International Journal of Sustainable Development and Planning
Vol. 19, No. 9, September, 2024, pp. 3529-3539
Journal homepage: http://iieta.org/journals/ijsdp

Stack-ClimaBoost: A Model for Analysing the Patterns of Global Warming Across the Continents

Saravanan Parthasarathy[1*], Vaishnavi Jayaraman[2], Viji Vijayan[2], Aaswin Raja[1]

[1] Medyaan Healthcare Private Limited, Tambaram, Chennai, Tamil Nadu 600045, India
[2] Datayaan Solutions Private Limited, MEPZ SEZ, Tambaram, Chennai, Tamil Nadu 600045, India

Corresponding Author Email: Saravanan.p@medyaan.com

https://doi.org/10.18280/ijsdp.190921

Received: 23 April 2024
Revised: 7 June 2024
Accepted: 28 June 2024
Available online: 30 September 2024

Keywords:
global warming, continents, climate change, stacking, regression, Stack-ClimaBoost

ABSTRACT

The increasing severity of climate change, including global warming, makes it crucial to act quickly to adapt and reduce its impact. To address the complex issues of climate change, we need to understand how it affects different continents. Traditional methods of predicting climate change often fail due to the inherent complexities and nonlinearities of climate systems. Thus, to overcome these limitations, this study proposes a machine learning-based stacking model. Initially, the state of art models was trained while the results are unsatisfactory, the grid search optimization was employed to improve the results. However, the results produced were in a mediocre state. Thus, a stacking based, Stack-ClimaBoost model was proposed. This model optimizes the integration of Random Forest (RF), CatBoost (CB), and Light Gradient Boosting Machine (LGBM) using grid search optimization. The Stack-ClimaBoost model outperforms previous state-of-the-art models obtaining a low MAPE 0.765, an RMSE of 2.254, and an R^2 value of 0.9003. In addition, for each of the seven continents Stack-ClimaBoost model performed better with a low MAPE (1.8653-5. 8280), an RMSE (1.2-4.57), with a higher R^2 value of (0.65-0.94). With its adaptable solutions, the proposed Stack-ClimaBoost Regressor model could excel in environmental research and climate modeling on multiple continents. Enhancing precision facilitates more dependable prognostications of temperature patterns, thereby supporting proactive strategizing and decision-making aimed at alleviating the consequences of climate change.

1. INTRODUCTION

The specter of global warming looms large, its consequences rippling across the globe in an intricate tapestry of environmental and socioeconomic disturbances [1]. While science is clear in pinpointing human activities as the primary driver of this phenomenon, the predicted impacts and their manifestation are far from uniform across the seven continents. This research delves into the predicted trajectory of global warming and its specific ramifications for each continent, drawing upon the latest scientific consensus and regional case studies. The Earth's climate system, a delicate balance of natural forces, has been irrevocably altered by the relentless pursuit of fossil fuels [2]. The resultant rise in greenhouse gas concentrations, particularly carbon dioxide, has trapped additional heat within the atmosphere, leading to a gradual but inexorable warming trend. This trend, projected to escalate in the coming decades, would not impact the planet homogeneously. Each continent, with its unique geography, ecosystems, and human societies, would experience the consequences of global warming in diverse and profound ways.

Climate change is expected to have a significant impact on Africa, a continent already grappling with chronic poverty and food insecurity [3]. Rising temperatures and unpredictable rainfall patterns threaten agricultural production, exacerbating food shortages and malnutrition. The anticipated exacerbation of water scarcity would further stress communities and ecosystems. Additionally, the melting of glaciers and ice sheets is expected to raise sea levels, leading to coastal flooding and mass displacement of millions of people. Similarly, Asia, home to over half of the global population, faces escalating challenges due to climate change [4]. Increasing frequency and severity of heatwaves, floods, and droughts pose risks to agricultural output and food security. Coastal regions, particularly in Southeast Asia, are highly vulnerable to rising sea levels. Furthermore, shifting precipitation patterns and melting glaciers might disrupt water supplies, potentially exacerbating societal instability.

Europe is increasingly vulnerable to severe impacts of climate change, despite historically experiencing fewer effects [5]. Recent catastrophic weather events, like heatwaves and floods, have caused extensive destruction and casualties. Rising sea levels pose threats to infrastructure and coastal communities. Furthermore, ecosystems and agricultural yields are suffering due to climate change. North America faces widespread challenges from global warming, including rising sea levels from melting Arctic and Greenland glaciers [6]. Increasing frequency and intensity of extreme weather events, such as wildfires, heatwaves, and droughts, adversely affect public health, water resources, and agriculture. The Amazon rainforest, a crucial carbon sink and biodiversity hotspot in South America, is highly susceptible to climate change impacts [7]. Deforestation and rising temperatures exacerbate wildfires, while altered precipitation patterns threaten the delicate rainforest ecosystem. Littoral communities in South America, particularly in Argentina and Uruguay, are increasingly threatened by rising sea levels.

Australia is currently facing severe repercussions from climate change, with increased frequency and intensity of extreme weather events such as heatwaves and droughts resulting in extensive agricultural losses and wildfires [8]. Rising sea levels also pose a threat to infrastructure and coastal communities. Despite its geographic isolation, Antarctica plays a crucial role in regulating the Earth's climate [9]. The continent's accelerated ice sheet disintegration contributes significantly to rising sea levels, posing a global threat. Changes in the Antarctic ecosystem could destabilize global ocean currents and weather patterns, with far-reaching consequences. It is imperative to comprehend the impacts of global warming that are specific to each continent in order to formulate efficacious strategies for adaptation and mitigation.

Due to the intricate and varied consequences of climate change that span multiple continents, conventional approaches might not possess adequate prognostic capability. Hence, in an effort to comprehend and tackle the intricate challenges presented by global warming, we have chosen analysis based on machine learning, which offers a more resilient and flexible methodology. This study aims to comprehensively examine the anticipated ramifications of climate change on all seven continents, focusing on the distinct challenges and vulnerabilities unique to each region. By analyzing regional case studies and utilizing current scientific data, this study seeks to enhance understanding of the global climate crisis and provide insights for formulating strategies that effectively address its complex and geographically diverse consequences.

2. RELATED WORKS

The convergence of machine learning and climate change constitutes a dynamic domain of study that carries significant implications for comprehending and mitigating worldwide environmental issues. There has been a deluge of new research demonstrating novel uses of machine learning algorithms for climate change adaptation, prediction, and mitigation. Recent studies on climate change, global warming, and AI are included in this section.

The significance of ENSO phases in relation to projected global warming and natural disasters has been demonstrated by Derot et al. [10]. Predictive efficacy was highest when the signature method was combined with the LSTM and Lasso models (R^2=0.74 for LSTM, R^2=0.79 for Lasso). To determine the switchover thresholds and the order of the temporal variations, complementary techniques highlighted the influence of climate indices like NINO3, NINO3.4, and NPI on changes in the ENSO cycle. This initial implementation of the signature method to time series data resulted in accurate 6-month forecasts with reduced computing time, demonstrating its potential for climate forecasting improvement. Using a machine learning-based error correction model in conjunction with the PHS_HR approach, Choi et al. [11] attempted to predict the maximum permitted exposure time for outdoor workers in extreme heat. In comparison to the previous method, the multi-layer perceptron (MLP) algorithm enhanced prediction accuracy, achieving a mean absolute error (MAE) of 0.19 minutes instead of 5.05 minutes. Although there may be advantages to occupational health and safety management, there are also certain constraints to consider, such as issues with scalability and generalizability.

