearman correlation method was used to measure the strength and direction of the monotonic relationship between variables to improve anomaly sensitivity. In addition, this study is among the pioneering studies on HTM-based AD tasks on meteorological data.

This paper is organized into five main sections to contribute to a comprehensive study of the AD task on meteorological TS data using HTM. The study is organized according to the following sections: The second section, methodology, presents the HTM, its components, and the applied correlation method. The details of the proposed approach are explained in detail in the third section. The fourth section presents the experimental studies, the dataset and hyperparameters used, HTM-based anomaly results, correlation analysis results and findings. The last section is reserved for discussion and conclusion.

2. Methodology

This section provides an overview of the methodologies used to implement the proposed approach, focusing on the HTM approach and correlation techniques. The correlation techniques help in the meaningful interpretation of the output analyses of HTM conducted univariately to address the AD analysis process effectively. Thus, correlation methods effectively assess the interrelationships and dependencies in the data and provide an accurate and reliable handling of HTM outputs.

2.1. HTM and components

In statistical methods, a model is generated based on the data obtained from normal behavior. Then, the conformity of new samples to the model is performed by statistical inference test. According to the test results, the samples with a low probability of being generated from the learned model are removed as outlier samples [28]. In ML methods, labeled data are needed to solve anomaly classification problems. Although ML approaches have the potential to exhibit high performance in domain-specific AD problems, factors such as labeling cost and labeling accuracy of the data should be considered. However, distinguishing data exhibiting anomalous behavior in large volumes of flow data and parsing model input data sets may not produce reliable results. Unlike statistical and ML approaches, HTM, an unsupervised ML approach, can continuously learn evolving data models with the hierarchical components in its architecture and capture the temporal context efficiently. In addition, HTM has various advantages, such as applying to dynamic and complex data models, operating in real-time, being adaptable to AD scenarios involving univariate or multivariate analyses, and being robust to noise [19]. HTM is a detailed computational theory that mimics the working principles of the neocortex. HTM incorporates time-based learning techniques for capturing and recalling spatial and temporal patterns [55]. The primary data structure in HTM is sparse distributed representation (SDR) [19] [56] [57]. The HTM methodology for processing TS data includes essential components such as Encoder, Spatial Pooling (SP), and Temporal Memory (TM) [58]. The encoder part normalizes the data by converting the input data into binary representations called fixed-length SDRs [59]. This stage prepares the data for further processing. SP ensures that the encoded data maintains a constant sparsity, typically not exceeding 2%, to preserve essential features. The TM stage focuses on recognizing and extracting temporal information patterns after the SP step, thus providing a higher-level representation [60] [61]. Moreover, the other component in HTM represents the task of classification [62]. However, the classifier process is not included in this study. Figure 1 presents a visual representation of the HTM pipeline designed to handle AD.

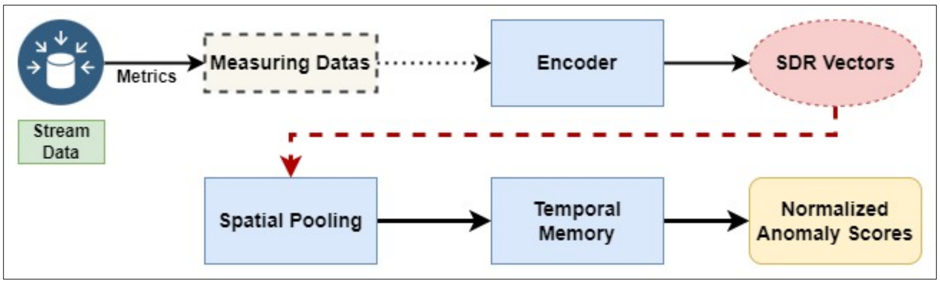

Figure 1. A general HTM pipeline for detecting anomalies in TS data [15]

In HTM, which takes an approach similar to the Hebbian learning rule [63], the temporal pattern learning process first captures spatial patterns between input bits by frequently identifying bits that occur together. When these bits overlap, mini columns are activated. A set of active mini columns represents a spatial pattern. The HTM then starts to learn the transitions between these spatial patterns. Active cells connect with recently active cells via distal synapses and are pattern recognizers to predict future patterns. When the distal basal segment recognizes a memorized pattern, it depolarizes the cell, making it predictive. If the expected input occurs, the cell correctly predicts and activates; if an unexpected input arrives, all cells within the selected mini columns collectively recognize it as a new pattern and activate. This collective process enables HTM to store, retrieve, and predict subsequent patterns smoothly [19] [64].

2.1.1. Encoder

The encoder is responsible for normalization by transforming the input data into binary representations called fixed-length SDRs, making them suitable for the input forms of other components. SDRs are the main data type used in the cortex and is used throughout HTM systems. SDRs match memory space representations in terms of mismatch probability, robustness to noise, subsampling, and classification of vectors and combinations [59]. The encoder process is performed according to the fundamental parameters such as the length of the SDR vector to be generated, the number of open bits (value of 1) and the sparsity rate. SDRs generated with encoders designed with the right parameters have a strong potential for high performance in learning temporal sequences. In the HTM, there is a sequence of SDRs that can be designed differently depending on the specific purposes with various functions in different parts of the neuron and in the algorithm structure [57]. SDRs differ from standard computer notation, such as ASCII, because they are encoded for direct representation. They consist of an extensive sequence of bits, mostly zeros and a few ones, and are, therefore, sparse. In addition, each bit in an SDRs has semantic properties, so if two SDRs have more bits of overlapping value, the two SDRs have comparably similar meanings at different proportional values. Consequently, the first step in using an HTM system is to convert a data source into an SDRs using an encoder that determines which output bits will have a value of one or zero for a given input [65].

Figure 2 (a) shows the SDR vector produced by processing the time and the first feature together on the dataset used with the encoder whose parameters are set. Figure 2 (b) shows the SDR vector that includes time and all features of the dataset, representing the first input of the dataset. In addition, the 1-valued bits of the presented SDR vectors are shown in red, and the 0-valued bits in black color, and their numerical occurrence numbers are also presented.

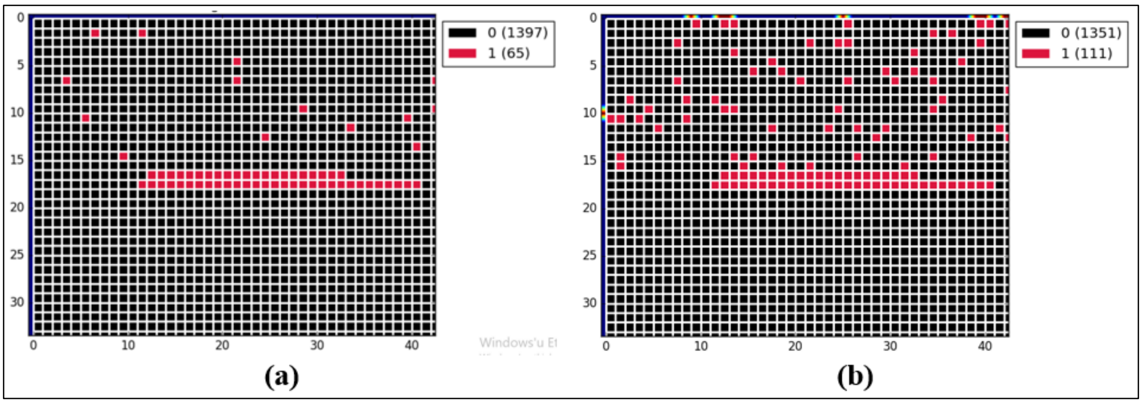

Figure 2. (a) SDR produced by processing the timestamp and the first feature. (b) SDR containing all features of the first input of the timestamp dataset, representing the first input.

2.1.2. Spatial pooling (SP) and temporal memory (TM)

The SP phase ensures that the encoded data maintains a constant sparsity, typically not exceeding 2%, to preserve essential features. The TM phase focuses on recognizing and extracting temporal information patterns after the SP step, thus providing a higher-level representation [55] [56]. Moreover, the other component in HTM represents the classification task [62]. However, the classifier process is not included in this study. SP, responsible for SDR encoding of binary input data, combines competitive Hebbian learning rules and homeostatic excitability control [66]. In SP, binary vectors in the input space are transformed into sparse sequences. The SP ensures that the sparsity and overlap of the input space of the HTM are preserved. Therefore, similar input data have high overlap, while different input data have low overlap [67]. The bits (columns) that are active after the SP task to represent the first input of the dataset used in this study are given in Figure 3.

