INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 05, Issue 03, March 2025, pp : 958-966
www.ijprems.com
editor@ijprems.com
e-ISSN : 2583-1062
Impact Factor : 7.001

AI & ML MODELS TO PREDICT CLIMATE CHANGE MITIGATION THROUGH HIGH-PRECISION ANALYTICS

Dr. Santosh Singh[1], Dr. Omkar Singh[2], Pratibha Jaiswar[3], Rutuja Shinde[4]

[1,2]H.O. D, Department of IT, Thakur College of Science and Commerce, Thakur Village, Kandivali (East), Mumbai, Maharashtra, India.

[3,4]PG student, department of IT, Thakur College of Science and Commerce,Thakur Village, Kandivali(East), Mumbai, Maharashtra, India.

sksingh@tcsc.edu.in, omkarsing@tcsc.edu.in, pratibhajaiswar1234@gmail.com, rutuja22022002@gmail.com

ABSTRACT

Intelligent techniques are only beginning to play a significant role in improving forecasts forSevere natural degradation and studying the changes in the climate through time. Advances inArtificial intelligence have significantly strengthened prediction techniques, such as newMachine learning and computational intelligence techniques. When it comes to the risksArtificial intelligence (AI) and machine learning present our rights, privacy, constitutional Protections, and democracy, many people are more focused on these perceived dangers.However, beyond such valid worries, AI has the potential to improve efficiency, provideMore accurate forecasts, and maximize the efficiency of the community as a whole by enhancing Various operations. Climate change threatens the functionality of societies, requiring aSignificant amounts of adaptability to keep up with changing climate changes in the future.Machine learning (ML) techniques have seen vast advances in the last several years. Prompting major discoveries in other fields of study, and researchers are now predictingMachine learning techniques may help climate studies. While a substantial number ofIndividual Planet System components have been studied using ML methods, it still hasn't Been used more broadly to comprehend the whole climate system. Incorporating the knownClimate links, artificial intelligence (AI) can construct improved weather alerts, particularly Severe occurrences.

Keywords: Temperature, Relative Humidity, Dew Point, Apparent Temperature, Precipitation, Rain, Mean Sea Level Pressure, Surface Pressure, Wind Speed, Wind Direction, Wind Gusts.

1. INTRODUCTION

There are many dangers in society today.That one of the most threatening is the Rising temperatures caused by climate climate Change. However, since the most Devastating effects are expected to happen Tomorrow, it is probable that more will present And less systemic issues will become Prominent in politics. Although we know That individuals act defensively. Downplaying the biggest risks in Community, experts recommend reminding Us of this. Denial has a miraculous effect On one's well-being, but it has no bearing On one's grandkids. The introduction to This section explains climate change in aNutshell. It's all about technology. And data difficulties, and it offers real-World climate change mitigation examples With the help of AI. Today's civilization is saturated with Artificial intelligence. Although the phrase "Data scientist" dates back to the 1950s. Is now a popular term. 31 But the vast majority of publications on artificial such Intelligence begins by admitting that this Word is vague, malleable, and susceptible to numerous interpretations. Starting with The big picture, therefore, is essential. At least explicitly define the three core Ideas we're going to be using: data However, to expand on that previous Point, you might say that when one speaks About artificial intelligence, no one is talking About "a collection of methods used to Simulate some aspects of social or Nonhuman intelligence uses computers [2]" One possible explanation for this Format's obscurity is because humans still I have yet to completely comprehend how Their own mind works. It may be simpler to embrace AI's gray Regions if we acknowledge our own limits. Because of our narrow experience, it is Likely that AI is often categorized using Basis for examining a wide assortment of Mental tasks that computers can perform Machine learning and artificial intelligence (AI) have revolutionized how we analyze and interpret vast amounts of data, enabling advanced functionalities such as voice recognition and face detection. These technologies are distinguished by their ability to resolve conflicts, fully comprehend natural language, and learn from experience. AI and machine learning can incur additional costs and complexity, but many argue that they are worth the investment. By integrating hypotheses, experiments, and deeper understanding, machine learning elevates the research process beyond mere data analysis. This approach relies on various methods that utilize large datasets to continuously enhance system capabilities through ongoing learning. Human input is critical to this process, providing essential specifications and context. Each time the software makes a prediction, it begins with an informed estimate about the type of knowledge it should pursue and then evaluates the accuracy of its prior assumptions by seeking detailed information. These processes can be seen as layered yet distinct, all of which require substantial data to be effective. In contrast, traditional data analytics often involves human analysts collecting large datasets to identify correlations among variables, searching for verifiable relationships based on pre-existing assumptions. Within the energy sector, business intelligence plays a pivotal role in predicting outcomes, optimizing operations, and facilitating transactions.