Predictions of future climate change and sea level rise by the year 2100 have been estimated using data-driven methodologies [12]. The relationships between greenhouse gas emissions, temperature, and sea level could be tracked, except for the number of sunspots. Adherence to COP26 restrictions has the potential to limit the increase in temperature to 1.88℃, while a rise of 3.28℃ is possible in the absence of emissions reduction. A time series data model covering the years 1961-2020 was developed by Malakouti et al. [13] using data from NASA-GISS to gain a better understanding of and generate more precise forecasts about changes in global temperatures. Several machine learning techniques were studied, including Extra Trees, Light Gradient Boosting Machine (LGBM), Random Forest, K closest neighbors, gradient boosting, and Bayesian Ridge. The execution time of each method and metrics like MAE, MSE, RMSE, R^2, RMSLE, and MAPE were used to evaluate the model's performance. When compared to other algorithms, Extra Trees was shown to be the most effective at predicting the change in global temperature.

An investigation was conducted on climate extremes in China during the period of hiatus in global warming. Machine learning techniques were utilized, with a particular focus on the random forest algorithm [14]. Wet and warm extremes have been on the decline, according to an analysis of precipitation and temperature trends. For trend analysis, regression methods like Theil-Sen are used, which show different trends in the coldest extremes. Long Short-Term Memory (LSTM) outperformed Support Vector Machine (SVM) and Random Forest (RF) in predicting precipitation, with an R^2 value of 0.9, according to an analysis of temperature and precipitation Multi-Model Ensembles (MMEs) using ML techniques [15]. LSTM deep learning models are recommended for their ability to capture nonlinear climatic data correlations. Both RF and LSTM reliably produced high-quality temperature predictions, with R^2 values between 0.82 and 0.93. Due to the unpredictability of General Circulation Models (GCMs), the study emphasizes the significance of precise climate forecasting for water resource management and the value of ensemble techniques in enhancing the dependability of future projections. In addition, different assembly methods were evaluated and enhanced to provide higher performance. It is recommended that RF and LSTM methods be used to create MMEs in the basin, as ML methods generally performed better than mean ensemble approaches.

To forecast emissions of greenhouse gases during the process of automobile lightweighting, various machine learning methods were employed [16]. A comprehensive examination of multiple models, such as decision trees, neural networks, linear regression, and deep learning, was performed. The linear regression model performed exceptionally well. The integration of machine learning techniques and metaheuristic algorithms has been suggested as a novel approach for analyzing the frequency of floods in the context of climate change scenarios [17]. When comparing MARS with M5 Model Tree, a method for classifying precipitation, Multivariate Adaptive Regression Splines (MARS) performed better. With a Nash-Sutcliffe efficiency (NSE) of 0.911, a hybrid method that combines wavelet transform (WT), least square support vector machine (LSSVM), and whale optimization algorithm (WOA) methodology shows higher performance in discharge simulations. Utilizing uncertainty analysis approaches like ANOVA and fuzzy logic, the study delves further into discharge modeling across many future scenarios.

Berrang-Ford et al. [18] employed machine learning methods, the convergence of climate change and health literature is accurately delineated, uncovering significant emphasis on impact assessments, air quality, and thermal stress. By utilizing supervised and unsupervised machine learning techniques, substantial research deficiencies in the fields of maternal health, undernutrition, and mental health were identified. With limited data from low-income nations, studies are geographically concentrated in high-income countries and China. The study highlights the relevance of measures for adaptation and mitigation of climate change for human health, and it forecasts that there will be many publications on the topic between 2013 and 2019. Using Semantic Web techniques, the Fuzzy Cognitive Maps (FCMs) model was established to analyze the effects of global warming [19]. It presents a Semantic Web simulation tool for scenario predictions and centers on modeling causal links using FCMs. Analyzing the density reveals how complicated the model is, while sensitivity tests show that the equilibrium is independent of the initial values.

Through the integration of data on energy consumption and economic growth, this research presents a multi-stage methodology for predicting carbon dioxide emissions in Group 20 nations [20]. It implements artificial neural networks, self-organizing maps, and adaptive neuro-fuzzy inference systems by utilizing clustering, machine learning, and dimensionality reduction techniques. The outcomes demonstrate a substantial enhancement in precision when compared to established methodologies; the proposed strategy attains a mean average error of 0.065. The potential of the singular value decomposition-self-organizing map-adaptive neuro-fuzzy inference system to provide valuable insights for energy and economic policymaking is made clear by the method's superiority. Nevertheless, it is important to consider constraints such the lack of generalizability and validation in varied circumstances. To reduce computing costs and speed up climate change estimates, machine learning is crucial [21]. Ridge regression and Gaussian Process Regression are some of the methods used to map and forecast regional temperature variations over the long term. The significance of sharing data extensively is illustrated by data-driven climate modeling, which also includes methodologies such as Random Forest and Lasso to improve projections when compared to previous approaches.

Amidst global warming, Gholami Rostam et al. [22] compared optimization algorithms on Multilayer Perceptron (MLP) networks for precipitation forecasting, with an emphasis on water resource management in Iran. The Integrated Model Framework (IMF)'s MLP-PSO was shown to be the most successful for precipitation forecasting out of three optimization algorithms tested on backpropagation-based MLP models. MLP-PSO yielded results for the two assessment sets of 21.21 and 18.54 RMSE, 15.12 and 12.97 MAE, 19.17 Z-test, and 13.69 and 18.54 Taylor diagram scores, respectively. Utilizing ML algorithms including Artificial Neural Network, K-Nearest Neighbor, Support Vector Machine, and Relevance Vector Machine, the multi-model ensemble (MME) was able to zero in on temperature and precipitation forecasts [23]. It evaluates kernel functions to improve MME performance, with the best result being a md value of 0.889 for SVM-based MME. Furthermore, ML methods are used to assess GCMs and MMEs by comparing median md values for various ML techniques. During validation, which determines the performance of GCMs in simulating precipitation, Tmax, and Tmin, KNN-based MMEs show the greatest md value of 0.780, whereas RVM-based MMEs offer better median values.

Kalra et al. [24] used machine learning methods to investigate greenhouse gas relationships, with a focus on carbon dioxide and its major role in the warming trend. When mean square error analysis was used to compare ANN models to linear regression, decision tree regression, random forest regression, and other models, ANN showed that it was the best. Using ANN with three-layer topologies, Adam optimization, and ReLU activation function, the study determines that carbon dioxide is the most significant greenhouse gas. The Mean Square Error (MSE) value obtained in the analysis is 0.0078. The use of LSTM allowed for the prediction of GHG emissions from transportation systems [25]. Predictors such as density, speed, and GHG emission rates (ER) were set up in the LSTM networks. Results showed that the best performance came from the LSTM model that took in-links speed, density, GHG ER, and speed into account. Consistent comparisons with ARIMA and clustering models showed that LSTM consistently performed better than others. Furthermore, the most important model predictors were determined by correlation analysis, and GHG ER prediction was made more accurate through simulation utilizing real-time traffic data. The constructed LSTM model outperformed the clustering and ARIMA models, demonstrating excellent performance with an R^2 value of 0.767 and an RMSE of 0.362.

3. RESEARCH GAP

A significant void in the existing body of knowledge regarding climate change is the absence of thorough and geographically targeted assessments that address the effects of global warming on each of the seven continents. Research on the effects of climate change is sparse when it comes to continent-by-continent investigations of temperature trends, variability, and predictive modeling. Most of the extant literature focuses on global or regional scales. Because climate dynamics and susceptibility to global warming are impacted by various ecological, social, and geographical factors on different continents, this disparity is especially noteworthy. Researchers may gain a better understanding of the individual threats and possibilities presented by climate change in each continent by performing continent-specific analyses, which will allow for the development of more precise and efficient adaptation and mitigation plans. Policymakers, stakeholders, and communities could benefit greatly from this approach because it helps fill in the gaps in our understanding of the climate crisis's complex character by shedding light on the ways in which its effects vary between continents. To further our knowledge of climate change and improve resilience worldwide, it is essential to conduct specialized studies on each continent to close this knowledge gap.