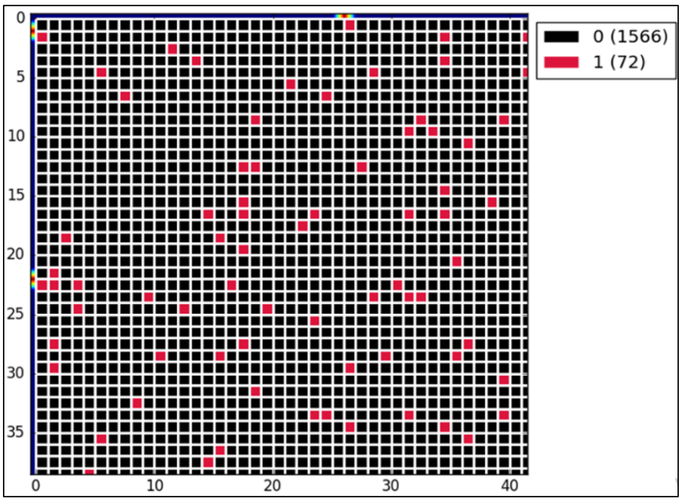

Figure 3. Active columns after SP process.

The TM phase focuses on recognizing and extracting temporal information patterns after SP phase, thus providing a higher-level representation [60] [61]. In this way, it learns from the SDR sequences generated over time by the SP and predicts the next incoming pattern based on the temporal context of each input. The TM task follows the SP and determines the current active temporal state. The working principle of TM is to activate cells in columns, make them predictable and update the persistence of synapses [68].

2.2. Spearman's correlation

Correlation analysis is a statistical method widely used to summarize scientific research data. Correlation analysis is used to determine whether there is a relationship between two variables and to assess the degree of interdependence by measuring the strength of the relationship (r) [69]. Regression approaches seek to express the specific form of the relationship between selected values of one independent variable and the means of all corresponding values of the second dependent variable. In this aspect, correlation differs from regression, but correlation is the basis of regression theory [70] [71]. The correlation coefficient, denoted as r, theoretically ranges between +1 and -1. When interpreting correlations, 0 indicates no linear relationship, +1 indicates a perfect positive linear relationship, and -1 indicates a perfect negative linear relationship. Furthermore, values between 0 and 0.3 (or -0.3) represent a weak positive (negative) linear relationship, values between 0.3 and 0.7 (or -0.7) indicate a moderately positive (negative) linear relationship, and values between 0.7 and 1.0 (or -0.7 and -1.0) indicate a strong positive (negative) linear relationship. The correlation coefficient assumes linearity, which makes it reliable for linear relationships but not practical for non-linear relationships and requires careful consideration of linearity assumptions when interpreting data [72].

In statistics, Spearman's rank correlation coefficient, denoted ρ (rho) or sometimes rs, is named after Charles Spearman [73]. It is a nonparametric metric used to measure rank correlation and measures the statistical dependence between the ranks of two variables. This coefficient assesses how effectively a monotonic function can characterize the relationship between two variables. The Spearman correlation between two variables is essentially the Pearson correlation applied to the rank values of these variables. While Pearson correlation is concerned with assessing linear relationships, Spearman correlation focuses on monotonic relationships, whether linear or not. Without repeated data values, a reliable Spearman correlation of +1 or -1 is obtained when each variable is an optimal monotonic function of the other. The Spearman coefficient is suitable for both continuous and discrete ordinal variables. It is important to note that Spearman's ρ and Kendall's τ can be seen as exceptional cases of a more general correlation coefficient, reflecting their versatility in various statistical applications [74] [75]. In line with the information presented above, Spearman's correlation was applied to statistically examine the relationships between features in this study.

$$p = 1 - [(6 * \sum d^2)/(n * (n^2 - 1))] \tag{1}$$

In Equation 1, p represents Spearman's rank correlation coefficient. $\sum d^2$ is the sum of squared differences between the ranks of the corresponding data pairs. n represents the number of data pairs compared.

3. Proposed Approach

In this study, the pipeline for the approach developed for performing AD on meteorological data with the HTM approach is presented in Figure 4. According to this structure, the data is considered as a whole and prepared as separate inputs by selecting features according to time information to transform them into a format suitable for the inputs of the ML model. The prepared input data were transformed into SDR vectors and prepared for the SP stage. Thus, the data were transformed into sparse sequences, and temporal information patterns were obtained. Overlaps were also detected with the SP. In the next phase, AD was performed by training the dataset with a hybrid network model that predicts the next model based on the temporal input. Then, each input was used in the TM phase to perform AD. Figure 3 shows the encoder outputs of the input data as described in Figure 2 and Figure 3. In addition, the parameters of each HTM component presented in Figure 4 were fine-tuned and applied. Information about the parameters is presented in detail in the sections on the experimental studies.

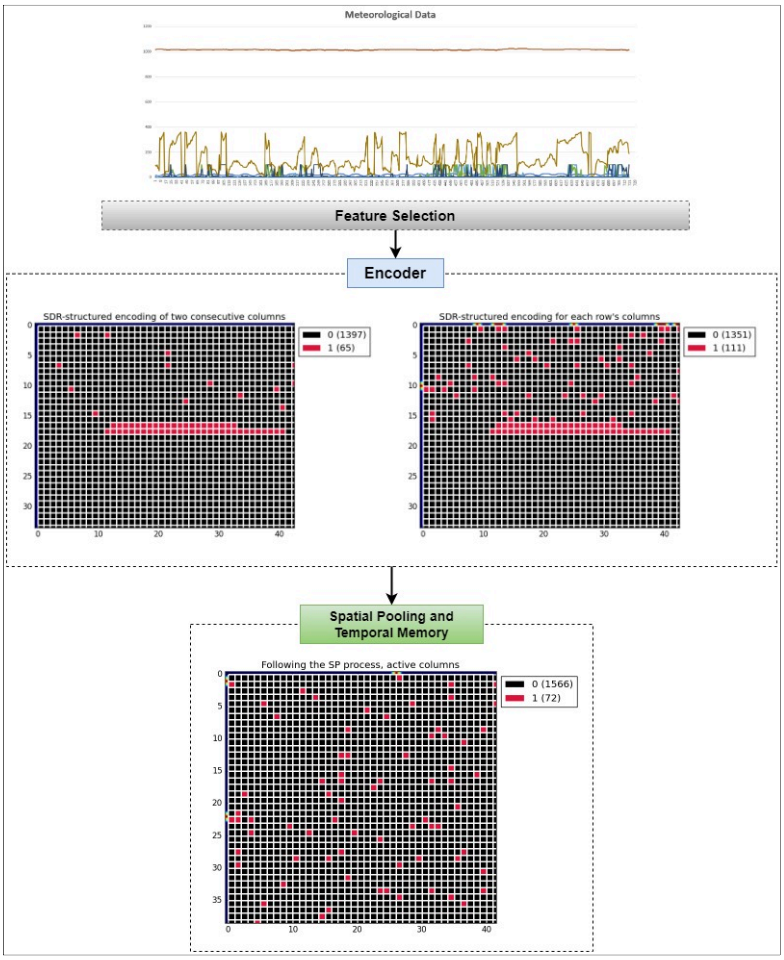

Figure 4. (a) SDR generated by processing only the first feature with the timestamp. (b) SDR containing all features of the first input of the timestamp dataset.

4. Experimental Studies

In this section, correlation calculations and evaluations are presented to analyze the relationships between features in the experimental setting, results, and interpretation of outputs for AD experiments performed with HTM.