2. REVIEW OF LITERATURE

The study demonstrates that machine learning (ML) models can effectively predict climate variable changes and assess the relationship between greenhouse gas (GHG) emissions and climate variables, particularly in the North-East African region. By applying ML techniques such as long short-term memory (LSTM), autoencoders, and convolutional neural networks (CNN) to essential climate variables (ECVs) data, the study successfully identifies the best-performing model for predicting climate states. The CNN model outperformed others in terms of key performance metrics, including root-mean-squared-error, mean-absolute-error, Pearson correlation, and R^2 coefficient for temperature, CO_2, and CH_4 variables. This high accuracy in linking GHG emissions to ECVs can aid in climate adaptation and mitigation efforts, providing valuable insights for managing GHG concentrations and addressing climate events and crises in the region.Ibrahim SK, Ziedan IE, Ahmed A. Study of climate change detection in North-East Africa using machine learning and satellite data. IEEE Journal of Selected Topics in Applied Earth Observations andRemote Sensing. 2021 Oct 19;14:11080-94.[1]

The chapter highlights the critical role of artificial intelligence (AI) in addressing climate change, a major driver of large-scale human migration and a "threat multiplier & quot; according to the United Nations. The main causes of global warming include natural and anthropogenic activities like greenhouse gas emissions and deforestation, leading to increased temperatures, altered rainfall patterns, and extreme weather events. The global average temperature has risen by 0.6°C and is expected to increase further, with CO_2 concentrations projected to exceed 700 µmol mol–1 by the century. AI is emerging as a key tool in mitigating these effects, with the potential to reduce global greenhouse gas emissions by up to 4% and improve prediction tools for extreme environmental events. The integration of AI into climate action aligns with the United Nations Sustainable Development Goals (SDGs), particularly "Climate Action." By leveraging AI, it is possible to enhance climate prediction accuracy and support effective climate change mitigation strategies.Sahil K, Mehta P, Bhardwaj SK, Dhaliwal LK. Development of mitigation strategies for climate change using artificial intelligence to attain sustainability. InVisualization techniques for climate change with machine learning and artificial intelligence 2023 Jan 1 (pp. 421-448). Elsevier.[2]

The review highlights the transformative impact of artificial intelligence (AI) and machine learning (ML) on climate change mitigation, with a focus on remote sensing, urban transportation, and buildings. Over the past two decades, the exponential growth of relevant literature underscores the potential of big data and ML to move beyond generic climate recommendations. These technologies facilitate the development of tailored, context-specific solutions for urban planning, infrastructure, and policy, which can be scaled to address global climate challenges. The proposed meta-algorithmic framework offers a structured approach to leveraging ML for optimizing urban infrastructure, thereby enhancing the effectiveness of climate mitigation efforts and accelerating progress towards sustainable urban development.Milojevic-Dupont N, Creutzig F. Machine learning for geographically differentiated climate change mitigation in urban areas. Sustainable Cities and Society. 2021 Jan 1;64:102526.[3] .