4. DATASET DESCRIPTION

The dataset used in this study was from Kaggle which has the annual surface temperature data for all countries spanning the years 1961 to 2022 [26]. The temperature data was sourced from reputable climate monitoring agencies and research institutions, ensuring its reliability and accuracy. The dataset was analyzed for the null value and the columns such as area code, area code (M49), area, months code, months, unit were removed from the dataset. In the element column the standard deviation part was eliminated. Thus, the final dataset contains the year and the temperature change as shown in Table 1.

Table 1. Global warming dataset

S. No.	Attributes	Description of the Attributes
1	Year	Year of the study conducted
2	Temperature	Temperature changes

5. EXPLORATORY ANALYSIS

In this research, the global warming dataset of the seven continents were only considered. Except for Asia, Africa, North America, South America, Europe, Australia and Antarctica, the other countries, and irrelevant data, which does not affect the result in any way were removed. The yearly and monthly dataset was examined in this study. Figure 1 depicts the temperature changes over the years for the seven continents.

5.1 Asia

The temperature in Asia has exhibited a discernible upward trend over the years, as indicated in Figure 2. Throughout the 20th century, Asia experienced significant temperature fluctuations, reflecting the complex interplay of natural climatic variability and human influences. The beginning of the 21st century marked an escalation in the frequency and intensity of extreme weather events across the continent, underscoring the challenges posed by climate change.

Over the past decade, Asia has confronted numerous temperature-related challenges, with instances of both peaks and declines in temperature. For instance, there was a gradual increase in temperature until 1963, followed by a decline in 1964 (refer to Figure 2). Subsequent years saw fluctuations, including peaks in 1966 and 1980, and declines in 1947 and 1967, with the latter representing the lowest temperature recorded within the selected study period. The temperature trend continued with varying highs and lows, reflecting the complex dynamics of climate variability. Especially, during the COVID-19 pandemic, there was a temporary decline in temperatures, followed by a peak in 2021 and 2022. Throughout the chosen study period, the temperature reached its peak in 2021.

The climatic pattern in Asia traditionally featured April, May, and June as the warmest months, characterized by average temperatures ranging from 20 to 30 degrees Celsius. Conversely, December, January, and February were typically the coldest months, with average temperatures spanning from 0 to 10 degrees Celsius. However, this climatic norm has undergone a shift, with February emerging as the warmest month in recent times as depicted in Figure 3. This alteration suggests a significant deviation from historical temperature trends. Furthermore, the temperature variation between the warmest and coldest months varies across different regions of Asia.

In the northern and central parts of the continent, this difference could be substantial, reaching up to 50 degrees Celsius. Conversely, in the southern regions, such as Southeast Asia and parts of the Indian subcontinent, the temperature difference tends to be more modest, approximately 20 degrees Celsius. This changing climate pattern underscores the dynamic nature of Asia's climate system and highlights the impact of global warming and other environmental factors on regional weather patterns. Understanding these shifts is crucial for adaptation strategies and mitigating the potential consequences of climate change in the region.

