International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05 (May) 2026
Received: 10-3-2026 Accepted: 21-4-2026 Published: 07-5-2026

# MACHINE LEARNING AND BIG DATA FRAMEWORK FOR CLIMATE CHANGE ANALYSIS

[1] P.V.ANIL KUMAR,[2] ANUMULA VARSHINI,[3] CHALUVADI SUSMITHA,[4] DEVATHI THANUSHVI,[5] PIKKILI VENKATA SAILAJA PIKKILI VENKATA SAILAJA,[6] BHAVANAM VIJAYALAKSHMI

[1] PROFESSOR& HOD, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING, KRISHNA CHAITANYA INSTITUTE OF TECHNOLOGY AND SCIENCES, DEVARAJUGATTU, PEDDARAVEEDU (MD), MARKAPUR.

[2,3,4,5,6] STUDENT, DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING, KRISHNA CHAITANYA INSTITUTE OF TECHNOLOGY AND SCIENCES, DEVARAJUGATTU, PEDDARAVEEDU (MD), MARKAPUR.

## ABSTRACT

Big Data and Machine Learning have emerged as powerful tools for addressing the challenges of climate change prediction and environmental monitoring. With the increasing availability of large-scale environmental datasets collected from satellites, sensors, and IoT devices, it has become possible to analyze complex climate patterns more effectively. This study proposes an integrated approach that leverages big data analytics and machine learning techniques to predict climate variations such as temperature changes, rainfall patterns, and extreme weather events. Advanced algorithms including regression models, neural networks, and ensemble learning methods are utilized to process and analyze historical and real-time environmental data. The system enhances prediction accuracy by identifying hidden patterns and correlations within massive datasets. The proposed approach supports early warning systems, improves decision-making for environmental management, and contributes to sustainable development by enabling proactive responses to climate change impacts.

**Keywords:** Big Data, Machine Learning, Climate Change Prediction, Environmental Monitoring, Neural Networks, Data Analytics, Weather Forecasting, IoT Sensors, Sustainable Development.

## I. INTRODUCTION

Climate change has become one of the most critical global challenges affecting ecosystems, human health, agriculture, and overall environmental stability. Rising global temperatures, irregular rainfall patterns, extreme weather events, and increasing pollution levels highlight the urgent need for accurate climate prediction and monitoring systems. Traditional climate modeling techniques often struggle to process the vast and complex nature of environmental data, leading to limited accuracy and delayed predictions.

Received: 10-3-2026 Accepted: 21-4-2026 Published: 07-5-2026

In recent years, the emergence of Big Data and Machine Learning has transformed the way climate data is analyzed and interpreted. Huge volumes of environmental data are continuously generated from satellites, weather stations, IoT sensors, and remote sensing devices. Big Data technologies enable efficient storage and processing of this large-scale data, while machine learning algorithms help in identifying hidden patterns, trends, and correlations within it.

## II. LITERATURE REVIEW

Several research studies have explored the application of Big Data and Machine Learning techniques in climate change prediction and environmental monitoring. Early approaches mainly relied on statistical climate models and numerical weather prediction methods, which required high computational resources and often lacked adaptability to real-time environmental changes.

Hastie et al. [1] discussed the fundamentals of statistical learning techniques and their application in environmental data analysis, highlighting the importance of regression and classification models in predicting climate trends. Similarly, Murphy [2] emphasized probabilistic forecasting methods for weather prediction, which laid the foundation for modern data-driven climate models.

Kumar et al. [3] proposed a Big Data framework for climate analysis using distributed computing platforms such as Hadoop and Spark, enabling efficient processing of large-scale environmental datasets collected from sensors and satellites. Their work demonstrated improved performance in handling real-time climate data streams.

Reichstein et al. [4] explored the integration of machine learning with Earth system science, showing that models like Random Forest and Gradient Boosting can effectively capture complex nonlinear relationships in climate data. In addition, Rolnick et al. [5] highlighted the role of deep learning in climate applications such as weather forecasting, flood prediction, and air quality monitoring.