The review highlights the challenges and opportunities in managing and interpreting the vast volumes of data generated by remote sensing and in situ instruments used to monitor climate change. It emphasizes that traditional methods are often inadequate for handling Big Datasets due to their size, complexity, and the need for rapid processing. By focusing on phenomena like droughts, floods, and sea-level rise, the review underscores the importance of developing advanced techniques for data analysis and modeling to enhance climate monitoring and sustainability. These innovations are crucial for timely geohazard warnings and effective climate response strategies.Montillet JP, Kermarrec G, Forootan E, Haberreiter M, He X, Finsterle W, Fernandes R, Shum CK. A review on how Big Data can help to monitor the environment and to mitigate risks due to climate change.[4]

The paper addresses the critical challenges of analyzing the massive volumes of climate change data generated by various sources, emphasizing that current algorithms struggle with the scale of this data. It underscores the importance of big data analytics in monitoring seasonal changes, assessing health risks, and optimizing natural resource management. By discussing various big data analytic methods, their strengths and weaknesses, and implementation frameworks, the paper provides insights into how these approaches can address climate change and sustainability challenges. It highlights the necessity of advancing these techniques to overcome data analysis issues and achieve sustainable development goals effectively.Ikegwu AC, Nweke HF, Mkpojiogu E, Anikwe CV, Igwe SA, Alo UR. Recently emerging trends in big data analytic methods for modeling and combating climate change effects. Energy Informatics. 2024 Feb 7;7(1):6.[5]

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 05, Issue 03, March 2025, pp : 958-966
www.ijprems.com
editor@ijprems.com
e-ISSN : 2583-1062
Impact Factor : 7.001

The paper explores the integration of artificial intelligence (AI) with Climate-Smart Agriculture (CSA) to enhance agricultural adaptation and productivity. It analyzes how AI can be applied across key areas such as crop and livestock management, ecosystem services, and farm and land management. By leveraging AI technologies, CSA aims to improve food security, adaptability to climate change, and sustainability, ultimately enhancing agricultural productivity and income in a cost-effective manner.Gryshova I, Balian A, Antonik I, Miniailo V, Nehodenko V, Nyzhnychenko Y. Artificial intelligence in climate smart in agriculture: toward a sustainable farming future. Access to science, business, innovation in the digital economy, ACCESS Press. 2024;5(1):125-40.[6]

3. CONCEPTUAL FRAMEWORK

3.1 Artificial Intelligence

AI seems to naturally be positioned to tackle the complex problems associated with climate change and environmental pollution. This section offers a quick introduction to AI and highlights various instances of its application in addressing issues of potential pollution that may be irreversible. Today, civilization is saturated with artificial intelligence, influencing numerous facets of our lives and industries. The term "artificial intelligence" was first popularized by software developers in the 1950s, and it has since become ubiquitous. However, many publications on AI begin by acknowledging the vagueness and malleability of the term, which can lead to numerous interpretations. While there is ongoing debate about how best to define AI, it can be understood as “a collection of methods for mimicking some aspects of social or nonhuman cognition using machines.” This ambiguity may stem from the fact that humans themselves have yet to fully comprehend the intricacies of their own minds. Acknowledging the limits of our understanding may help us embrace the uncertainties associated with AI. Currently, AI is often illustrated through examples that focus on specific, limited processes like speech recognition or computer vision, as well as broader cognitive tasks such as problem-solving, natural language processing (NLP), and learning. Thus, it is essential to clearly define three crucial constructs: advanced analytics, artificial intelligence, and machine learning. While these concepts are layered and interconnected, each relies heavily on large data sets to be effective. Traditional data analytics involves humans collecting substantial amounts of data to analyze it, seeking connections between variables based on assumptions. In the energy industry, effective use of business intelligence can improve prediction, control, and payment collection, handling many utility needs with less expense and complexity. However, when considering the added costs and time, machine learning can justify its higher price tag. In contemporary business intelligence, machine learning elevates the analysis process from simple data examination to one that involves assumption testing, experimentation, and self-learning. Quantum computing also plays a role, leveraging vast amounts of data to retrain algorithms and continuously enhance systems. Humans provide the foundational information and set crucial specifications, but as the algorithms process each piece of data, they formulate informed estimates about what knowledge to seek, continuously updating their understanding based on previous results.