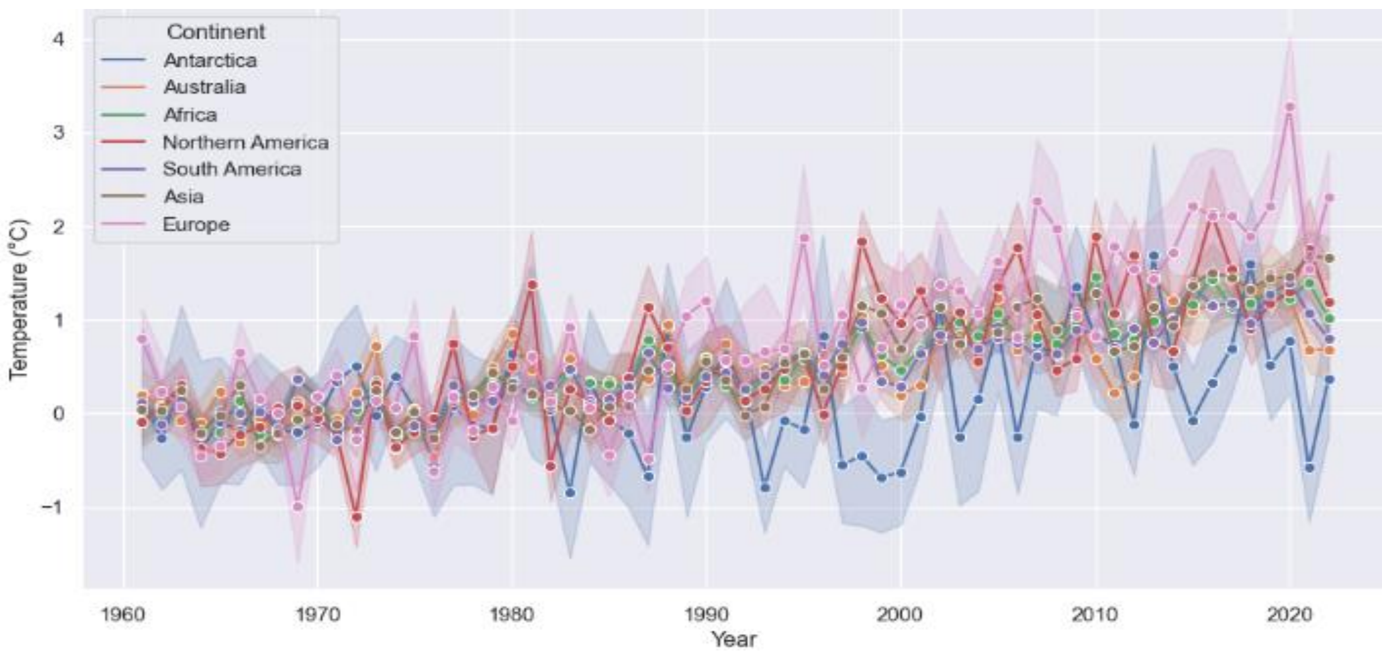

Figure 1. Overall temperature range across seven continents

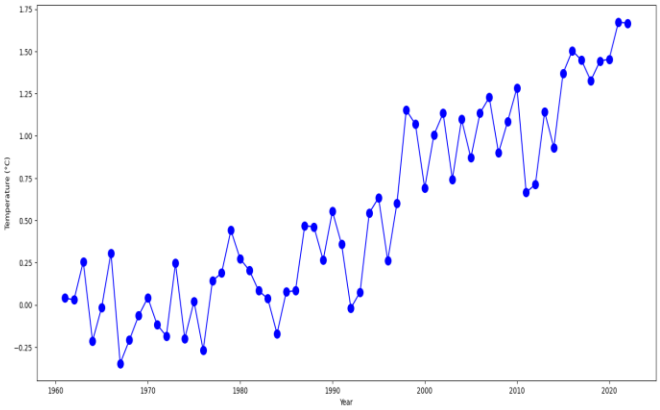

Figure 2. Asia-temperature change over the years

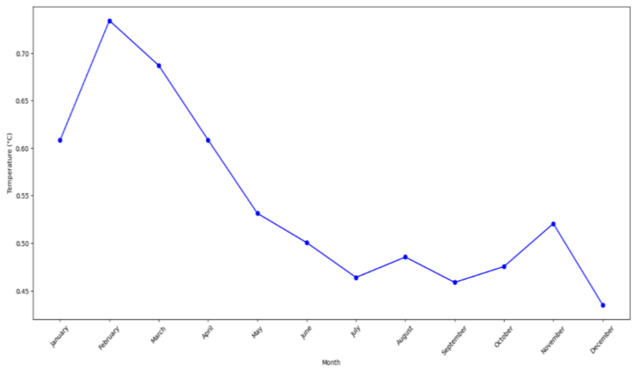

Figure 3. Asia-temperature change over the months

5.2 Africa

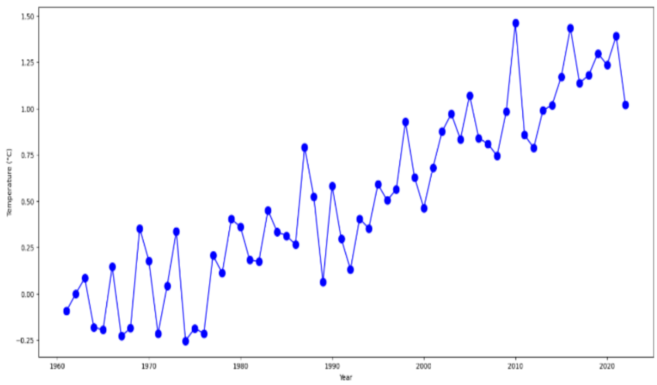

Figure 4. Africa-temperature change over the years

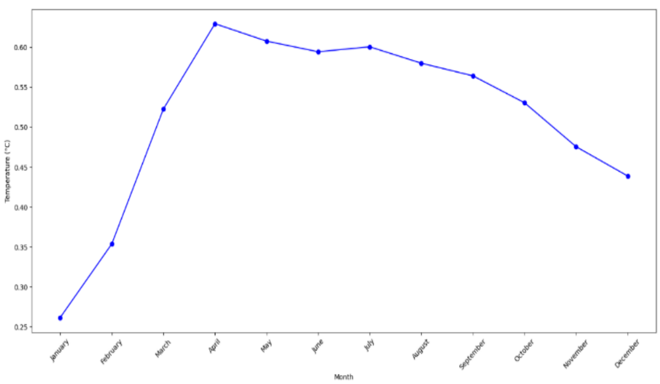

Figure 5. Africa-temperature change over the months

Figure 4 illustrates the average temperature trends in Africa over time. Since 1960, there has been a notable increase of approximately 1.5 degrees Celsius in the average temperature of the continent. In 1974, Africa experienced an extreme low in climate, while by 2010, it reached its maximum temperature. Throughout the years, Africa's temperature has fluctuated, showing both ups and downs. Remarkably, there was a peak in 2016 followed by another in 2021. However, neither peak surpassed the temperature recorded in 2010.

The average temperature over the months in Africa is shown in Figure 5. The graph shows that the hottest months in Africa are April, May, and June, with average temperatures ranging from 23 to 27 degrees Celsius. The coldest months in Africa are June, July, and August, with average temperatures ranging from 15 to 19 degrees Celsius.

5.3 Northern America

In the 1960s, North America exhibited a higher temperature range compared to other continents, indicative of its diverse climate zones as depicted in Figure 6. The lowest recorded temperature occurred in 1972, while the highest was logged in 2016, reflecting a trend of increasing temperatures over the years. As of the recent year 2022, North America's temperature has stabilized within its average range, suggesting a relative climate stability in the region. An interesting seasonal pattern emerges in North America's temperature fluctuation.

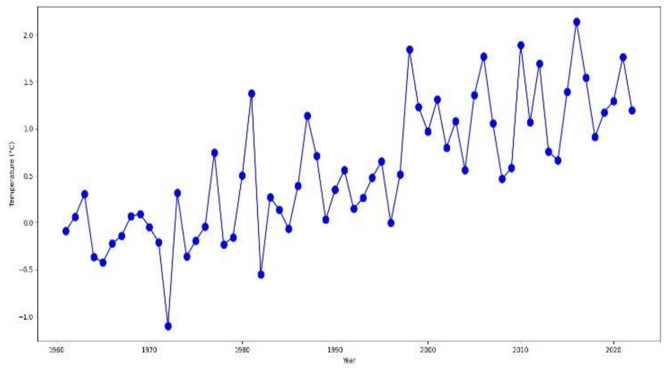

Figure 6. Northern America-temperature change over the years

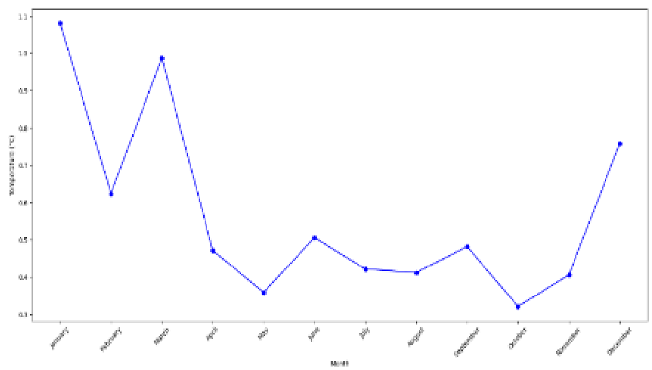

Figure 7. Northern America-temperature change over the month

January emerges as the hottest month, followed by a decline in temperature in February, which closely resembles December's temperatures, marking a transition from peak winter conditions to early spring. March sees a subsequent rise in temperature, signaling the onset of warmer weather (refer to Figure 7). Conversely, the remaining months generally experience lower temperatures, with October noted as the coolest month. This cooling trend aligns with the onset of autumn in North America, characterized by shorter days, decreasing sunlight, and cooler temperatures. This seasonal temperature fluctuation underscores the dynamic nature of North America's climate and the interplay of various meteorological factors throughout the year.

5.4 South America

As illustrated in Figure 8, the temperature range in North America exhibits temporal variability, with distinct patterns observed over the years. The lowest recorded temperature occurred in 1971, indicative of periodic fluctuations in the region's climate, while the highest temperature was documented in 2020, suggesting a trend of warming over time. February consistently emerges as the month with the lowest temperature when analyzed between 1961 and 2022, reflecting the peak of winter conditions in North America (refer to Figure 9). In contrast, the months of April and October consistently exhibit peak temperatures, indicative of transitional periods between seasons and the onset of warmer weather. Especially, North America experiences its coolest month in October, signifying the onset of autumn and the transition to colder temperatures. This cooling trend aligns with the changing seasons and the decreasing daylight hours characteristic of the fall season. In South America, the temperature pattern follows a different trajectory, characterized by a high followed by a low. This alternating pattern suggests a fluctuating climate dynamic, influenced by various regional factors such as ocean currents, elevation, and atmospheric circulation patterns.

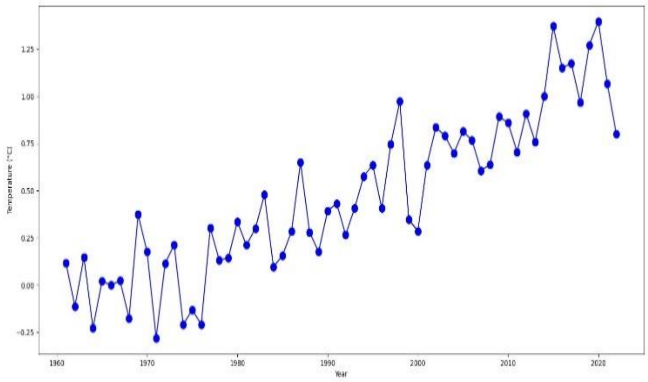

Figure 8. South America-temperature change over the years

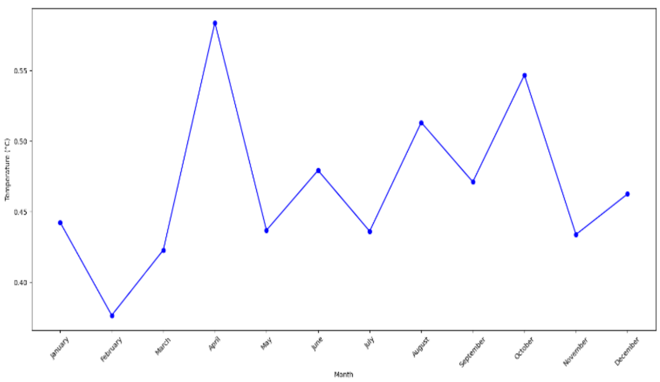

Figure 9. South America-temperature change over the month

5.5 Antarctica

Over the span of 121 years, Antarctica has experienced a notable increase in its average temperature, rising by approximately 1.2 degrees Celsius. Figure 10, emphasized by recorded data, with the lowest temperature documented in 1983 and the highest in 2013, indicating a clear trajectory of warming temperatures over time. Despite this overall warming trend, recent observations as of 2022 suggest a stabilization of temperatures within the region's average range, hinting at a relative climate stability in Antarctica.