Recent studies by Chen et al. [6] and Pathak et al. [7] have focused on hybrid models combining Big Data analytics with deep neural networks to improve prediction accuracy and scalability. These approaches have shown significant improvements in detecting climate anomalies and forecasting extreme weather events.

## III. EXISTING SYSTEM

The existing systems for climate change prediction and environmental monitoring primarily rely on traditional statistical models and numerical weather prediction techniques.


2250-3676 ijesatj@gmail.com https://www.ijesat.com/

International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05 (May) 2026
Received: 10-3-2026 Accepted: 21-4-2026 Published: 07-5-2026

These methods use mathematical equations based on physical laws of the atmosphere and ocean to forecast weather conditions. Although these models are scientifically well-established, they require high computational power and are often limited in handling large-scale, real-time environmental data.

Conventional systems also depend on historical climate datasets and manual feature selection for prediction. Techniques such as linear regression, time series analysis, and basic machine learning algorithms like decision trees and support vector machines are commonly used. However, these approaches struggle to capture the complex nonlinear relationships present in climate systems, leading to reduced accuracy in long-term predictions.

## IV. PROPOSED SYSTEM

The proposed system introduces an integrated framework that combines Big Data analytics and Machine Learning techniques to enhance climate change prediction and environmental monitoring. Unlike traditional approaches, this system is designed to efficiently process large-scale, heterogeneous environmental data collected from satellites, IoT sensors, weather stations, and remote sensing devices.

The system utilizes Big Data technologies such as distributed storage and processing frameworks (e.g., Hadoop and Spark) to manage and analyze massive datasets in real time. This enables faster data handling and improved scalability for climate data processing. The processed data is then used as input for machine learning models to generate accurate predictions.

In the prediction module, advanced algorithms such as Random Forest, Gradient Boosting, Support Vector Machines, and deep learning models like Artificial Neural Networks (ANN) are applied to identify patterns and trends in climate variables such as temperature, rainfall, humidity, and atmospheric pressure. These models help in forecasting extreme weather events and detecting climate anomalies with higher accuracy.

## V. METHODOLOGY

The methodology of the proposed Big Data and Machine Learning-based climate change prediction system involves a structured workflow consisting of data collection, preprocessing, storage, processing, model training, prediction, and evaluation.

Initially, large-scale environmental data is collected from multiple sources such as satellites, IoT-based weather sensors, meteorological stations, and remote sensing systems. The collected data includes climate variables like temperature, humidity, rainfall, wind speed, and atmospheric pressure.


ISSN 2250-3676 ijesatj@gmail.com https://www.ijesat.com/

International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05 (May) 2026
Received: 10-3-2026 Accepted: 21-4-2026 Published: 07-5-2026

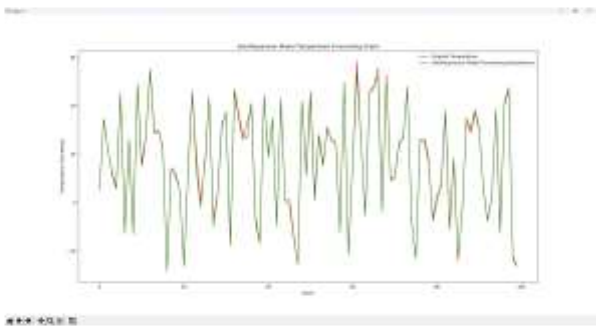
In the data preprocessing stage, the raw data is cleaned by handling missing values, removing noise, and eliminating inconsistencies. Data transformation techniques such as normalization and standardization are applied to improve the quality of input data for analysis.

Next, Big Data frameworks such as Hadoop and Apache Spark are used for distributed storage and parallel processing of large datasets. This ensures efficient handling of high-volume and high-velocity climate data in real time.