4.1. Experimental Setting

4.1.1. Dataset

For the detection of anomalies, a dataset [54] consisting of 11 features and data from 1642 days is used in this study. For all features in the dataset, measures of central tendency such as mean (μ) and median, as well as measures of dispersion such as maximum (max), minimum (min), and standard deviation (σ), are calculated and presented. The dataset was reorganized by calculating daily means or sums of hourly temperature (temp), rainfall (), snowfall (s), total_cloud (t_c), high_level_cloud (h_c), mid_level_cloud (ml_c), low_level_cloud (l_c), pressure (pr), wind_speed (w_s), wind_direction (w_d) and turbo_wind (t_w). In addition, p-values were calculated separately to assess the statistical significance of the data and to perform regular distribution tests to determine whether each attribute was normally distributed. Since the p-values for each attribute in the table were calculated below the generally accepted threshold value of 0.05, it was determined that it did not have a normal distribution according to the condition presented and expressed with a TRUE value. In addition, as indicated in Table 1, the daily values of some features were calculated by averaging, and the daily values of some features were summed.

Table 1. Features and statistical details about the dataset used [54].

Date Range			:	January 1, 2019- June 30, 2023							
# Samples			:	1642							
# Features			:	11							
Averaged features			:	[temp], [t_c], [h_c], [ml_c], [l_c], [pr], [w_s], [w_d], [t_w]							
Summed features			:	[r],[s]							
				i∈{1,2,3,4,5,6,7,8,9,10, 11} -> S_i							
Type	**temp (°c)**	**r (mm)**	**s (mm)**	**t_c (%)**	**h_c (km)**	**ml_c (km)**	**l_c (km)**	**pr (hpa)**	**w_s (m/s)**	**w_d (degrees)**	**t_w (m/s)**
μ	12.7	1.3	0.1	46.4	18.7	29.2	35.2	1017.4	3.2	198.4	5.0
median	12.2	0.0	0.0	48.3	9.7	22.5	30.2	1017.8	2.7	201.8	4.2
max	30.4	49.2	18.3	100.0	98.9	100.0	100.0	1041.5	13.5	344.9	21.3
min	-4.8	0.0	0.0	0.0	0.0	0.0	0.0	977.1	0.7	54.2	1.0
σ	7.1	3.3	0.7	31.9	21.4	28.6	30.8	8.4	1.8	60.2	2.7
p value p<0.05	~0, True	0, True	~0, True	~0, True	~0, True	~0, True	~0, True	~0, True	~0, True	0, True	0, True

A normal distribution test is a statistical procedure used to determine whether a data set conforms to a normal distribution, also known as a Gaussian distribution. Normality tests assess how closely data distribution conforms to the theoretical characteristics of a normal distribution, such as symmetry and specifically associated mean and standard deviation values. Commonly used normality tests include the Kolmogorov-Smirnov test, the Shapiro-Wilk test, the Anderson-Darling test, and graphical methods such as the Q-Q plot. These tests are essential in statistical analysis as they enable the selection of appropriate statistical techniques and help the reliability of statistical inferences [76] [77] [78]. This study used the Kolmogorov-Smirnov technique to calculate the regular distribution tests for the characteristics.

4.1.2. Hyperparameters

Table 2 presents the hyperparameters adapted for the AD task in the HTM model. Each table row introduces hyperparameters related to a different component of the HTM framework, namely the Encoder, SP, and TM. In this context, Table 2 shows the configuration settings required to implement the HTM model with optimal performance in the AD context, separating the hyperparameters according to each HTM component.

Tablo 2. Hyperparameters for the applied model.

Encoder	resolution: 0.88, size: 700, timeday: 30,1, weekend:21, sdralpha (predictor):0.1
SP	boostStrength: 3.0, columnCount: 1638, localAreaDensity: 0.04, potentialPct: 0.85, synPermActiveInc: 0.04, synPermConnected: 0.1, synPermInactiveDec: 0.006
TM	activationThreshold: 17, cellsPerColumn: 13, initialPerm: 0.21, maxSegmentsPerCell: 128, maxSynapsesPerSegment: 64, minThreshold: 10, newSynapseCount: 32, permanenceDec: 0.1, permanenceInc: 0.1, anomaly period: 1000

4.2. Experimental results

This section includes HTM-based anomaly results with graphs for each feature. In addition, Spearman correlation results are also presented to evaluate the anomaly outputs more precisely by determining the strength of the relationship between the features.

4.2.1. Results of HTM-based AD Approach

This section presented the results of the experimental studies to detect significant anomalies from typical climate models using various meteorological features with the HTM approach. The figures below (Figures 5-15) show the anomaly scores and the metric values for each meteorological feature of interest.

Figure 5 shows the anomaly scores and metric values for the Temperature feature. Notable high-level anomalies occurred on June 20, 2022, and July 20, 2022, indicating sudden temperature increases. In addition, moderate anomalies were observed on 13.07.2022 and several dates in August 2022, highlighting significant deviations from expected temperature trends.

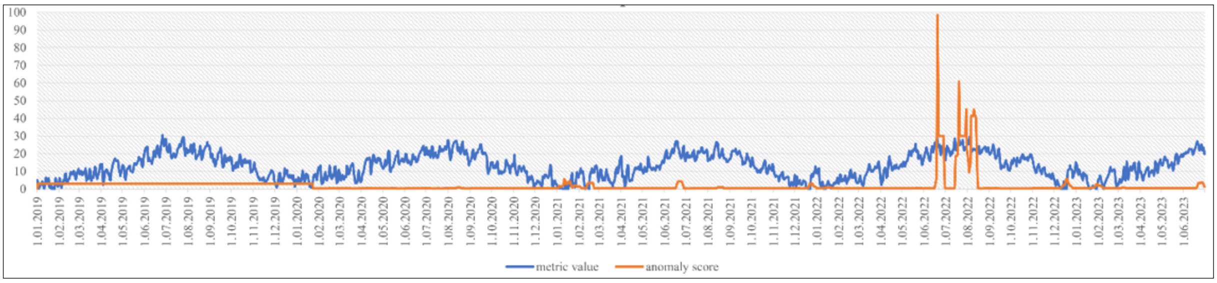

Figure 5. Plot of anomaly scores and metric values for the temperature feature.

Figure 6 shows the anomaly scores for rainfall and the resulting metric values. High-level anomalies were detected on August 8, 2021, and July 14, 2021, and indicate anomalous precipitation events. Moderate anomalies were also recorded, indicating significant changes in precipitation patterns, especially on August 7, 2022, and April 15, 2022.

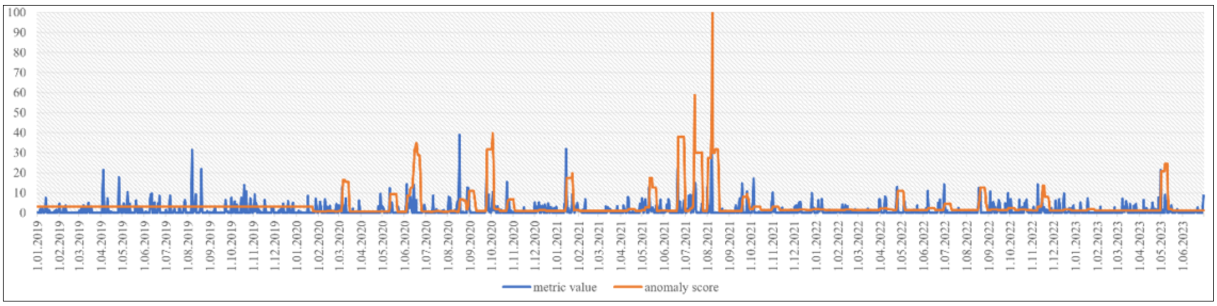

Figure 6. Plot of anomaly scores and metric values for the rainfall feature.

Figures 7 and 8 represent Snowfall and Total Cloud features, respectively. Snowfall anomalies were observed on February 11, 2021, January 15, 2021, December 16, 2022, and April 10, 2022. Total Cloud anomalies were relatively sparse, with only moderate anomalies detected on March 6, 2021.

