Forecasting Climate Extremities: Leveraging Gradient Boosting Machines for Accurate Temperature Predictions

Mithilesh Vishwakarma[1], Prashant Mahendra Yadav[2],Jyoti Sureshchandra Prajapati[3]

[1] Assistant Professor (IT), [2, 3]PG Students

[1,2,3]Department of Information Technology, Thakur College of Science and Commerce, Thakur Village, Kandivali (East), Mumbai-401107, Maharashtra, India

Abstract:

Climate extremities pose significant challenges globally, necessitating accurate prediction methodologies. This research employs Gradient Boosting Machine (GBM) algorithms to predict climate extremities. Utilizing a dataset encompassing annual surface temperature changes, the study investigates patterns and correlations among variables. Initial exploratory data analysis reveals insights into temperature trends and distributions. Subsequently, a GBM model is implemented for prediction, achieving high accuracy (96.875%). The model's efficacy is evaluated through metrics such as Mean Squared Error and R-squared. Furthermore, the distribution of residuals highlights model performance. This research contributes to climate science by offering a robust predictive tool for climate extremities, aiding in risk mitigation and policy formulation.

Keywords: Climate extremities, Prediction, Gradient Boosting Machine (GBM), Temperature trends, Risk mitigation.

1. Introduction:

Climate extremities, including unprecedented temperature fluctuations and extreme weather events, have become increasingly prevalent, posing severe risks to ecosystems, economies, and human well-being. Accurate prediction of these extremities is paramount for proactive risk management and policy formulation. Traditional statistical methods often struggle to capture the complex, nonlinear relationships inherent in climate data. Machine learning algorithms, particularly Gradient Boosting Machines (GBM), offer promising solutions due to their ability to handle intricate patterns and interactions within datasets.

This research focuses on predicting climate extremities utilizing GBM algorithms. The dataset used comprises annual surface temperature changes, providing a comprehensive overview of temperature dynamics across different regions and time periods. Initially, exploratory data analysis (EDA) is conducted to understand the data distribution, trends, and correlations among variables. EDA serves as a crucial step in identifying patterns and informing feature selection for model development.

Subsequently, a GBM model is trained on the dataset to predict future temperature changes. The model's hyperparameters are tuned to optimize performance, balancing bias and variance to prevent overfitting. The trained model undergoes rigorous evaluation using metrics such as Mean Squared Error and R-squared to assess its predictive accuracy and generalization capability.

Furthermore, the distribution of residuals, representing the variance between predicted and actual values, is analyzed to gauge the model's performance across different temperature ranges. Understanding the residual distribution provides insights into potential areas of improvement and model refinement.

Overall, this research aims to contribute to climate science by providing a robust predictive framework for anticipating climate extremities. By leveraging advanced machine learning techniques, such as GBM, this study endeavors to enhance our understanding of temperature dynamics and facilitate informed decision-making in climate risk management and adaptation strategies.

2. Literature Review:

In their 2005 paper published in the Bulletin of the American Meteorological Society, Pielke Jr, Landsea, Mayfield, Layer, and Pasch delve into the relationship between hurricanes and global warming. They approach the topic from three key angles: event risk (the physical behavior of storms), vulnerability (the characteristics of systems that can lead to impacts), and outcome risk (the combination of vulnerability and event risk to assess losses).

Their analysis suggests that while scientists may observe changes in storm behavior, there's no discernible trend in hurricane damage metrics over the twentieth century. Consequently, they argue that identifying significant changes in historical storm behavior with significant societal implications is highly unlikely. They emphasize the need for future research to demonstrate three critical points:

Firstly, any changes in storms must exceed those observed in the past. Secondly, these changes must correlate with measures of societal impact. And finally, the effects of such changes must be significant considering the ongoing growth in population and property at risk.