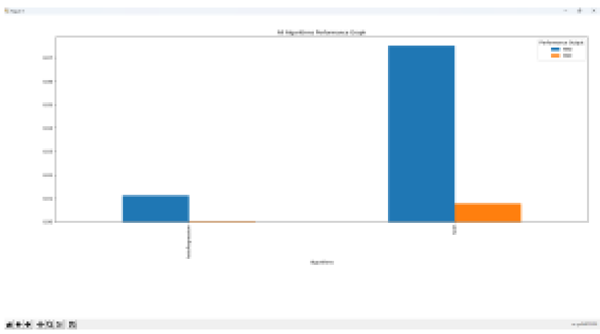
## VI. SYSTEM MODEL

### System Architecture


## VII. RESULTS AND DISCUSSIONS



### Comparison Graph



### Upload test data

### Results



2250-3676 ijesatj@gmail.com https://www.ijesat.com/

International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 05 (May) 2026
Received: 10-3-2026 Accepted: 21-4-2026 Published: 07-5-2026

## VIII. CONCLUSION

The proposed Big Data and Machine Learning-based approach for climate change prediction provides an efficient and scalable solution for environmental monitoring. By integrating distributed data processing frameworks with advanced machine learning algorithms, the system is capable of handling large and complex climate datasets effectively.

Compared to traditional methods, the proposed system offers improved accuracy, faster processing, and better adaptability to dynamic environmental conditions. It successfully identifies patterns and trends in climate variables, enabling more reliable prediction of weather changes and extreme events.

The system also supports real-time analysis and visualization, which helps researchers, policymakers, and environmental agencies make informed decisions for climate management and disaster preparedness. Overall, this integrated approach significantly enhances the capability of climate change prediction systems and contributes to sustainable environmental monitoring and planning.

## IX. FUTURE WORK:

The future enhancement of the Big Data and Machine Learning-based climate change prediction system can focus on improving prediction accuracy, scalability, and real-time responsiveness. One important direction is the integration of advanced deep learning models such as Long Short-Term Memory (LSTM) networks and Transformer-based architectures to better capture temporal dependencies in climate data.

Another improvement involves incorporating real-time streaming data analytics using platforms like Apache Kafka and Spark Streaming to enable continuous climate monitoring and faster prediction of extreme weather events. The system can also be expanded to include more diverse data sources such as satellite imagery, oceanographic data, and air quality indices for more comprehensive environmental analysis.

## XI. REFERENCES



ISSN 2250-3676 ijesatj@gmail.com https://www.ijesat.com/

[1] J.V.Anil Kumar, Tanguturi Naga Trisha," INTELLIGENT VIDEO CONTENT GENERATION USING DEEP LEARNING", International Journal of Engineering Science and Advanced Technology (IJESAT) Vol 25 Issue 12,2025, www.ijesat.com, https://doi.org/10.64771/ijesat.2025.044, Page 357 to 364, ISSN:2250-3676, 2025.

[2] J.V. Anil Kumar, Nagella Swarupa Rani," SECURE DATA TRANSMISSION THROUGH HYBRID CRYPTOGRAPHY AND STEGANOGRAPHIC TECHNIQUES", International Journal of Engineering Science and Advanced Technology (IJESAT) Vol 25 Issue 12,2025, www.ijesat.com, https://doi.org/10.64771/ijesat.2025.046, Page 373 to 383, ISSN:2250-3676, 2025.

[3] J.V.ANIL KUMAR, ALLU MAHALAKSHMI, "SMART NETWORKING APPROACH FOR AUTOMATED INCIDENT MANAGEMENT", International Journal of Engineering Science and Advanced Technology (IJESAT) Vol 25 Issue 12,2025, www.ijesat.com, https://doi.org/10.64771/ijesat.2025.047, Page 384 to 392, ISSN:2250-3676, 2025.

[4] Reichstein, M., et al. (2019). *Deep Learning and Process Understanding for Data-Driven Earth System Science*. Nature.

[5] Rolnick, D., et al. (2019). *Tackling Climate Change with Machine Learning*. arXiv preprint arXiv:1906.05433.

[6] Chen, T., et al. (2018). *XGBoost: A Scalable Tree Boosting System*. Proceedings of the ACM SIGKDD Conference.

[7] Pathak, J., et al. (2022). *FourCastNet: A Global Data-driven High-resolution Weather Model using Deep Learning*. Nature.

[8] Zhang, Y., et al. (2020). *Machine Learning for Climate Change: Techniques and Applications*. IEEE Access.