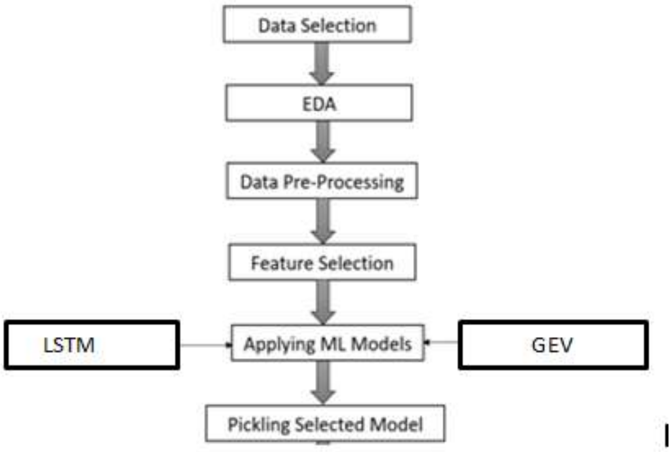
3.2 Machine Learning

Machine learning is revolutionizing how we understand and respond to climate extremities, offering advanced tools to process vast amounts of data from various sources and make precise predictions. Climate systems are highly complex, with numerous interacting variables such as temperature, humidity, wind patterns, ocean currents, and solar radiation. These interactions are difficult to model using traditional approaches, but machine learning algorithms excel at recognizing hidden patterns in large datasets. By processing environmental data from satellites, sensors, and weather stations, machine learning systems can analyze these interactions to predict extreme weather events such as hurricanes, floods, heatwaves, and droughts with greater accuracy. One of the significant advantages of machine learning is its ability to improve both short-term weather forecasts and long-term climate models. While traditional models often struggle with the unpredictable nature of climate systems, machine learning techniques like neural networks and decision trees can handle more variables and create models that better simulate the dynamics of our atmosphere. For example, by analyzing historical data on ocean temperature and atmospheric pressure, machine learning models can predict the intensity and path of hurricanes, allowing for earlier warnings and better preparation. These models are continuously learning, improving their accuracy as more data becomes available, enabling more reliable predictions of climate extremities.

3.3 Climate Change:

The management of energy consumption and the adjustment to its consequences are crucial in addressing environmental challenges that have emerged as global concerns. Throughout this project, a diverse array of research teams collaborated to uncover the fundamental principles underlying global climate change, particularly focusing on the significant role of greenhouse gas (GHG) emissions from living organisms and human activities. Their findings indicate that the increase in greenhouse gasses since the onset of industrialization has been a "primary mechanism" driving the remarkable rise in global atmospheric pressure. Extensive and compelling research indicates that the rise in atmospheric carbon dioxide is primarily due to fossil fuel combustion, which, if left unchecked, will further exacerbate global warming. Addressing these challenges requires not only immediate action to curb emissions but also a long-term commitment to sustainable practices and technologies that can effectively mitigate climate change impacts. Climate change refers to long-term shifts in temperatures and weather patterns, primarily caused by human activities like burning fossil fuels, deforestation, and industrialization. These activities release large quantities of greenhouse gasses (GHGs) such as carbon dioxide (CO_2), methane (CH_4), and nitrous oxide (N_2O) into the atmosphere. These gasses trap heat, leading to global warming and causing profound changes in the Earth's climate system. The effects of climate change on the environment are widespread and multifaceted, ranging from rising sea levels to biodiversity loss, with serious implications for ecosystems, weather patterns, and human societies.

4. METHODOLOGY



1. Problem Definition

Define the specific climate extremity event you want to predict, such as floods, heatwaves, droughts, hurricanes, or other extreme weather events. Clearly state whether this is a **regression** problem (e.g., predicting the intensity of an event) or a **classification** problem (e.g., determining whether an event will occur or not).

LSTMs are particularly well-suited for **time-series forecasting**, so they work best when predicting climate extremities based on historical patterns.

2. Data Collection

Collect data from reliable sources such as weather stations, satellites, and historical climate records. This data may include:

- **Temperature**
- **Humidity**
- **Precipitation**
- **Wind speed**
- **Atmospheric pressure**

Additionally, use **geospatial data** (e.g., topography, land use, population density) to improve predictions, especially for localized events. Since LSTMs rely on sequential data, **time-ordered data points** are essential to capture temporal dependencies.