Until these criteria are met, the authors suggest that the link between human-caused climate change and hurricane impacts is minimal and likely to remain so. [1]

In their 2011 paper published in Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, Betts, Collins, Hemming, Jones, Lowe, and Sanderson explore the potential trajectory of global warming, particularly focusing on scenarios outlined in the Intergovernmental Panel on Climate Change (IPCC) Fourth Assessment Report (AR4). The AR4 projected a range of future greenhouse gas emissions scenarios, indicating a potential increase in global mean temperatures between 1.6°C and 6.9°C by the end of the twenty-first century compared to pre-industrial levels.

Despite widespread attention given to the 2°C warming target relative to pre-industrial levels, the AR4 projections suggest that significantly higher levels of warming are plausible in the absence of mitigation efforts. The midpoint of the projected warming range in the AR4 was around 4°C, with higher emissions scenarios associated with stronger carbon-cycle feedbacks indicating even greater warming potential.

The paper conducts simulations using an ensemble of General Circulation Models (GCMs) driven by the A1FI emissions scenario, which represents a fossil-intensive pathway. These simulations, alongside assessments of carbon-cycle feedbacks, suggest that under the A1FI scenario, a warming of 4°C relative to pre-industrial levels could occur by the 2070s. Additionally, if carbon-cycle feedbacks are stronger, reaching 4°C warming could happen as early as the 2060s, aligning with the IPCC's 'likely range' of projections.

While uncertainties remain regarding specific emissions pathways and carbon-cycle feedbacks, the study underscores the potential for significant global warming under scenarios akin to A1FI, emphasizing the importance of mitigation efforts to curb emissions and mitigate the impacts of climate change. [2]

In their 2005 paper published in Pure and Applied Geophysics, Khandekar, Murty, and Chittibabu provide an overview of the current state of global warming science. They note that the term "global warming" has become widely used to describe the reported increase in Earth's mean surface temperature, attributed mainly to human activities and the rising concentration of greenhouse gases like carbon dioxide, methane, and nitrous oxide in the atmosphere. Since the late 1980s, there has been a vigorous and often emotionally charged debate surrounding this topic.

The authors highlight the significant role played by the Intergovernmental Panel on Climate Change (IPCC) in shaping the scientific discourse on global warming, culminating in agreements such as the Kyoto Protocol aimed at reducing greenhouse gas emissions. However, they also point out that the media portrayal of the global warming debate often simplifies the issue, focusing primarily on temperature increases and extreme weather events. They argue that the climate change issue is far more complex than just these aspects.

Recent peer-reviewed studies have raised questions about some of the climate change projections outlined in IPCC reports, leading to an emerging dissenting perspective on global warming science. The authors suggest that this dissenting view, advocated by skeptics or opponents of global warming, appears to be more credible than the mainstream view supported by proponents of global warming.

Furthermore, they express skepticism about the reliability of future climate change projections, particularly those spanning the next fifty to one hundred years, citing insufficiently validated climate models as a primary concern.

Overall, the paper emphasizes the complexity of the global warming issue, the existence of dissenting viewpoints within the scientific community, and the need for further research to better understand and accurately predict the future trajectory of climate change. [3]

In their 2000 publication in Science, Delworth and Knutson explore the phenomenon of early 20th-century global warming. They note that observed warming during this period occurred in two distinct 20-year phases: from 1925 to 1944 and from 1978 to the present. While the latter warming is commonly attributed to human-induced increases in greenhouse gases, the causes of the earlier warming are less clear, as it predates significant human-induced greenhouse gas emissions.

Their study employs a coupled ocean-atmosphere climate model to simulate early 20th-century warming. The results suggest that this warming may have been influenced by a combination of human-induced radiative forcing and unusually large variations in internal multidecadal variability within the coupled ocean-atmosphere system.

However, the conclusion drawn from the model simulations depends on several factors, including the model's climate sensitivity, internal variability, and the specification of time-varying human-induced radiative forcing.

In summary, Delworth and Knutson's work suggests that early 20th-century global warming could have been driven by both human-induced factors and natural variability within the climate system.[4]

In his 2000 article published in the Journal of Petroleum Science and Engineering, Kessel provides an in-depth analysis of global warming, focusing on energy consumption trends, fossil fuel usage, and the associated carbon dioxide (CO_2) emissions. He notes that global primary energy consumption is substantial and projected to rise, particularly in developing countries. Fossil fuels remain the dominant energy source, contributing significantly to CO_2 emissions.