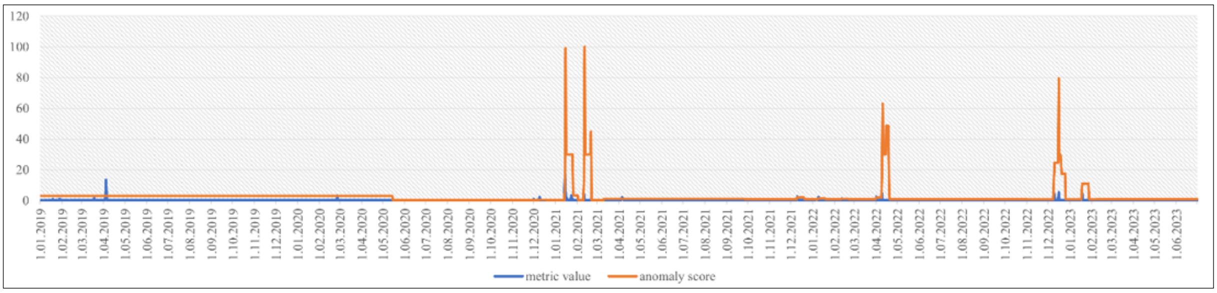

Figure 7. Plot of anomaly scores and metric values for the snowfall feature.

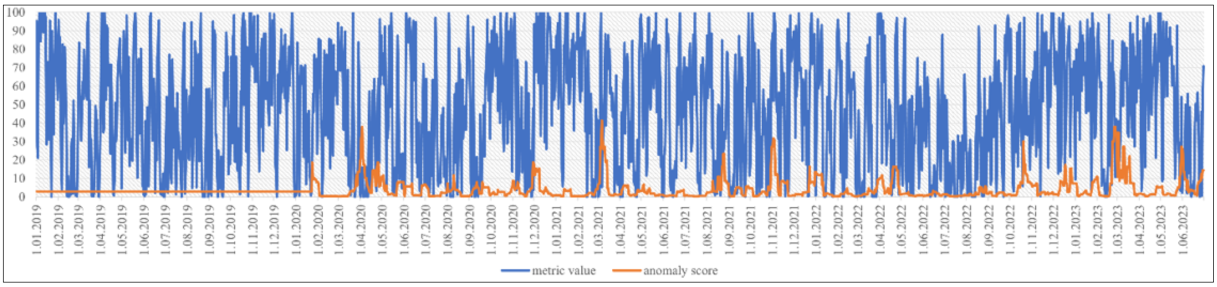

Figure 8. Plot of anomaly scores and metric values for the total cloud feature.

Figure 9 shows anomalies in the High Cloud feature, with high-level anomalies recorded on December 28, 2020. The mid-level anomalies were also spread over various dates in 2020 and 2021, highlighting the continuous deviations in the high cloud patterns.

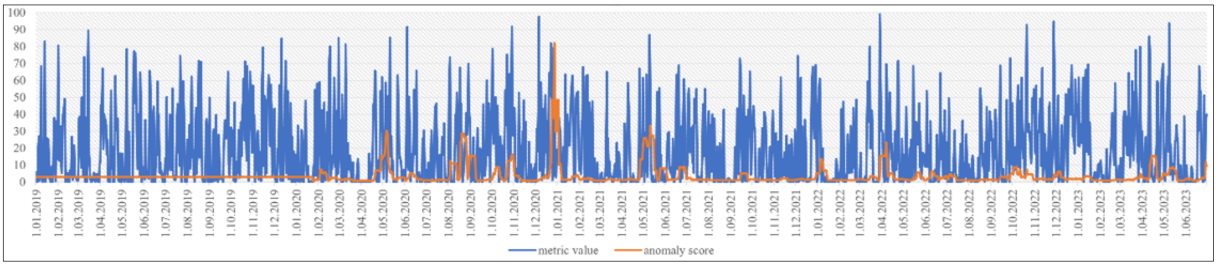

Figure 9. Plot of anomaly scores and metric values for the high-level cloud feature.

Figures 10 and 11 show the anomalies in the Mid-level Cloud and Low Cloud features, respectively. For the Mid-level Cloud, significant high-level anomalies occurred on February 15, 2022, August 20, 2020, April 3, 2022, and August 29, 2020. The Low Cloud, on the other hand, showed high-level anomalies on October 1, 2020, and June 18, 2020. Mid-level anomalies were also observed to be spread over various dates.

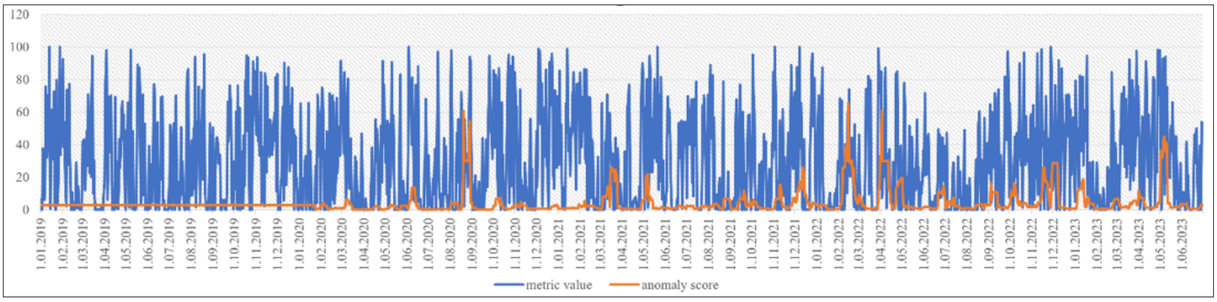

Figure 10. Plot of scores and metric values for the mid-level cloud feature.

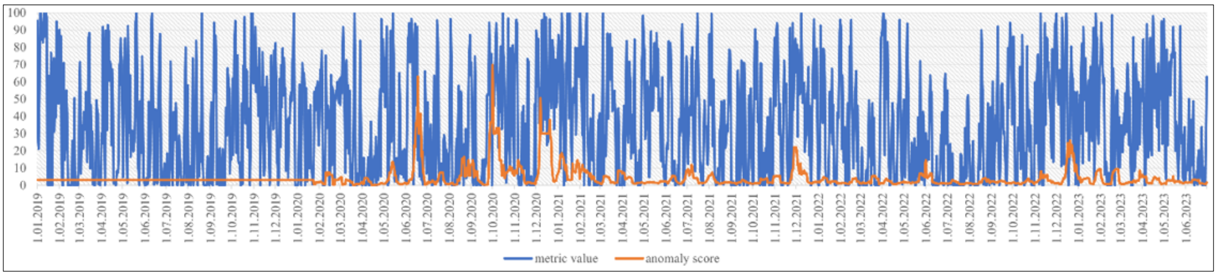

Figure 11. Plot of anomaly scores and metric values for the low-level cloud feature.

Figure 12 represents the results of AD in the pressure feature. No high-level anomalies were observed in the pressure feature. However, there is a moderate anomaly on February 12, 2023.

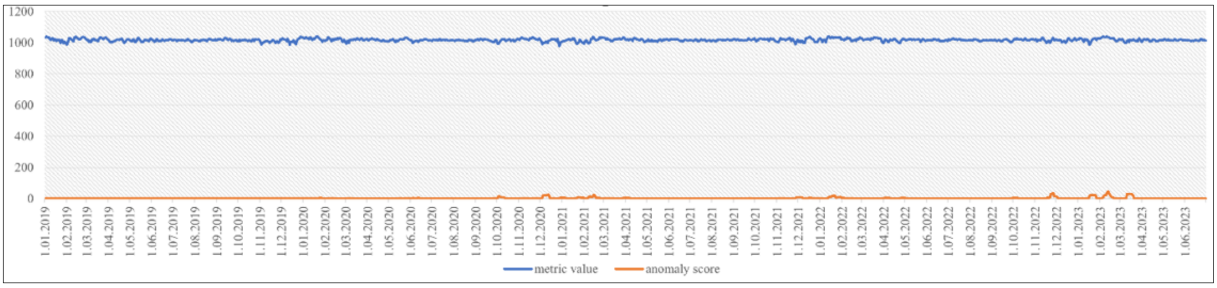

Figure 12. Plot of anomaly scores and metric values for the pressure feature.

Figure 13 represents wind speed, showing the absence of medium- and high-level anomalies and indicating consistent wind speed trends. In contrast, Wind Direction (Figure 14) exhibited high-level anomalies on February 28, 2023, medium-level anomalies on December 17, 2022, and February 27, 2023, indicating significant changes in wind direction patterns.