3. Data Preprocessing

Since LSTMs require properly formatted time-series data, the preprocessing steps include:

Handling Missing Data

- Use **imputation techniques** (e.g., forward filling, mean imputation) or remove records with too many missing values.
- Ensure that all data points are consistently time-stamped and properly aligned.

Feature Engineering

- Create additional **time-based features** such as moving averages, rolling statistics, and time lags.
- Convert categorical variables (e.g., locations, seasons) into numerical representations using **one-hot encoding** or **label encoding**.

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 05, Issue 03, March 2025, pp : 958-966
www.ijprems.com
editor@ijprems.com
e-ISSN : 2583-1062
Impact Factor : 7.001

Scaling the Data

LSTMs are sensitive to the magnitude of input values, so normalize or scale features using:

- **Min-Max Scaling** (scales values between 0 and 1)
- **Standardization** (zero mean and unit variance)

4. Train-Test Split

- Split the dataset into **training** (e.g., 80%) and **test** (e.g., 20%) sets.
- **For time-series forecasting:** Ensure that the split maintains temporal order (i.e., train on past data, test on future data).
- Convert the dataset into sequences of time steps (e.g., if predicting extreme weather for tomorrow, use the past **n days** as input).

Example: If using **past 30 days** to predict extreme events, reshape the data into **(samples, time steps, features)** format.

5. Model Training with LSTM

Building an LSTM Model

1. **Import Required Libraries**
2. **Define LSTM Model**
3. **Train the Model**

6. Model Evaluation

- **Make Predictions**
- **Convert Predictions (if Classification)**
- **Evaluate Model Performance**

7. Feature Importance Analysis

8. Model Optimization

- **Hyperparameter Tuning**
- **Cross-Validation in Time Series**

9. Deploying the Model

10. Monitoring and Maintenance

LSTM Algorithm

Long Short-Term Memory (LSTM) networks are well-suited for predicting climate extremities due to their ability to capture long-term dependencies in time-series data. When applying LSTM to climate event prediction, the choice of an appropriate objective function depends on the problem type. For **binary classification**, where the goal is to determine whether an extreme event will occur or not, the **binary cross-entropy** loss function is used. In **multi-class classification**, where different types of extreme events (such as floods, droughts, and storms) are predicted, the **categorical cross-entropy** loss function is applied when using one-hot encoded labels, while **sparse categorical cross-entropy** is used for integer labels. For **regression-based predictions**, such as forecasting the intensity of an extreme event, loss functions like **mean squared error (MSE)** or **mean absolute error (MAE)** are used to quantify the difference between actual and predicted values.Since climate extremities are relatively rare compared to normal conditions, class imbalance is a significant challenge when training LSTM models. To address this issue, class weights can be assigned to give more importance to rare extreme events during training. Additionally, oversampling techniques such as **SMOTE (Synthetic Minority Over-sampling Technique)** or random oversampling can be used to increase the number of extreme event cases in the dataset. Alternatively, undersampling methods can reduce the number of normal condition instances to balance the dataset. Another approach is adjusting the decision threshold to improve sensitivity to extreme events, ensuring that the model does not overlook critical occurrences.

LSTM models require **sequential data** as input, making it crucial to incorporate both meteorological and temporal features. Key meteorological data includes **temperature, precipitation, humidity, wind speed, and atmospheric pressure**, which influence extreme weather events. Seasonal variations, such as the **year, month, day, and time of day**, help capture cyclic climate patterns. Geographical factors such as **latitude, longitude, elevation, land type, and proximity to water bodies** also contribute to climate event predictions. Additionally, historical data on past extreme events, along with derived features like **lag features (past values of meteorological variables), rolling statistics (moving averages, standard deviations), and anomaly detection (differences from historical norms)**, enhance the model's ability to recognize patterns. External factors, such as climate cycles like **El Niño and La Niña**, ocean temperature anomalies, and greenhouse gas levels, may also be incorporated to improve prediction accuracy.By leveraging **sequential time-series data**, LSTM networks effectively capture complex climate patterns, allowing for improved forecasting of extreme weather events. Proper handling of data imbalance, feature engineering, and model tuning can significantly enhance predictive performance.