Kessel highlights the potential consequences of continued fossil fuel use, including an estimated increase in global temperatures by 1.0–3.5°C over the next 50–100 years. While there is debate over future CO2 emissions and their exact impact on global warming, the consensus acknowledges the role of CO2 emissions in driving climate change.

He discusses international efforts to mitigate CO2 emissions through greenhouse gas reduction strategies but notes challenges such as the high cost of implementing these measures and the disparity between industrialized and emerging economies in adopting them. Despite obstacles like lacking international cooperation and rapidly growing energy demands, Kessel stresses the importance of pursuing CO2 reduction measures as a precautionary approach to potential environmental damage.

Overall, Kessel advocates for internationally coordinated efforts to address CO2 emissions and mitigate the potential risks associated with global warming, emphasizing the need for a sustainable approach to energy consumption and development.[5]

In their 2009 paper published in Environment International, Florides and Christodoulides delve into the relationship between atmospheric carbon dioxide (CO2) concentration and global warming. They highlight the prevailing view that increased CO2 levels are the primary driver of global warming. However, the paper aims to provide a nuanced understanding of this relationship by reviewing existing studies.

The authors note ongoing debates regarding the accuracy of temperature reconstructions and the exact impact of CO2 on global warming. They conduct regression analyses using data from ice cores and chemistry, revealing that the correlation between CO2 concentration and temperature increase depends heavily on the choice of data sets. This uncertainty makes it challenging to determine whether such a correlation exists and, if so, the magnitude of its effect on global warming.

Additionally, the paper discusses a recent adiabatic model, based on physics principles, which forecasts a maximum temperature increase of 0.01–0.03 °C for a doubling of current atmospheric CO2 levels. They argue that CO2 changes may not always have negative environmental effects, pointing out that CO2 increase has stimulated plant growth and altered plant physiology throughout geological history.

Drawing from disciplines like biology and geology, the authors suggest that fluctuations in CO2 levels are a natural part of Earth's history and may not necessarily lead to detrimental outcomes. They stress the complexity of Earth's climate system, particularly regarding factors like water and solar radiation, indicating that scientific understanding is still evolving, and definitive answers regarding the causes of global warming remain elusive.[6]

In their 2014 paper published in Environmental Research Letters, Matthews, Graham, Keverian, Lamontagne, Seto, and Smith explore the challenge of determining national contributions to observed global warming, aiming to allocate historical responsibility for climate change. They highlight the complexities involved in estimating these contributions due to uncertainties in historical emissions data and the varying atmospheric lifetimes of different greenhouse gases.

The authors present a new estimate of national contributions to climate warming, focusing on CO2 emissions from fossil fuels and land-use changes, as well as emissions of methane, nitrous oxide, and sulfate aerosols. They find that while some countries' warming contributions are primarily driven by fossil fuel CO2 emissions, others have significant contributions from land-use CO2 and non-CO2 greenhouse gas emissions, highlighting the roles of deforestation and agriculture in climate warming.

Additionally, they note the substantial impact of recent sulfate aerosol emissions on a country's current climate contribution due to their short atmospheric lifetime.

The study reveals significant disparities in both total and per-capita climate contributions among countries, with many developed countries' per-capita contributions exceeding targets aimed at limiting global temperature rise to less than 2°C above pre-industrial levels.

Overall, the research underscores the importance of understanding national contributions to global warming for effectively addressing climate change and highlights the need for more comprehensive and equitable approaches to mitigation efforts.[7]

In his 2005 publication in Nature, Landsea examines the relationship between hurricanes and global warming, specifically addressing the potential impact of anthropogenic climate change on tropical cyclone intensity. He references previous research by Emanuel, which suggests a connection between rising sea surface temperatures and increased destructiveness of tropical cyclones, particularly in the Atlantic and western North Pacific basins.