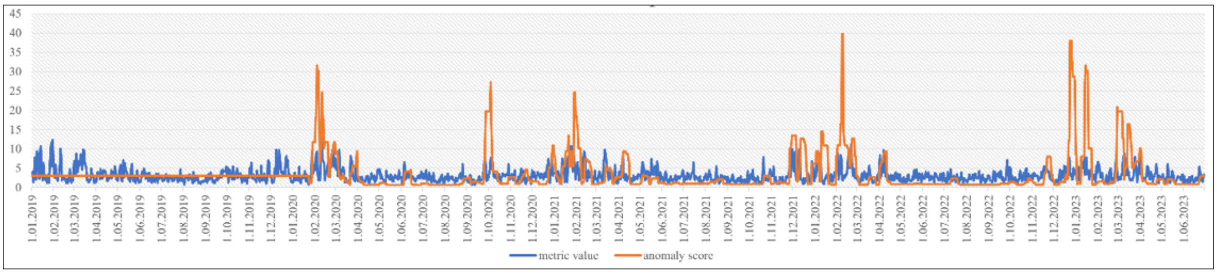

Figure 13. Plot of anomaly scores and metric values for wind speed feature.

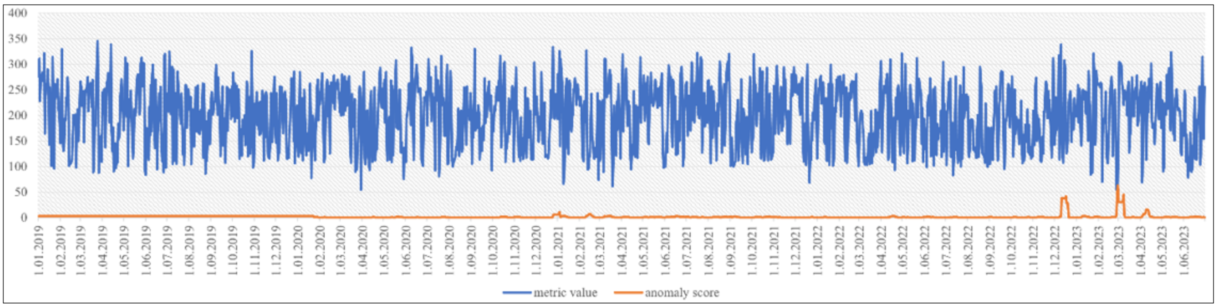

Figure 14. Plot of anomaly scores and metric values for wind direction feature.

Finally, Figure 15 focuses on turbo wind, showing high-level anomalies on December 30, 2022, and January 14, 2023. Mid-level anomalies extend to various dates in late December 2022, revealing anomalous wind behavior.

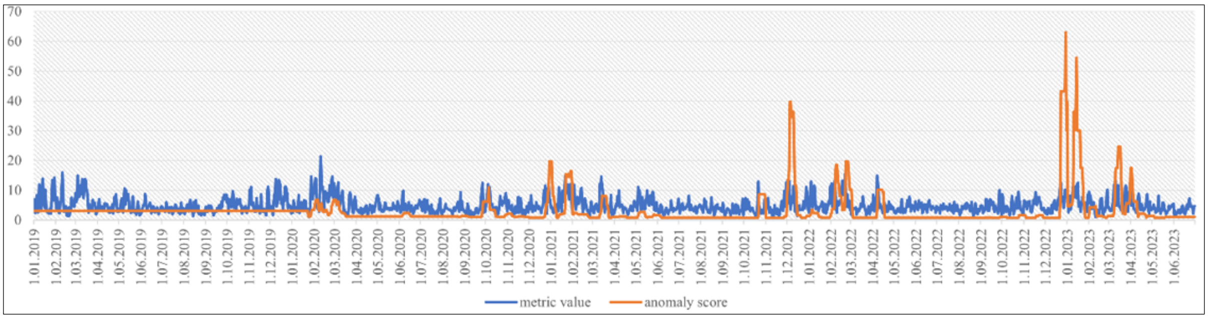

Figure 15. Plot of anomaly scores and metric values for the turbo wind feature.

Table 3 summarizes the high and medium AD results obtained from the HTM-based AD method for various meteorological features. These features include Temperature, Precipitation, Snowfall, Total Cloud, High-level Cloud, Mid-level Cloud, Low-level Cloud, Pressure, Wind Speed, Wind Direction, and Turbo Wind. The table is divided into three columns: "Characteristics," "High-Level Anomaly Dates," and "Medium-Level Anomaly Dates." Anomaly levels are normalized between 0 and 100.

Tablo 3. Dates for high and medium anomalies obtained with the HTM-Based AD approach for each feature.

Features	Anomaly Dates	
	High Level	**Medium Level**
Temparature	20.06.2022	30.07.2022
	20.07.2022	10.08.2022
		6.08.2022
		7.08.2022
		8.08.2022
		9.08.2022
		11.08.2022
		12.08.2022
Rainfall	8.08.2021	7.08.2021
	14.07.2021	15.04.2022
Snowfall	11.02.2021	16.04.2022
	15.01.2021	17.04.2022
	16.12.2022	18.04.2022
	10.04.2022	20.02.2021
Total Cloud	None	6.03.2021
High-level Cloud	28.12.2020	24.12.2020
		27.12.2020
		1.01.2021
		2.01.2021
		25.12.2020
Mid-level Cloud	15.02.2022	28.08.2020
	20.08.2020	7.05.2023
	3.04.2022	6.05.2023
	29.08.2020	13.02.2022
		10.05.2023
Low-level Cloud	1.10.2020	7.12.2020
	18.06.2020	21.06.2020
		22.06.2020
Pressure	None	12.02.2023
Wind Speed	None	None
Wind Direction	28.02.2023	8.03.2023
		17.12.2022
		27.02.2023
Turbo Wind	30.12.2022	29.12.2022
	14.01.2023	23.12.2022
		24.12.2022
		25.12.2022
		26.12.2022
		27.12.2022
		28.12.2022

Thresholds of 60 for high anomaly level and 40 for medium anomaly level were used. The table provides examples of high and medium anomaly dates detected using the HTM-based AD approach for each meteorological feature. These dates are listed in descending order of the anomaly scores calculated for each feature. Notably, some meteorological features have multiple high and mid-level anomaly dates, indicating periods of significant deviation from their expected behavior. In contrast, total cloud, pressure, and wind speed do not have any high-level anomalies recorded during the observed period. However, this does not mean that low-level anomalies did not occur. The findings in Table 4 are a valuable reference for researchers and meteorologists, providing information on when certain meteorological features deviate from their usual patterns. These deviations can be further analyzed to understand the underlying causes and consequences of climate or weather conditions during the periods in question.

4.2.2. Result of Spearman's correlation

This study used Spearman's rank correlation technique to calculate the correlation values. This technique was applied because the characteristics of the data set used were not normally distributed. Unlike Pearson's correlation, Spearman's is a nonparametric measure that assesses monotonic relationships. Spearman correlation is valuable when dealing with data that do not meet linearity assumptions, such as ordinal data or data with outliers. By considering the order of data points rather than their actual values, this correlation technique provides a robust measure of association less sensitive to outliers and the shape of the relationship between variables, making it a versatile tool in various statistical analyses and research fields.

Table 4. Spearman correlation results between the features in the dataset.