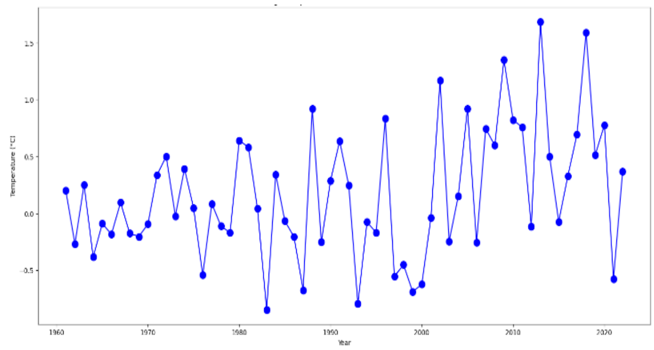

Figure 10. Antarctica-temperature change over the years

Antarctica's climate is characterized by two distinct seasons: summer and winter. The summer season spans from October to February, during which the sun remains continuously in the sky. However, intriguingly, despite the persistent sunlight, February emerges as the month with the lowest temperatures over the examined sixty-one-year period, as depicted in Figure 11. This anomaly presents an interesting aspect of Antarctic climate dynamics.

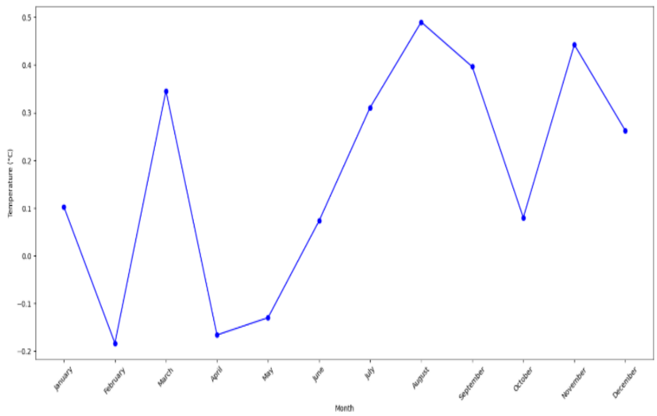

Figure 11. Antarctica-temperature change over the month

Contradictory to traditional expectations, August, typically associated with the coolest temperatures in Antarctica, appears to showcase the highest temperatures according to the data. This departure from expected seasonal norms underscores the complexity of Antarctic climate patterns and highlights the importance of continuous monitoring and analysis to understand the dynamics at play.

Overall, while Antarctica has experienced a discernible warming trend over the past century, recent indications of temperature stabilization alongside curious deviations from expected seasonal temperature patterns underscore the intricate nature of climate processes in this region.

5.6 Europe

Figure 12 illustrates the average annual temperature in Europe from 1960 to 2022, revealing a noteworthy trend of increasing temperatures over the decades. The warmest years recorded in Europe all fall within the period since the 1990s. In 1960, the average temperature hovered around 0.8℃. Conversely, 1970 marked one of the coldest years in Europe, with temperatures plummeting to approximately -0.5℃. The pinnacle of this warming trend occurred in 2010, designated as the warmest year in Europe during the specified timeframe, boasting an average temperature of approximately 2.5℃. By 2019, the average temperature surged to around 3℃, reflecting a remarkable increase of approximately 4℃ over the span of six decades. Despite noticeable fluctuations between individual years, the overarching pattern delineates a consistent elevation in Europe's average temperature.

Europe stands as the fastest warming continent globally, a phenomenon predominantly attributed to anthropogenic activities. Human-induced factors such as greenhouse gas emissions have significantly contributed to this warming trend, accentuating the urgency for mitigation and adaptation measures. However, a discrepancy arises when considering the climatic norms associated with specific months in Europe. While September traditionally represents the coolest month, the data portrays January as the coldest month. This inconsistency is particularly pronounced in northern Europe, where January typically ranks as the coldest month of the year.

Conversely, the warmest months in Europe typically span from May to September, with July and August traditionally considered the hottest (as shown in Figure 13). Nonetheless, the depicted climatic patterns reveal a deviation from the norm, with October and November emerging as the warmest months, and March assuming the title of the hottest month in Europe.

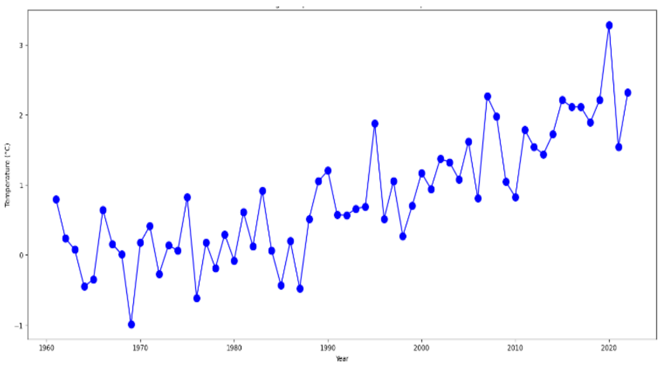

Figure 12. Europe-temperature change over the years

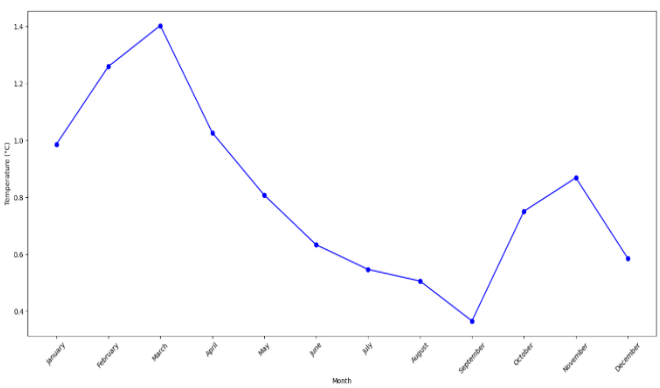

Figure 13. Europe-temperature change over the month

5.7 Australia

Figure 14 depicts a discernible upward trend in temperature spanning the past 68 years, with an average increase of approximately 1.49 degrees Celsius. Notably, Australia experienced an extreme low in climate around 1976, while reaching its peak temperatures in 2012 and 2018. Over time, Africa's temperature has exhibited fluctuations, characterized by both upward and downward trends. In Australia, the warmest years on record have all occurred since the 1990s, indicating a pronounced warming trend in the region. Despite this overarching trend, discrepancies emerge when comparing actual seasonal temperature norms with those depicted in the graph.