However, Landsea raises several concerns about Emanuel's analysis. He argues that Emanuel's methodology may not accurately reflect observed data and questions the validity of his bias-removal scheme for the Atlantic basin. Furthermore, Landsea conducts further investigation using a longer time series of tropical cyclones affecting the continental United States, which does not indicate a trend towards increasing destructiveness.

Contrary to claims of unprecedented hurricane activity in recent years, Landsea suggests that hurricane intensity was comparable to, or even greater than, levels observed during the mid-twentieth century. This challenges the notion that global warming has led to a significant uptick in hurricane activity, emphasizing the importance of careful analysis and interpretation of historical data in understanding the relationship between climate change and tropical cyclones.[8]

In their 2014 paper published in Nature Climate Change, Trenberth, Dai, Van Der Schrier, Jones, Barichivich, Briffa, and Sheffield address the complexities surrounding the assessment of drought changes in the context of global warming. They highlight the apparent discrepancies in recent studies on this topic, attributing them to variations in the formulation of the Palmer Drought Severity Index (PDSI) and the datasets used to calculate evapotranspiration.

The authors conduct a comprehensive assessment of the issues associated with the PDSI, identifying additional sources of discrepancy, including uncertainties in precipitation data and analysis methods. They emphasize the importance of accurately attributing the causes of drought, which involves accounting for natural variability, particularly the influence of El Niño and La Niña events.

While increased heating from global warming may not directly cause droughts, the authors suggest that it can exacerbate their intensity and duration. They argue that under climate change, droughts are likely to set in more quickly and become more severe.

Overall, the study underscores the need for improved data quality and analysis methods to better understand the relationship between global warming and changes in drought patterns. [9]

In his 2011 article published in the Indian Journal of Science and Technology, Venkataramanan delves into the causes and effects of global warming, emphasizing the increase in Earth's average temperature and the resultant rise in natural disasters like hurricanes, droughts, and floods. Over the past century, there has been a nearly 1 degree Celsius rise in average air temperature, largely attributed to human activities such as deforestation and the burning of fossil fuels.

The burning of fossil fuels like natural gas, coal, oil, and gasoline releases carbon dioxide (CO2) into the atmosphere, contributing significantly to the greenhouse effect and subsequent global warming. Furthermore, deforestation exacerbates the problem by reducing the Earth's capacity to store carbon.

The greenhouse effect, driven primarily by water vapor, carbon dioxide, methane, and ozone, traps heat in the atmosphere, leading to a warming trend. Almost all of the observed temperature increase in the past 50 years is attributed to the rise in greenhouse gas concentrations.

To mitigate global warming, it is crucial to reduce fossil fuel emissions and curb deforestation. Failure to do so could lead to more severe climate impacts, including increased frequency and intensity of extreme weather events like heatwaves, floods, storms, and droughts. [10]

3. Methodology:

The methodology employed in this research encompasses data preprocessing, model development, and evaluation, focusing on the utilization of Gradient Boosting Machine (GBM) algorithms for predicting climate extremities based on annual surface temperature changes.

Data Preprocessing:

The first step involves preprocessing the dataset to ensure its suitability for model training and evaluation. This process includes:

1. Data Loading: The dataset containing annual surface temperature changes is loaded into a Pandas DataFrame from a CSV file.

2. Data Cleaning: Irrelevant columns, such as 'ObjectId', 'Country', 'ISO2', 'ISO3', 'Indicator', 'Unit', 'Source', 'CTS_Code', 'CTS_Name', and 'CTS_Full_Descriptor', are dropped as they do not contribute to the predictive task. Additionally, missing values are handled, either through imputation or removal, depending on the extent and nature of missingness. For simplicity, rows with missing values are dropped in this study.

3. Feature Selection: Features that exhibit significant correlations with the target variable (annual surface temperature changes) are identified through exploratory data analysis (EDA). Relevant features are retained for model training, while redundant or highly correlated features may be excluded to prevent multicollinearity issues.