GEV Model for Predicting Climate Extremities

The **Generalized Extreme Value (GEV) model** is a statistical approach used to analyze and predict the occurrence and intensity of extreme climate events, such as floods, heatwaves, storms, and droughts. Unlike machine learning models that capture sequential dependencies, GEV focuses on modeling the **probability distribution of extreme values** in climate data. The choice of the objective function in GEV depends on the nature of the data. The **maximum likelihood estimation (MLE)** method is typically used to estimate the three key parameters of the GEV distribution: **shape, location, and scale**, which define the behavior and magnitude of extreme events.Since extreme climate events are **rare**, handling class imbalance is an inherent part of GEV modeling. Instead of dealing with class distributions like classification models, GEV relies on **block maxima** or **peaks-over-threshold (POT)** approaches. In the **block maxima method**, the dataset is divided into time blocks (e.g., annual maximum temperature or rainfall), and only the most extreme value per block is retained. The **POT method** selects values above a predefined threshold, ensuring that the model captures meaningful extremes without losing important variations. Proper selection of block sizes and thresholds is crucial to avoid bias and ensure statistical validity.

Feature selection in GEV modeling is different from machine learning-based models since it focuses on capturing **extreme event distributions** rather than sequential dependencies. Essential meteorological features include **temperature, precipitation, humidity, wind speed, and atmospheric pressure**, as these factors contribute to extreme weather conditions. Seasonal trends such as **yearly and decadal cycles** help in understanding long-term variations in climate extremes. Geographical characteristics like **latitude, longitude, elevation, land type, and proximity to water bodies** influence the regional occurrence of extreme events. Additionally, external climate drivers, such as **El Niño and La Niña phenomena**, ocean temperature anomalies, and greenhouse gas levels, can impact the likelihood and intensity of extremes.The strength of GEV lies in its ability to estimate **return levels**, which predict the intensity of extreme climate events over specific return periods (e.g., the estimated rainfall level for a 50-year flood event). This makes GEV particularly useful for **risk assessment** and **infrastructure planning**, such as designing flood defenses and setting temperature thresholds for heatwave warnings. However, GEV assumes that climate extremes follow a stable statistical distribution, meaning that if climate patterns change significantly due to global warming or other factors, periodic recalibration of the model is necessary to maintain accuracy..

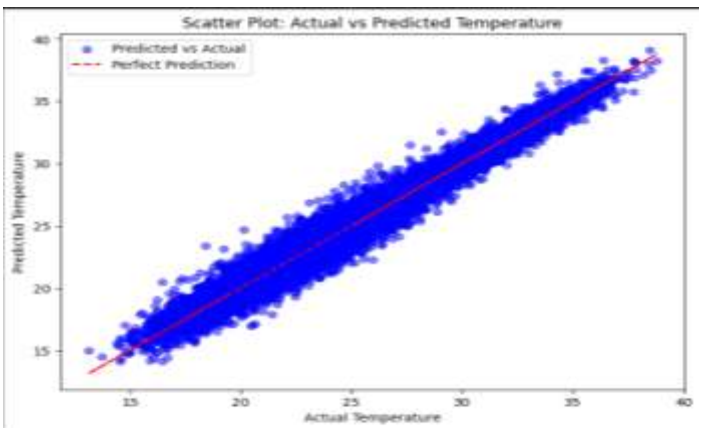
5. RESULT AND DISCUSSION

```
Training Set Performance:
R² Score: 0.999859784401012
MSE: 0.004697127068176566

Test Set Performance:
R² Score: 0.9998529263712498
MSE: 0.0049202625024369374
The model seems to be performing well.
```

The results suggest that the model is performing exceptionally well, as it has very high R^2 scores and low MSE values for both the training and test datasets. The close values of R^2 and MSE across both sets indicate that the model is generalizing well and not overfitting to the training data.