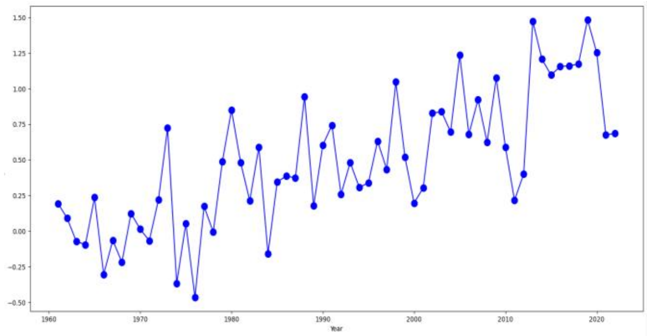

Figure 14. Australia-temperature change over the years

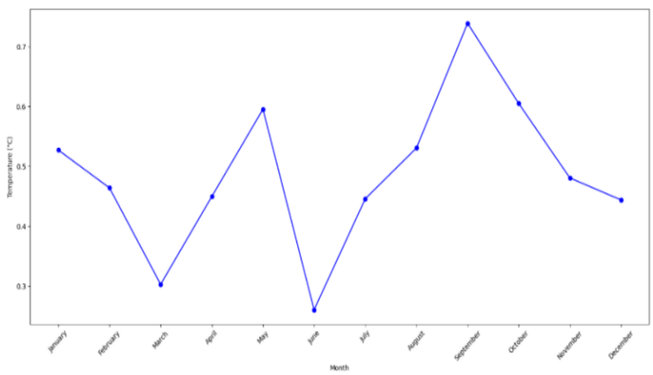

Figure 15. Australia-temperature change over the months

According to traditional climate patterns, December to February constitutes the summer season in Australia, while March to May marks autumn, June to August signifies winter, and September to November represents spring. However, Figure 15 presents deviations from these norms. For instance, while July typically registers as the coldest month in Australia, with daytime temperatures dropping as low as 12 degrees Celsius, the graph suggests June as the coolest month. Similarly, the graph indicates April, July, and December as the warmest months in Australia, contradicting the actual warmest months, which are December, January, and February. Furthermore, discrepancies extend to the identification of the hottest month in Australia. While December, January, and February traditionally occupy this status, the graph suggests September as the hottest month.

The comprehensive analyses of temperature fluctuations for all seven continents offer significant insights into the intricate nature of climate dynamics on a global scale. Each region has its own patterns that are affected by different things. For instance, the temperature of Asia has been increasing with significant fluctuations, however, there was a short-term cooling effect during the COVID-19 pandemic. Similarly, the warmth of Africa has been rising since the 1960s. Temperatures in North America have risen steadily over time, exhibiting distinct seasonal variations, whereas South American temperature trends are subject to temporal variability, with February consistently being the coldest month. Despite a century-long warming trend, Antarctica illustrates peculiar deviations from anticipated seasonal patterns. Europe is distinguished as the continent experiencing the most rapid global warming, as evidenced by significant disparities between recorded and customary seasonal temperatures. In contrast, Australia reveals a noticeable progressive in temperature, although there are inconsistencies between the projected trends and the observed seasonal averages. These analyses emphasize the significance of comprehending regional climate variations in order to develop effective strategies for mitigating and adapting to climate change. Further accentuate the necessity for ongoing monitoring and analysis in order to precisely track climate trends.

6. PROPOSED METHODOLOGY

Initially, the dataset underwent preprocessing to focus solely on temperature data across seven continents and various years. Categorical variables representing continents were encoded using label encoding, ensuring compatibility with machine learning algorithms. Subsequently, the dataset was split into training (70%) and testing sets (30%). State-of-the-art machine learning models were employed on the preprocessed dataset. Despite their implementation, the algorithms demonstrated mediocre performance. These algorithms were subsequently applied with grid search optimization to enhance their accuracy [27, 28]. However, the model yielded inadequate results. Therefore, to enhance predictive performance, a novel approach called the Stack-ClimaBoost Regressor model was proposed [29, 30]. This model employed column stacking, amalgamating predictions from grid search optimized LGBM and CB regressors. Additionally, the optimized RF was utilized as the meta-regressor. The overall architecture flow of the proposed Stack-ClimaBoost model is illustrated in Figure 16.

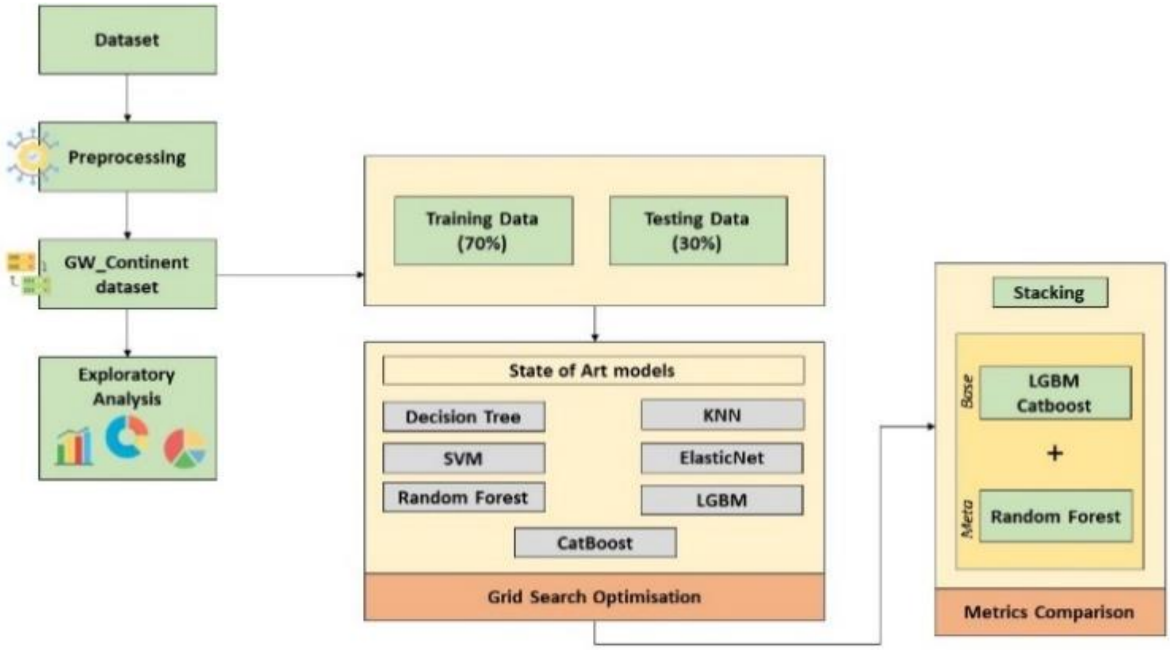

Figure 16. Flow diagram-Stack-ClimaBoost model

6.1 Evaluation metrics

The following metrics were chosen since a lower MAPE indicates higher predictive capability, which is crucial for precise applications like climate change prediction. A lower RMSE signifies smaller deviations, essential for reliable forecasts in public health and infrastructure planning. A higher R^2 shows the model explains more variance, enhancing long-term climate planning reliability. Interpretation from these metrics ensures informed decisions and better climate-related preparation.

6.1.1 Mean absolute percentage error (MAPE)

It measures the accuracy of a model by calculating the average absolute percentage error between predicted and actual values. It is expressed as a percentage, making it easy to interpret.

$$\text{MAPE} = \frac{1}{\text{n}}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| 100$$

6.1.2 Root mean squared error (RMSE)

RMSE measures the standard deviation of the residuals. It indicates how spread out the residuals are and is sensitive to large errors.

$$\text{RMSE} = \sqrt{\frac{1}{\text{n}}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

6.1.3 Coefficient of determination (R^2)

R^2 measures the proportion of variance in the dependent variable that is predictable from the independent variables. It ranges from 0 to 1, with 1 indicating perfect prediction.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}$$

7. DISCUSSION

The preprocessed dataset was implemented with various machine learning models to analyze complex datasets and make accurate predictions. In this study, the following algorithms were employed Decision Tree (DT), K Nearest Neighbor (KNN), Support Vector Machine (SVM), ElasticNet, Random Forest (RF), CatBoost (CB), and Light Gradient Boosting (LGBM). The rationale for choosing these algorithms is that the DT offers flexibility in analysing the time series data while the KNN algorithm is a non-parametric method that is useful for identifying patterns based on historical data. Similarly, the SVM could handle if there are any non-linear relationships in the data. ElasticNet model was chosen as it combines the L1 and L2 regularizations which helps to manage the multicollinearity in time series data. RF, an ensemble technique improves accuracy by averaging multiple decision trees. CB algorithm works well without extensive preprocessing and works effectively. Similarly, LGBM is optimized for speed and efficiency which Handles complex data structures. However, the results obtained from these models revealed varying degrees of performance, with most models exhibiting suboptimal predictive accuracy.