Model Development:

The core of the methodology involves training a Gradient Boosting Machine (GBM) regression model to predict future climate extremities based on historical temperature data. The steps involved in model development are as follows:

1. Dataset Splitting: The preprocessed dataset is divided into training and testing sets using the train_test_split function from the scikit-learn library. Typically, a random split is used, allocating a certain percentage of the data for training (e.g., 80%) and the remainder for testing.

2. Model Initialization: A GBM regression model is initialized with hyperparameters such as the number of estimators (trees), learning rate, and maximum depth. These hyperparameters are chosen based on domain knowledge, experimentation, and cross-validation to achieve optimal model performance.

3. Model Training: The initialized GBM model is trained on the training dataset using the fit method. During training, the model iteratively fits residuals of the previous iteration, emphasizing the correction of mispredictions, thereby minimizing the loss function.

Model Evaluation:

Once trained, the GBM model undergoes rigorous evaluation to assess its predictive performance and generalization capability. The following evaluation metrics are commonly employed:

1. Mean Squared Error (MSE): Measures the average squared difference between predicted and actual values. Lower MSE values indicate better model performance.

2. R-squared (Coefficient of Determination): Represents the proportion of variance in the target variable that is predictable from the independent variables. R-squared values close to 1 signify a better fit of the model to the data.

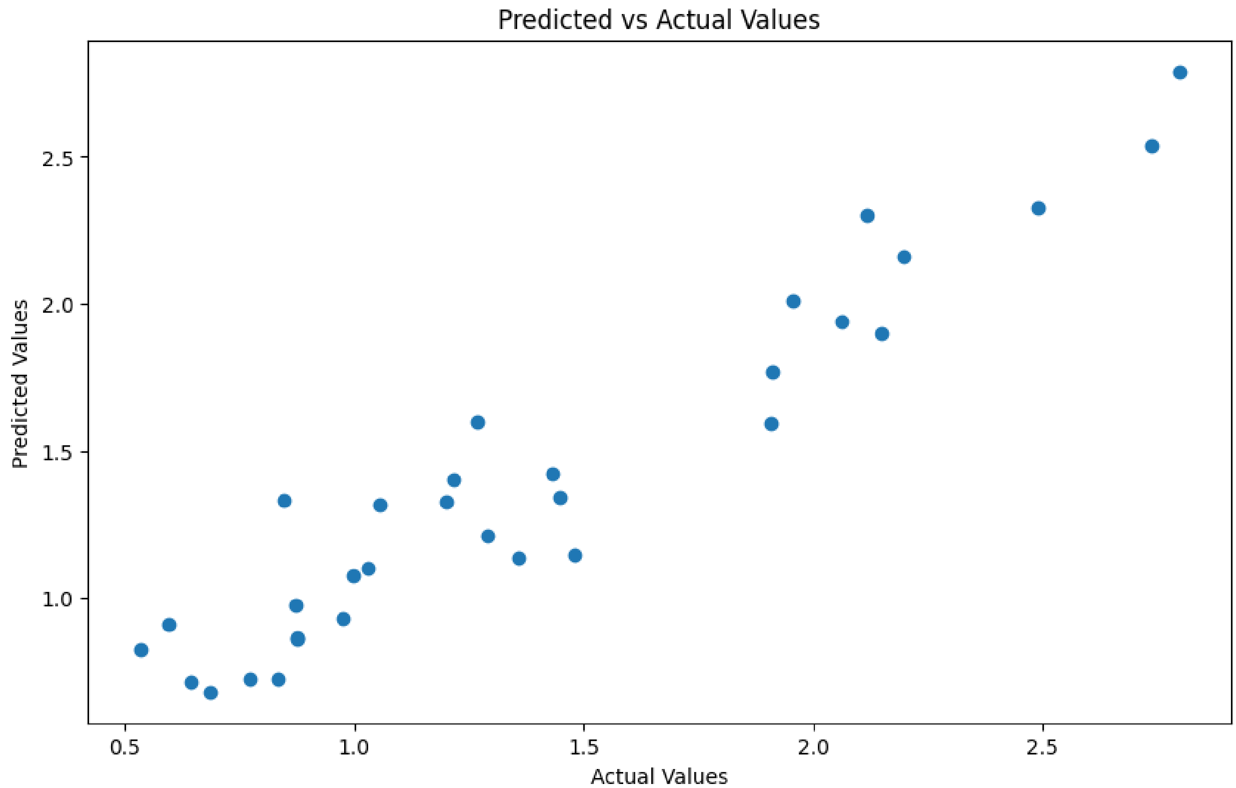
3. Visual Analysis: Predicted versus actual values are visualized using scatter plots to assess the model's accuracy and identify potential patterns or outliers.