Korelasyonlar												
		s1	**s2**	**s3**	**s4**	**s5**	**s6**	**s7**	**s8**	**s9**	**s10**	**s11**
s1	**rho**	1,000										
	sig. (2-tailed)	.										
	n	1642										
s2	**rho**	-,014	1,000									
	sig. (2-tailed)	,584	.									
	n	1642	1642									
s3	**rho**	-,298**	,238**	1,000								
	sig. (2-tailed)	,000	,000	.								
	n	1642	1642	1642								
s4	**rho**	-,259**	,647**	,223**	1,000							
	sig. (2-tailed)	,000	,000	,000	.							
	n	1642	1642	1642	1642							
s5	**rho**	,100**	,344**	-,023	,517**	1,000						
	sig. (2-tailed)	,000	,000	,358	,000	.						
	n	1642	1642	1642	1642	1642						
s6	**rho**	-,083**	,633**	,152**	,814**	,726**	1,000					
	sig. (2-tailed)	,001	,000	,000	,000	,000	.					
	n	1642	1642	1642	1642	1642	1642					
s7	**rho**	-,293**	,658**	,234**	,907**	,249**	,596**	1,000				
	sig. (2-tailed)	,000	,000	,000	,000	,000	,000	.				
	n	1642	1642	1642	1642	1642	1642	1642				
s8	**rho**	-,176**	-,455**	-,102**	-,416**	-,364**	-,527**	-,346**	1,000			
	sig. (2-tailed)	,000	,000	,000	,000	,000	,000	,000	.			
	n	1642	1642	1642	1642	1642	1642	1642	1642			
s9	**rho**	-,266**	,384**	,242**	,441**	,172**	,381**	,436**	-,309**	1,000		
	sig. (2-tailed)	,000	,000	,000	,000	,000	,000	,000	,000	.		
	n	1642	1642	1642	1642	1642	1642	1642	1642	1642		
s10	**rho**	-,113**	,365**	,185**	,445**	,051*	,263**	,524**	-,182**	,319**	1,000	
	sig. (2-tailed)	,000	,000	,000	,000	,040	,000	,000	,000	,000	.	
	n	1642	1642	1642	1642	1642	1642	1642	1642	1642	1642	
s11	**rho**	-,251**	,374**	,180**	,407**	,177**	,375**	,404**	-,288**	,881**	,288**	1,000
	sig. (2-tailed)	,000	,000	,000	,000	,000	,000	,000	,000	,000	,000	.
	n	1642	1642	1642	1642	1642	1642	1642	1642	1642	1642	1642
*. The correlation is significant at the 0.05 level **(2-tailed)** **.The Correlation is significant at the 0.01 level **(2-tailed)** **i ∈ features {1,2,3,4,5,6,7,8,9,10,11} -> Si**												

Table 4 presents the results of the Spearman rank correlations between the variables (s1 to s11) based on the data set of 1,642 data points used in this study. The table, obtained with SPSS software, provides detailed information on the strength of the relationships between variables using nonparametric statistical measures. The table identifies correlations by significance values (Sig.) and rho values. In the table, the significance level is emphasized with asterisks; correlations marked with one asterisk (*) are significant at the 0.05 level, while correlations marked with two asterisks (**) are highly significant at the 0.01 level.

In the correlation table showing the outputs of the analyses, the acceptability, strength, and direction of the relationship between the variables are presented. The correlation coefficient (rho) can take values between -1 and +1. While -1 means that one of the variables increases while the other decreases, +1 means that the variables move in the same direction. If this coefficient is 0, it indicates no relationship between the variables [79]. However, the strength of the relationship is determined by the absolute value of the correlation coefficient [80]. According to the correlation coefficients, researchers consider the degree of relationship between two variables as 0.7 and above as strong and 0.9 and above as very strong. Other levels indicate moderate (rho>0.5) and weak relationships [81].

Table 4 shows significant strong and very strong correlations within the dataset. The table shows a significant relationship between 'total_cloud' and 'min_level_cloud' with a strong positive correlation of 0.814. A strong correlation of 0.907 between 'total_cloud' and 'low_cloud' indicates a strong and positive relationship between these two parameters. In addition, a strong correlation of 0.726 between 'high_level_cloud' and 'mid_level_cloud.' Furthermore, the correlation coefficient 0.881 between 'wind_speed' and 'turbo_wind' indicates a strong relationship between these variables. In terms of moderate correlations, the table shows that there is a positive correlation between "rainfall" and "total_cloud," "mid_level_cloud" and "low_level_cloud," a positive correlation between "total_cloud" and "high_level_cloud," a positive correlation between "mid_level_cloud" and "low_level_cloud," a negative correlation with "pressure" and finally a positive moderate correlation between "low_level_cloud" and "wind_direction."

5. Discussion and Conclusion

This study aims to perform AD tasks on time-indexed data obtained for meteorological events between January 1, 2019, and June 30, 2023, in Kazakhstan, the ninth-largest country in the world. AD tasks are performed using HTM, an unsupervised machine-learning model inspired by the human neocortex. The HTM model is one of the state-of-the-art approaches known for its effectiveness on TS problems. In addition, correlation analyses were performed to determine the strength and direction of monotonic relationships between meteorological features. In correlation analyses, the Spearman method was applied depending on the distribution characteristics of the data. Thus, the correlations between variables were considered for more precise and accurate AD.

According to the results obtained from the experimental studies, ten anomalies were detected in temperature, 4 in precipitation, 8 in snowfall, 1 in total cloud cover, 6 in high-level cloud cover, 9 in medium-level cloud cover, 5 in low-level cloud cover, and 1 in pressure. No anomalies were detected in wind speed. In the wind direction and turbo wind features, 4 and 9 high and medium anomalies were detected, respectively. The HTM model successfully detected 18 high-level and 39 mid-level anomalies. It was concluded that the HTM model detected a total of 57 different anomalies in the detection of anomalies in meteorological conditions. According to the results of Spearman correlation analysis between meteorological features, strong relationships were found between some variables. This shows that meteorological parameters are closely related to each other. The HTM model and correlation analysis results guide researchers in detecting abnormal conditions and effectively detecting high-level and low-level abnormal conditions. However, in the face of limitations such as the univariate rather than multivariate addressing of AD problems, this has been addressed with the help of correlation analyses. The results of the experimental studies show the HTM model's effectiveness, which can detect anomalies in various meteorological features. In this context, the findings show that the HTM model is an effective tool for detecting anomalies in meteorological data and has an effective potential in solving TS problems related to meteorological events.

Future studies are planned to use successful models such as HTM and LSTM to analyze early forecasts of meteorological events. These models have significant potential to process complex data to identify important trends and detect potential anomalies. This will provide a stronger and more effective basis for a better understanding of meteorological events and forecasting future weather events.

References

[1] Friston K and Buzsáki G, “The functional anatomy of time: what and when in the brain,” *Trends in cognitive sciences*, vol. 20, no. 7, pp. 500–511, 2016.

[2] Choi K, Yi J., Park C., and Yoon S., “Deep learning for anomaly detection in time-series data: review, analysis, and guidelines,” *IEEE Access*, vol. 9, pp. 120043–120065, 2021.

[3] Längkvist M., Karlsson L., and Loutfi A., “A review of unsupervised feature learning and deep learning for time-series modeling,” *Pattern Recognition Letters*, vol. 42, no. 1, pp. 11–24, 2014, doi: 10.1016/j.patrec.2014.01.008.

[4] Idrees SM, Alam SM, and Agarwal P., “A prediction approach for stock market volatility based on time series data,” *IEEE Access*, vol. 7, pp. 17287–17298, 2019.

[5] Box G., “Box and Jenkins: Time Series Analysis, Forecasting and Control,” in *A Very British Affair: Six Britons and the Development of Time Series Analysis During the 20th Century*, Springer, 2013, pp. 161–215.

[6] Fu TC, “A review on time series data mining,” *Engineering Applications of Artificial Intelligence*, vol. 24, no. 1, pp. 164–181, 2011, doi: 10.1016/j.engappai.2010.09.007.

[7] Donatelli RE, Park JA, Mathews SM, and Lee JS, “Time series analysis,” *American Journal of Orthodontics and Dentofacial Orthopedics*, vol. 161, no. 4. Elsevier, pp. 605–608, 2022.

[8] Dietterich TG, “Machine learning for sequential data: A review,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2002, vol. 2396, pp. 15–30, doi: 10.1007/3-540-70659-3_2.

[9] Agrawal R., Faloutsos C., and Swami A., “Efficient similarity search in sequence databases,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 1993, vol. 730 LNCS, pp. 69–84, doi: 10.1007/3-540-57301-1_5.

[10] Abonyi J., Feil B., Nemeth S., and Arva P., “Modified Gath-Geva clustering for fuzzy segmentation of multivariate time-series,” *Fuzzy Sets and Systems*, vol. 149, no. 1, pp. 39–56, 2005, doi: 10.1016/j.fss.2004.07.008.