Scatter plot :



INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 05, Issue 03, March 2025, pp : 958-966
www.ijprems.com
editor@ijprems.com
e-ISSN : 2583-1062
Impact Factor : 7.001

The scatter plot illustrates the relationship between actual and predicted temperature values .From the plot, it is evident that there is a strong linear relationship between actual and predicted values, as most of the points are closely clustered along the red line. This indicates that the model has effectively learned temperature patterns and is making accurate predictions. Additionally, there is minimal deviation from the ideal prediction line, suggesting that errors are relatively small. While some points do deviate slightly, these variations are likely due to natural fluctuations in data or minor model imperfections.

Furthermore, there is no significant systematic bias, meaning the model does not consistently overpredict or underpredict values. If such a bias were present, the points would show a clear shift above or below the red line. Instead, the scatter plot confirms that the model's predictions are well-balanced and reliable.Overall, the visualization demonstrates that the LSTM-based temperature prediction model is performing well, with high accuracy and strong predictive capability. Minor deviations suggest areas for potential improvement, such as additional feature engineering or fine-tuning model parameters, but the model's current performance is robust and effective.

Model Performance Evaluation

LSTM model utilizes **Mean Squared Error (MSE)** as the loss function, which measures the average squared difference between actual and predicted values. Unlike accuracy metrics used in classification models, MSE provides an indication of how well the model predicts continuous numerical values.

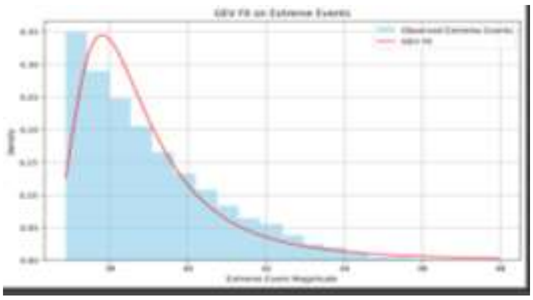
Final Performance Metrics:

- **Training Loss (MSE):** 2.8027e-04 (0.00028027)
- **Validation Loss (MSE):** 1.4913e-04 (0.00014913)

Key Insights:

1. **Lower MSE values indicate better predictive performance**, as they suggest that the predicted values are closely aligned with the actual values.
2. **Progressive improvement is evident** as the loss consistently decreases, signifying effective learning from the dataset.
3. **The validation loss is slightly lower than the training loss**, suggesting that the model generalizes well to unseen data and does not suffer from overfitting.
4. **The small magnitude of loss values** (in the range of 0.0001 - 0.0002) indicates a high degree of precision in temperature prediction.

Graph GEV Fit on Extreme Events



What the Graph Represents:

- The graph shows a **Generalized Extreme Value (GEV) distribution** fitted to observed extreme events.
- The **histogram (blue bars)** represents the empirical distribution of observed extreme events (e.g., extreme temperatures, precipitation, or any climate-related variable).
- The **red curve** is the **GEV fit**, which models the probability distribution of extreme events.
- The **GEV parameters** displayed (Shape, Location, and Scale) define the characteristics of this fitted distribution.

Key Observations:

1. **Right-Skewed Distribution:**
   - The histogram shows that extreme events mostly occur at lower magnitudes, with fewer occurrences at higher magnitudes.
   - This suggests that extreme climate events **follow a decreasing probability** as magnitude increases.
2. **GEV Fit Matches Observed Data Well:**
   - The red curve closely follows the histogram, indicating that the **GEV distribution is a good model** for the dataset.