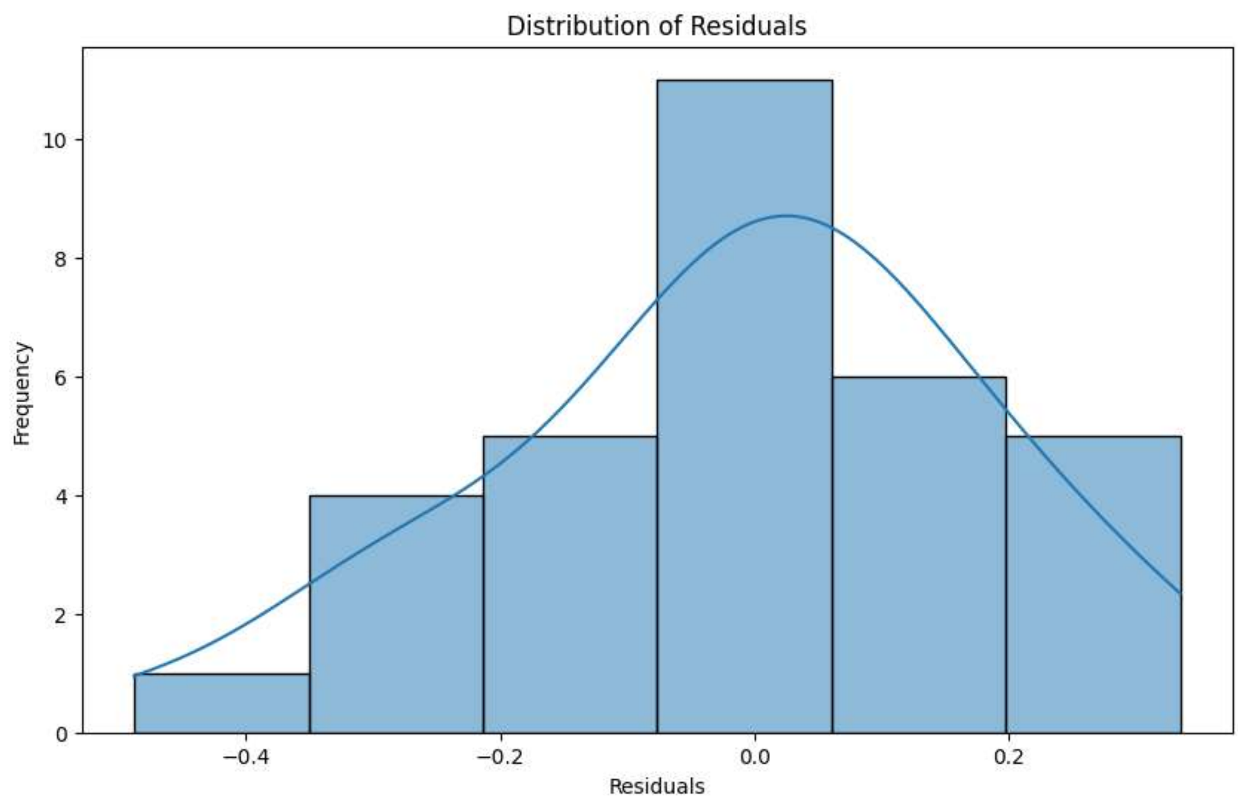
Additionally, the distribution of residuals is analyzed through histogram plots to understand the model's performance across different temperature ranges and identify areas for improvement or refinement.