[11] Keogh E. and Ratanamahatana CA, “Exact indexing of dynamic time warping,” *Knowledge and information systems*, vol. 7, pp. 358–386, 2005.

[12] Keogh E. and Pazzani M., “An enhanced representation of time series which allows fast and accurate classification, clustering and relevance feedback,” in *Kdd*, 1998, vol. 98, pp. 239–243, [Online]. Available: http://www.aaai.org/Papers/KDD/1998/KDD98-041.pdf.

[13] Liu G., Zhong K., Li H., Chen T., and Wang Y., “A state of art review on time series forecasting with machine learning for environmental parameters in agricultural greenhouses,” *Information Processing in Agriculture*, 2022.

[14] UA. Bhatti *et al.*, “Time series analysis and forecasting of air pollution particulate matter (PM 2.5): an SARIMA and factor analysis approach,” *Ieee Access*, vol. 9, pp. 41019–41031, 2021.

[15] Saka F., Karaoğlan KM, “Detecting Anomalies in Dam Water Levels using Hierarchical Temporal Memory: A Case Study in Istanbul Province,” in *4th International Symposium of Engineering Applications on Civil Engineering and Earth Sciences 2023 (IEACES2023)*, 2023, pp. 139–150.

[16] Thoppil NM., Vasu V., and Rao CSP., “Deep Learning Algorithms for Machinery Health Prognostics Using Time-Series Data: A Review,” *Journal of Vibration Engineering and Technologies*, vol. 9, no. 6, pp. 1123–1145, 2021, doi: 10.1007/s42417-021-00286-x.

[17] Abanda A., Mori U., and Lozano JA., “A review on distance based time series classification,” *Data Mining and Knowledge Discovery*, vol. 33, no. 2, pp. 378–412, 2019, doi: 10.1007/s10618-018-0596-4.

[18] Li H.and Du T., “Multivariate time-series clustering based on component relationship networks,” *Expert Systems with Applications*, vol. 173, p. 114649, 2021, doi: 10.1016/j.eswa.2021.114649.

[19] Wu J., Zeng W., and Yan F., “Hierarchical Temporal Memory method for time-series-based anomaly detection,” *Neurocomputing*, vol. 273, pp. 535–546, 2018, doi: 10.1016/j.neucom.2017.08.026.

[20] Li G.and Jung JJ., “Deep learning for anomaly detection in multivariate time series: Approaches, applications, and challenges,” *Information Fusion*, vol. 91, pp. 93–102, 2023, doi: 10.1016/j.inffus.2022.10.008.

[21] Ahmed M., Mahmood AN., and Islam MR., “A survey of anomaly detection techniques in financial domain,” *Future Generation Computer Systems*, vol. 55, pp. 278–288, 2016.

[22] Shaukat K. *et al.*, “A Review of Time-Series Anomaly Detection Techniques: A Step to Future Perspectives,” in *Advances in Intelligent Systems and Computing*, 2021, vol. 1363 AISC, pp. 865–877, doi: 10.1007/978-3-030-73100-7_60.

[23] Lin J., Keogh E., Fu A., and Van HH., "Approximations to magic: Finding unusual medical time series," in *Proceedings - IEEE Symposium on Computer-Based Medical Systems*, 2005, pp. 329–334, doi: 10.1109/cbms.2005.34.

[24] Shyu ML., Chen SC., Sarinnapakorn K., and Chang L., "A novel anomaly detection scheme based on principal component classifier," in *Proceedings of the IEEE foundations and new directions of data mining workshop*, 2003, pp. 172–179.

[25] Angiulli F. and Pizzuti C., "Fast outlier detection in high dimensional spaces," in *European conference on principles of data mining and knowledge discovery*, 2002, pp. 15–27.

[26] Hosseinzadeh M., Rahmani AM., Vo B., Bidaki M., Masdari M., and Zangakani M., "Improving security using SVM-based anomaly detection: issues and challenges," *Soft Computing*, vol. 25, no. 4, pp. 3195–3223, 2021, doi: 10.1007/s00500-020-05373-x.

[27] Hu M., Feng X., Ji Z., Yan K., and Zhou S., "A novel computational approach for discord search with local recurrence rates in multivariate time series," *Information Sciences*, vol. 477, pp. 220–233, 2019.

[28] Chandola V., Banerjee A., and Kumar V., "Anomaly detection: A survey," *ACM computing surveys (CSUR)*, vol. 41, no. 3, pp. 1–58, 2009.

[29] Lavin A. and Ahmad S., "Evaluating real-time anomaly detection algorithms--the Numenta anomaly benchmark," in *2015 IEEE 14th international conference on machine learning and applications (ICMLA)*, 2015, pp. 38–44.

[30] Sgueglia A., Di Sorbo A., Visaggio CA., and Canfora G., "A systematic literature review of IoT time series anomaly detection solutions," *Future Generation Computer Systems*, vol. 134, pp. 170–186, 2022.

[31] Terrades OR., Berenguel A., and Gil D., "A flexible outlier detector based on a topology given by graph communities," *Big Data Research*, vol. 29, p. 100332, 2022.

[32] Li C., Mo L., Tang H., and Yan R., "Lifelong condition monitoring based on NB-IoT for anomaly detection of machinery equipment," *Procedia Manufacturing*, vol. 49, pp. 144–149, 2020, doi: 10.1016/j.promfg.2020.07.010.

[33] Kim TY. and Cho SB., "Web traffic anomaly detection using C-LSTM neural networks," *Expert Systems with Applications*, vol. 106, pp. 66–76, 2018.

[34] He Q., Zheng YJ., Zhang CL., and Wang HY., "MTAD-TF: Multivariate time series anomaly detection using the combination of temporal pattern and feature pattern," *Complexity*, vol. 2020, pp. 1–9, 2020.

[35] Audibert J., Michiardi P., Guyard F., Marti S., and Zuluaga MA., "Do deep neural networks contribute to multivariate time series anomaly detection?," *Pattern Recognition*, vol. 132, p. 108945, 2022.

[36] Lindemann B., Maschler B., Sahlab N., and Weyrich M., "A survey on anomaly detection for technical systems using LSTM networks," *Computers in Industry*, vol. 131, p. 103498, 2021.

[37] Ahmed M., Naser MA., and Hu J., "A survey of network anomaly detection techniques," *Journal of Network and Computer Applications*, vol. 60, pp. 19–31, 2016, doi: 10.1016/j.jnca.2015.11.016.

[38] Fernando T., Gammulle H., Denman S., Sridharan S., and Fookes C., "Deep learning for medical anomaly detection--a survey," *ACM Computing Surveys (CSUR)*, vol. 54, no. 7, pp. 1–37, 2021.

[39] Kwon D., Kim H., Kim J., Suh SC., Kim I., and Kim KJ., "A survey of deep learning-based network anomaly detection," *Cluster Computing*, vol. 22, pp. 949–961, 2019.

[40] Aldweesh A., Derhab A., and Emam AZ., "Deep learning approaches for anomaly-based intrusion detection systems: A survey, taxonomy, and open issues," *Knowledge-Based Systems*, vol. 189, p. 105124, 2020.

[41] Rashid ANMB., Ahmed M., Sikos LF., and Haskell-Dowland P., "Anomaly detection in cybersecurity datasets via cooperative co-evolution-based feature selection," *ACM Transactions on Management Information Systems (TMIS)*, vol. 13, no. 3, pp. 1–39, 2022.

[42] Hodge V. and Austin J., "A survey of outlier detection methodologies," *Artificial intelligence review*, vol. 22, pp. 85–126, 2004.

[43] Chen Y., Sui Y., and Shayilan A., "Constructing a high-performance self-training model based on support vector classifiers to detect gold mineralization-related geochemical anomalies for gold exploration targeting," *Ore Geology Reviews*, p. 105265, 2022.

[44] Nassif AB., Talib MA., Nasir Q., and Dakalbab FM., "Machine learning for anomaly detection: A systematic review," *Ieee Access*, vol. 9, pp. 78658–78700, 2021.

[45] Schmidl S., Wenig P., and Papenbrock T., "Anomaly Detection in Time Series: A Comprehensive Evaluation," *Proceedings of the VLDB Endowment*, vol. 15, no. 9, pp. 1779–1797, 2022, doi: 10.14778/3538598.3538602.