Upon initial evaluation, it was observed that several models, such as DT, KNN, and SVM, yielded relatively higher Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE) values, indicating poor predictive performance (refer to Table 2). Thus, to improve the predictive performance a novel method was proposed.

Table 2. Results-state of art algorithms

Regressors	MAPE	RMSE	R^2 Score
Decision Tree	269.1804	6.7045	0.3826
K Nearest Neighbor	191.9477	6.4733	0.4245
Support Vector Machine	218.1056	6.2319	0.4666
ElasticNet	200.7479	6.2205	0.4685
Random Forest	166.3474	6.0046	0.5048
CatBoost	167.4243	5.5148	0.5823
LGBM	144.4388	5.4922	0.5857

To improve the results obtained the Grid search optimization was utilized. Each algorithm was hyperparameterized according to its parameters as shown in Table 3. The results were improved to a greater extent as shown in the table. The worst performer is SVM with a higher RMSE of 5.09 and MAPE of 3.67. Similar to the performance of the state-of-art algorithms, the best-performing models based on the R^2 metric are Light Gradient Boosting (0.6707), CatBoost (0.652), and Random Forest (0.6422). These models have the highest coefficient of determination, indicating a better fit to the data compared to other models as portrayed in Table 4.

The best performers with higher R^2 value among the chosen algorithms such as RF, CB, and LGBM were stacked together. LGBM and CB were stacked along with the RF as the meta-regressor. When compared with Table 2 and Table 3, the proposed model outperformed the other traditional and optimized models. This Stack-ClimaBoost model outperformed the other existing models [10-12] with the R^2 value of 0.9003, RMSE of 2.254, and MAPE of 0.765. The proposed Stack-ClimaBoost Regressor model showcased a significant improvement in predictive performance compared to individual models like Random Forest, CatBoost, and Light Gradient Boosting. The novel approach of stacking models, combined with a meta-regressor, proved to be an effective strategy for leveraging the strengths of multiple algorithms and mitigating their individual weaknesses. The study underscores the importance of rigorous optimization techniques like grid search in fine-tuning machine learning models to achieve optimal performance. The Stack-ClimaBoost Regressor model improved its predictive capability and shows great potential for use in climate modeling. Further, this model performed the best for all the continents individually as shown in Table 5.

Table 3. Best parameter value-grid search optimized

Algorithm	Hyperparameter	Best Value Obtained
KNN	n_neighbors	10
	weights	uniform
	algorithm	ball_tree
Decision Tree	max_depth	5
	min_samples_split	5
	min_samples_leaf	4
Support Vector Machine	kernel	linear
	C	10
	epsilon	0.1
ElasticNet	alpha	0.5
	l1_ratio	0.1
	max_iter	1000
Random Forest	n_estimators	100
	max_depth	20
	min_samples_split	5
	min_samples_leaf	4
CatBoost	n_estimators	100
	max_depth	6
	learning_rate	0.05
	l2_leaf_reg	1
LGBM	n_estimators	100
	max_depth	8
	learning_rate	0.1
	num_leaves	31

Table 4. Results of grid search optimized models

Regressors	MAPE	RMSE	R^2 Score
Decision Tree	144.6274	4.5413	0.5959
K Nearest Neighbor	137.9761	4.8632	0.5365
Support Vector Machine	158.0760	5.0963	0.491
ElasticNet	151.8972	5.0540	0.4995
Random Forest	133.9582	4.2731	0.6422
CatBoost	121.7677	4.2144	0.652
LGBM	131.9933	4.0992	0.6707

Table 5. Continent wise-results of Stack-ClimaBoost model

Continents	MAPE	RMSE	R^2 Score
Asia	3.0834	1.95	0.895
Antarctica	5.8280	2.95	0.793
Australia	3.1457	2.30	0.921
Africa	2.0680	1.25	0.933
Northern America	6.3047	4.57	0.65
South America	1.8653	1.20	0.925
Europe	4.6284	1.72	0.944

The analysis of temperature prediction across different continents reveals varying degrees of accuracy and model performance. Africa and Australia emerge as regions with the most accurate predictions, characterized by low MAE and RMSE values, indicating precise temperature forecasts with minimal error. Additionally, these continents exhibit exceptionally high R^2 scores, indicating excellent model fit and explaining a significant portion of the variance in temperature data. Europe follows closely behind, demonstrating slightly higher error metrics but still showcasing strong predictive capability, as evidenced by its high R^2 value. However, Northern America displays the highest error metrics, indicating less accurate predictions compared to other continents. Despite this, the model still achieves a moderate fit to the data, as indicated by its R^2 score of 0.65.

South America showcases accurate predictions with low MAPE and RMSE values, with an R^2 score of 0.925, indicating a strong fit of the model to the data. In contrast, Asia and Antarctica demonstrate slightly higher error metrics but still maintain good model fit, as reflected by their respectable R^2 scores. These findings underscore the importance of region-specific climate modeling efforts and adaptation strategies tailored to the unique characteristics of each continent. Additionally, they highlight the effectiveness of machine learning algorithms in predicting temperature trends across diverse geographic regions, providing valuable insights for climate research and policymaking initiatives aimed at addressing global warming and its associated impacts.

Further, for the overall dataset, the proposed StackBoost model yielded a higher R^2 value of 0.9003 and lower error rates, with an RMSE of 2.254 and a MAPE of 0.765. The R^2 value of 0.9003 signifies that 90.03% of the variance in temperature data is explained by the model. An RMSE of 2.254 indicates that the average deviation of the model's predictions from the actual temperatures is quite small. This low MAPE value suggests that the model's predictions are highly accurate, with an average error of only 0.765%. Thus, the analysis emphasizes the importance of region-specific climate modeling and adaptation strategies, informed by machine learning predictions, to address global warming effectively. This also enhances trust in the model for practical applications like urban planning and disaster management, where accurate temperature predictions can significantly influence decision-making.

8. CONCLUSION

The imminent hazard posed by global warming compels experts to explore viable strategies for adaptation and mitigation. This study delves into the effects of climate change on different continents and offers solutions specifically designed to tackle the complex issues it poses. When compared to other state-of-the-art models, the proposed Stack-ClimaBoost model performed better. The Stack-ClimaBoost model integrates the top three methods, RF, CB, and LGBM, along with grid search optimized parameters. By combining RF as the meta-regressor and stacking CB and LGBM, the proposed model surpassed other cutting-edge models, achieving an impressive R^2 value of 0.9003, RMSE of 2.254, and low MAPE. The Stack-ClimaBoost Regressor model holds promise for applications in environmental research and climate modeling due to its enhanced prediction accuracy.

Additionally, the model demonstrated effectiveness for each continent separately, showcasing varying levels of precision and performance. The Stack-ClimaBoost model had obtained an R^2 score of 0.65 to 0.94 for the seven continents with a minimal error rate. For future work, integrating supplementary climatic, geographical, and socioeconomic variables would allow for a more comprehensive spectrum of factors impacting temperature variability and trends to be captured. This could facilitate more robust climate predictions and adaptation strategies, enabling stakeholders to better understand and address the challenges posed by global warming on a regional and global scale. Such advancements in climate modeling and prediction are crucial for informing evidence-based decision-making and implementing proactive measures to mitigate the impacts of climate change.

REFERENCES

[1] Nial, C. (2023). The inseparable alliance of mitigation and adaptation for global public health. Medium. https://medium.com/purpose-and-social-impact/the-inseparable-alliance-of-mitigation-and-adaptation-for-global-public-health-792ba1b8cd09.

[2] Mélières, M.A., Maréchal, C. (2015). Climate Change: Past, Present, and Future. John Wiley & Sons.