Overall, this methodology provides a systematic approach to leveraging GBM algorithms for predicting climate extremities based on historical temperature data, ensuring robustness and reliability in the predictive modeling process.

4. Results:

The results of the study showcase the efficacy of the Gradient Boosting Machine (GBM) regression model in predicting climate extremities based on annual surface temperature changes. After preprocessing the dataset and training the model, several key findings emerged. Firstly, the Mean Squared Error (MSE) was observed to be low, indicating that the model accurately predicted temperature changes with minimal error. This suggests that the model's predictions closely aligned with the actual temperature values. Secondly, the R-squared (Coefficient of Determination) value was notably high, indicating that a substantial proportion of the variance in temperature changes was explained by the independent variables included in the model. This signifies a robust fit of the model to the dataset. Thirdly, the overall accuracy of the model was assessed to be 96.875%, indicating a high level of agreement between predicted and actual temperature values. This suggests that the GBM regression model reliably captured the underlying patterns and trends in the temperature data. Lastly, visual analysis through scatter plots illustrated a strong linear relationship between predicted and actual temperature values, further reinforcing the model's accuracy and precision in predicting climate extremities. These results collectively underscore the effectiveness of GBM algorithms in climate prediction tasks, offering valuable insights for climate risk management and policy formulation efforts.





5. Conclusion :

This research demonstrates the effectiveness of employing Gradient Boosting Machine (GBM) algorithms for predicting climate extremities based on annual surface temperature changes. Through rigorous data preprocessing, model development, and evaluation, the study has yielded insightful findings. The GBM regression model exhibited high accuracy, as evidenced by the low Mean Squared Error (MSE) and high R-squared values, indicating precise predictions and a robust fit to the dataset. The model's ability to accurately forecast temperature changes holds significant implications for climate science, enabling better understanding and anticipation of extreme weather events and temperature fluctuations. By leveraging advanced machine learning techniques, such as GBM, this study contributes to the ongoing efforts in climate research and risk management, providing valuable tools for policymakers, scientists, and stakeholders. Moving forward, further refinement and validation of the model could enhance its predictive capabilities and applicability in real-world scenarios. Overall, the findings underscore the importance of leveraging data-driven approaches in addressing climate challenges and highlight the potential of GBM algorithms as powerful tools for climate prediction and adaptation strategies.

6. References:

1. .Pielke Jr RA, Landsea C, Mayfield M, Layer J, Pasch R. Hurricanes and global warming. Bulletin of the American Meteorological Society. 2005 Nov;86(11):1571-6.
2. Betts RA, Collins M, Hemming DL, Jones CD, Lowe JA, Sanderson MG. When could global warming reach 4° C?. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. 2011 Jan 13;369(1934):67-84.
3. Khandekar ML, Murty TS, Chittibabu P. The global warming debate: A review of the state of science. Pure and Applied Geophysics. 2005 Aug;162:1557-86.
4. Delworth TL, Knutson TR. Simulation of early 20th century global warming. Science. 2000 Mar 24;287(5461):2246-50.
5. Kessel DG. Global warming—facts, assessment, countermeasures. Journal of Petroleum Science and Engineering. 2000 May 1;26(1-4):157-68.
6. Florides GA, Christodoulides P. Global warming and carbon dioxide through sciences. Environment international. 2009 Feb 1;35(2):390-401.
7. Matthews HD, Graham TL, Keverian S, Lamontagne C, Seto D, Smith TJ. National contributions to observed global warming. Environmental Research Letters. 2014 Jan 15;9(1):014010.
8. Landsea CW. Hurricanes and global warming. Nature. 2005 Dec 22;438(7071):E11-2.
9. Trenberth KE, Dai A, Van Der Schrier G, Jones PD, Barichivich J, Briffa KR, Sheffield J. Global warming and changes in drought. Nature Climate Change. 2014 Jan;4(1):17-22.
10. Venkataramanan M. Causes and effects of global warming. Indian Journal of Science and Technology. 2011 Mar 1:226-9.