[46] Dong B. and Wang X., "Comparison deep learning method to traditional methods using for network intrusion detection," in *2016 8th IEEE international conference on communication software and networks*

(ICCSN), 2016, pp. 581–585.

[47] Kim K. and Aminanto ME., "Deep learning in intrusion detection perspective: Overview and further challenges," in *2017 International Workshop on Big Data and Information Security (IWBIS)*, 2017, pp. 5–10.

[48] Karatas G., Demir O., and Sahingoz OK., "Deep learning in intrusion detection systems," in *2018 International Congress on Big Data, Deep Learning and Fighting Cyber Terrorism (IBIGDELFT)*, 2018, pp. 113–116.

[49] Li J., Izakian H., Pedrycz W., and Jamal I., "Clustering-based anomaly detection in multivariate time series data," *Applied Soft Computing*, vol. 100, p. 106919, 2021.

[50] Ahmad S., Lavin A., Purdy S., and Agha Z., "Unsupervised real-time anomaly detection for streaming data," *Neurocomputing*, vol. 262, pp. 134–147, 2017.

[51] Ma T., Zhu Z., Wang L., Wang H., and Ma L., "Anomaly detection for hydropower turbine based on variational modal decomposition and hierarchical temporal memory," *Energy Reports*, vol. 8, pp. 1546–1551, 2022, doi: 10.1016/j.egyr.2022.02.286.

[52] Soares E., Costa Jr P., Costa B., and Leite D., "Ensemble of evolving data clouds and fuzzy models for weather time series prediction," *Applied Soft Computing*, vol. 64, pp. 445–453, 2018.

[53] Bamaqa A., Sedky M., Bosakowski T., and Bastaki BB., "Anomaly Detection Using Hierarchical Temporal Memory (HTM) in Crowd Management," in *Proceedings of the 2020 4th International Conference on Cloud and Big Data Computing*, 2020, pp. 37–42, doi: 10.1145/3416921.3416940.

[54] "Sample Weather Dataset in Kazakhstan | Kaggle." https://www.kaggle.com/datasets/armanzhalgasbayev/sample-weather-dataset (accessed Sep. 12, 2023).

[55] Cui Y., Surpur C., Ahmad S., and Hawkins J., "A comparative study of HTM and other neural network models for online sequence learning with streaming data," in *2016 International Joint Conference on Neural Networks (IJCNN)*, 2016, pp. 1530–1538.

[56] Hawkins J.and Ahmad S., "Why neurons have thousands of synapses, a theory of sequence memory in neocortex," *Frontiers in neural circuits*, vol. 10, p. 23, 2016, doi: 10.3389/fncir.2016.00023.

[57] Hawkins J., Ahmad S., and Dubinsky D., "Hierarchical temporal memory including HTM cortical learning algorithms," *Techical report, Numenta, Inc, Palto Alto http://www. numenta. com/htmoverview/education/HTM_CorticalLearningAlgorithms. pdf*, 2010.

[58] Wielgosz M. and Pietroń M., "Using Spatial Pooler of Hierarchical Temporal Memory to classify noisy videos with predefined complexity," *Neurocomputing*, vol. 240, pp. 84–97, 2017, doi: 10.1016/j.neucom.2017.02.046.

[59] Ahmad S. and Hawkins J., "Properties of Sparse Distributed Representations and their Application to Hierarchical Temporal Memory," *arXiv preprint arXiv:1503.07469*, 2015, [Online]. Available: http://arxiv.org/abs/1503.07469.

[60] Struye J. and Latré S., "Hierarchical temporal memory and recurrent neural networks for time series prediction: An empirical validation and reduction to multilayer perceptrons," *Neurocomputing*, vol. 396, pp. 291–301, 2020, doi: 10.1016/j.neucom.2018.09.098.

[61] Barua A., Muthirayan D., Khargonekar PP., and Al Faruque MA., "Hierarchical Temporal Memory-Based One-Pass Learning for Real-Time Anomaly Detection and Simultaneous Data Prediction in Smart Grids," *IEEE Transactions on Dependable and Secure Computing*, vol. 19, no. 3, pp. 1770–1782, 2020.

[62] Akila D., Pal S., Jayakarthik R., Chattopadhyay S., and Obaid AJ., "Deep Learning Enhancing Performance Using Support Vector Machine HTM Cortical Learning Algorithm," in *Journal of Physics: Conference Series*, 2021, vol. 1963, no. 1, p. 12144, doi: 10.1088/1742-6596/1963/1/012144.

[63] Attneave F., MB., and Hebb DO., *The Organization of Behavior; A Neuropsychological Theory*, vol. 63, no. 4. Psychology press, 1950.

[64] Hawkins J., "(BAMI) Biological and Machine Intelligence," *Numenta*, p. 64, 2019.

[65] Purdy S., "Encoding data for HTM systems," *arXiv preprint arXiv:1602.05925*, 2016.

[66] Cui Y., Ahmad S., and Hawkins J., "The HTM spatial pooler—a neocortical algorithm for online sparse distributed coding," *Frontiers in Computational Neuroscience*, vol. 11, p. 111, 2017, doi: 10.3389/fncom.2017.00111.

[67] Lai YK.and Nguyen MH., "A Real-Time DDoS Attack Detection and Classification System Using Hierarchical Temporal Memory," *APSIPA Transactions on Signal and Information Processing*, vol. 12, no. 2, 2023.

[68] Zyarah AM and Kudithipudi D., "Neuromorphic Architecture for the Hierarchical Temporal Memory," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 3, no. 1, pp. 4–14, 2019, doi: 10.1109/TETCI.2018.2850314.

[69] Taylor R., “Interpretation of the Correlation Coefficient: A Basic Review,” *Journal of Diagnostic Medical Sonography*, vol. 6, no. 1, pp. 35–39, 1990, doi: 10.1177/875647939000600106.

[70] Warren WG., “Correlation or Regression: Bias or Precision,” *Applied Statistics*, vol. 20, no. 2, p. 148, 1971, doi: 10.2307/2346463.

[71] Asuero AG., Sayago A., and González AG., “The correlation coefficient: An overview,” *Critical Reviews in Analytical Chemistry*, vol. 36, no. 1, pp. 41–59, 2006, doi: 10.1080/10408340500526766.

[72] Ratner B., “The correlation coefficient: Its values range between+ 1/- 1, or do they?,” *Journal of targeting, measurement and analysis for marketing*, vol. 17, no. 2, pp. 139–142, 2009.

[73] Spearman C., “The Proof and Measurement of Association between Two Things,” *The American Journal of Psychology*, vol. 15, no. 1, p. 72, 1904, doi: 10.2307/1412159.

[74] Lehman A. and Rourke NO., “JMP for Basic Univariate and Multivariate Statistics A Step-by-Step Guide,” *Analysis*, p. 146, 2005, Accessed: Sep. 18, 2023. [Online]. Available: http://books.google.com/books?id=1nlApuloc0AC&pgis=1.

[75] Eden SK., Li C., and Shepherd BE., “Nonparametric estimation of Spearman’s rank correlation with bivariate survival data,” *Biometrics*, vol. 78, no. 2, pp. 421–434, 2022, doi: 10.1111/biom.13453.

[76] Agostino RBD., “Tests for the Normal Distribution,” in *Goodness-of-Fit Techniques*, Routledge, 2018, pp. 367–420.

[77] Mohd RN. and Bee WY., “Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests,” *Journal of Statistical Modeling and Analytics*, vol. 2, no. 1, pp. 13–14, 2011.

[78] Yap BW. and Sim CH., “Comparisons of various types of normality tests,” *Journal of Statistical Computation and Simulation*, vol. 81, no. 12, pp. 2141–2155, 2011, doi: 10.1080/00949655.2010.520163.

[79] Kavak B., “Pazarlama araştirmalari tasarim ve analiz.,” *Ankara. Detay Yayıncılık*, 2013.

[80] Gegez AE., *Pazarlama araştırmacıları*. Beta Basım Yayım Dağıtım AŞ., 2007.

[81] Nakip M., *Pazarlama Araştırmaları Teknikler ve Uygulamalar*. Seçkin Yayıncılık, 2006.