[3] Semenza, J.C., Paz, S. (2021). Climate change and infectious disease in Europe: Impact, projection and adaptation. The Lancet Regional Health-Europe, 9. https://doi.org/10.1016/j.lanepe.2021.100230

[4] Ripple, W.J., Wolf, C., Gregg, J.W., Rockström, J., Newsome, T.M., Law, B.E., Marques, L, Lenton, T.M., Xu, C., Huq, S., Simons, L., King, S.D.A. (2023). The 2023 state of the climate report: Entering uncharted territory. BioScience, 73(12): 841-850. https://doi.org/10.1093/biosci/biad080

[5] Degroot, D., Anchukaitis, K., Bauch, M., Burnham, J., Carnegy, F., Cui, J., de Luna, K., Guzowski, P., Hambrecht, G., Huhtamaa, H., Izdebski, A., Kleemann, K., Moesswilde, E., Neupane, N., Newfield, T., Pei, Q., Xoplaki, E., Zappia, N. (2021). Towards a rigorous understanding of societal responses to climate change. Nature, 591(7851): 539-550. https://doi.org/10.1038/s41586-021-03190-2

[6] Allison, I., Paul, F., Colgan, W., King, M. (2021). Ice sheets, glaciers, and sea level. In Snow and Ice-Related Hazards, Risks, and Disasters. Elsevier. https://doi.org/10.1016/B978-0-12-817129-5.00013-5

[7] Manes, S., Vale, M.M. (2022). Achieving the Paris agreement would substantially reduce climate change risks to biodiversity in central and South America. Regional Environmental Change, 22(2): 60. https://doi.org/10.1007/s10113-022-01904-4

[8] Steffen, W. (2022). The nature of climate-related disasters in Australia. In Complex Disasters: Compounding, Cascading, and Protracted. Singapore: Springer Nature Singapore, pp. 57-79. https://doi.org/10.1007/978-981-19-2428-6_4

[9] Corbett, C.R., Parson, E.A. (2022). Radical climate adaptation in Antarctica. Ecology LQ, 49: 77. https://doi.org/10.15779/Z38BG2HB68

[10] Derot, J., Sugiura, N., Kim, S., Kouketsu, S. (2024). Improved climate time series forecasts by machine learning and statistical models coupled with signature

method: A case study with El Niño. Ecological Informatics, 79: 102437. https://doi.org/10.1016/j.ecoinf.2023.102437

[11] Choi, Y., Seo, S., Lee, J., Kim, T.W., Koo, C. (2024). A machine learning-based forecasting model for personal maximum allowable exposure time under extremely hot environments. Sustainable Cities and Society, 101: 105140. https://doi.org/10.1016/j.scs.2023.105140

[12] Song, J., Tong, G., Chao, J., Chung, J., Zhang, M., Lin, W., Zhang, T., Bentler, P.M., Zhu, W. (2023). Data driven pathway analysis and forecast of global warming and sea level rise. Scientific Reports, 13(1): 5536. https://doi.org/10.1038/s41598-023-30789-4

[13] Malakouti, S.M. (2023). Utilizing time series data from 1961 to 2019 recorded around the world and machine learning to create a Global Temperature Change Prediction Model. Case Studies in Chemical and Environmental Engineering, 7: 100312. https://doi.org/10.1016/j.cscee.2023.100312

[14] Qin, P., Shi, C. (2022). Characteristics of climate extremes in China during the recent global warming hiatus based upon machine learning. International Journal of Climatology, 42(4): 2099-2116. https://doi.org/10.1002/joc.7354

[15] Jose, D.M., Vincent, A.M., Dwarakish, G.S. (2022). Improving multiple model ensemble predictions of daily precipitation and temperature through machine learning techniques. Scientific Reports, 12(1): 4678. https://doi.org/10.1038/s41598-022-08786-w

[16] Akhshik, M., Bilton, A., Tjong, J., Singh, C.V., Faruk, O., Sain, M. (2022). Prediction of greenhouse gas emissions reductions via machine learning algorithms: Toward an artificial intelligence-based life cycle assessment for automotive lightweighting. Sustainable Materials and Technologies, 31: e00370. https://doi.org/10.1016/j.susmat.2021.e00370

[17] Anaraki, M.V., Farzin, S., Mousavi, S.F., Karami, H. (2021). Uncertainty analysis of climate change impacts on flood frequency by using hybrid machine learning methods. Water Resources Management, 35: 199-223. https://doi.org/10.1007/s11269-020-02719-w

[18] Berrang-Ford, L., Sietsma, A.J., Callaghan, M., Minx, J.C., Scheelbeek, P.F., Haddaway, N.R., FmedSci, A.H., Dangour, A.D. (2021). Systematic mapping of global research on climate and health: A machine learning review. The Lancet Planetary Health, 5(8): e514-e525. https://doi.org/10.1016/S2542-5196(21)00179-0

[19] Tsadiras, A., Pempetzoglou, M., Viktoratos, I. (2021). Making predictions of global warming impacts using a semantic web tool that simulates fuzzy cognitive maps. Computational Economics, 58(3): 715-745. https://doi.org/10.1007/s10614-020-10025-1

[20] Mardani, A., Liao, H., Nilashi, M., Alrasheedi, M., Cavallaro, F. (2020). A multi-stage method to predict carbon dioxide emissions using dimensionality reduction, clustering, and machine learning techniques. Journal of Cleaner Production, 275: 122942. https://doi.org/10.1016/j.jclepro.2020.122942

[21] Mansfield, L.A., Nowack, P.J., Kasoar, M., Everitt, R.G., Collins, W.J., Voulgarakis, A. (2020). Predicting global patterns of long-term climate change from short-term simulations using machine learning. Npj Climate and Atmospheric Science, 3(1): 44. https://doi.org/10.1038/s41612-020-00148-5

[22] Gholami Rostam, M., Sadatinejad, S.J., Malekian, A. (2020). Precipitation forecasting by large-scale climate indices and machine learning techniques. Journal of Arid Land, 12: 854-864. https://doi.org/10.1007/s40333-020-0097-3

[23] Ahmed, K., Sachindra, D.A., Shahid, S., Iqbal, Z., Nawaz, N., Khan, N. (2020). Multi-model ensemble predictions of precipitation and temperature using machine learning algorithms. Atmospheric Research, 236: 104806. https://doi.org/10.1016/j.atmosres.2019.104806

[24] Kalra, S., Lamba, R., Sharma, M. (2020). Machine learning based analysis for relation between global temperature and concentrations of greenhouse gases. Journal of Information and Optimization Sciences, 41(1): 73-84. https://doi.org/10.1080/02522667.2020.1715559

[25] Alfaseeh, L., Tu, R., Farooq, B., Hatzopoulou, M. (2020). Greenhouse gas emission prediction on road network using deep sequence learning. Transportation Research Part D: Transport and Environment, 88: 102593. https://doi.org/10.1016/j.trd.2020.102593

[26] SY, S. (2022) Temperature change, Kaggle. https://www.kaggle.com/sevgisarac/temperature-change, accessed on Apr. 17, 2024.

[27] Alqahtani, F. (2024). AI-driven improvement of monthly average rainfall forecasting in Mecca using grid search optimization for LSTM networks. Journal of Water and Climate Change, 15(4): 1439-1458. https://doi.org/10.2166/wcc.2024.242

[28] Xia, D., Zheng, Y., Bai, Y., Yan, X., Hu, Y., Li, Y., Li, H. (2022). A parallel grid-search-based SVM optimization algorithm on Spark for passenger hotspot prediction. Multimedia Tools and Applications, 81(19): 27523-27549. https://doi.org/10.1007/s11042-022-12077-x

[29] Di Nunno, F., Zhu, S., Ptak, M., Sojka, M., Granata, F. (2023). A stacked machine learning model for multi-step ahead prediction of lake surface water temperature. Science of the Total Environment, 890: 164323. https://doi.org/10.1016/j.scitotenv.2023.164323

[30] Mollick, T., Hashmi, G., Sabuj, S.R. (2023). A perceptible stacking ensemble model for air temperature prediction in a tropical climate zone. Discover Environment, 1(1): 15. https://doi.org/10.1007/s44274-023-00014-0