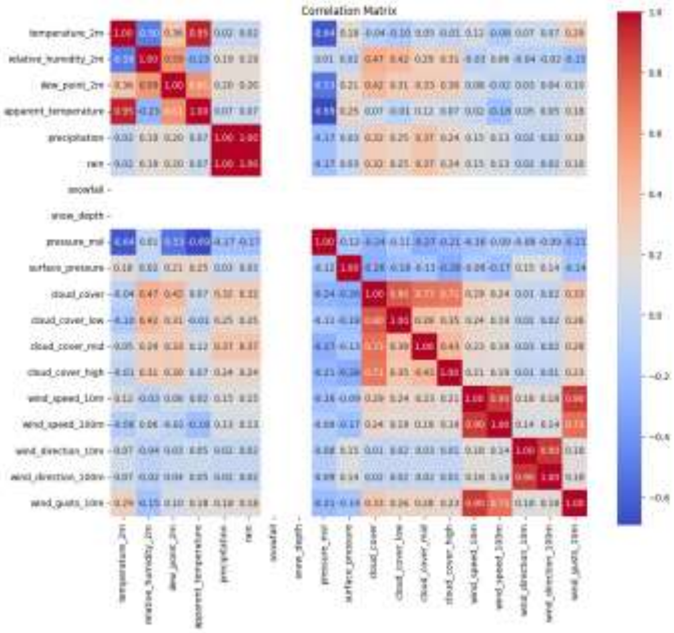
INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 05, Issue 03, March 2025, pp : 958-966
www.ijprems.com
editor@ijprems.com
e-ISSN : 2583-1062
Impact Factor : 7.001

3. GEV Parameters:

- **Shape (-0.2793):** A negative shape parameter suggests a **bounded distribution**, meaning there is an upper limit to the extreme event magnitudes.
- **Location (38.0605):** The central tendency of extreme event values is around **38**.
- **Scale (1.1056):** Controls the spread of the distribution. A higher scale value would indicate more variability in extreme events.

- The graph models **the likelihood of extreme climate events**, helping to predict their occurrence and severity.
- The GEV distribution is often used in **climate risk assessment**, infrastructure planning, and disaster preparedness.
- Since the shape parameter is negative, it implies that the most extreme values have a **finite upper limit**, meaning the dataset suggests there is a **cap on the severity of extreme events**.
- This analysis can be used for **forecasting future extreme events**, estimating return periods (e.g., "1-in-100-year events"), and making data-driven decisions in climate science.

CORRELATION MATRIX:



The correlation matrix provides valuable insights into the relationships between various meteorological variables, helping to identify patterns that can be useful for weather forecasting and climate modeling. Several strong positive correlations are observed, particularly among temperature-related variables. For instance, **temperature_2m** and **apparent_temperature** exhibit a high correlation of **0.95**, indicating that as actual temperature increases, the apparent temperature also rises. Similarly, **cloud cover levels at different altitudes (low, mid, and high) are strongly correlated with total cloud cover**, with values around **0.73 for cloud_cover_low and cloud_cover_mid, and 0.35 for cloud_cover_high**. This suggests that different cloud layers significantly contribute to the overall cloudiness. Additionally, **wind speed at different heights (10m and 100m) and wind gusts are strongly correlated (~0.73),** highlighting the relationship between wind speeds at various altitudes.On the other hand, strong negative correlations are also evident. **Temperature_2m and relative_humidity_2m** exhibit a negative correlation of **-0.50**, meaning that as temperature rises, relative humidity tends to decrease. Another significant negative correlation is found between **temperature_2m and pressure_msl (-0.64),** aligning with the general meteorological principle that higher temperatures are often associated with lower pressure systems. Additionally, **cloud cover and surface pressure show a negative correlation (-0.28),** indicating that increased cloudiness is often associated with low-pressure systems, which are typically linked to unsettled weather conditions.When examining precipitation, a strong correlation (~1.00) is observed between **precipitation and rain**, which is expected since rain is a direct contributor to total precipitation levels. Additionally, **precipitation shows a moderate correlation with cloud cover (~0.32),** suggesting that increased cloudiness often leads to higher chances of precipitation.Overall, this correlation matrix is a powerful tool for identifying key relationships between weather variables, helping to improve predictive models. Strong correlations indicate that some variables may provide redundant information, while weaker correlations highlight the need to consider multiple factors when modeling climate patterns. Understanding these relationships is crucial for feature selection in predictive modeling, ensuring that only the most relevant and non-redundant variables are used to enhance forecasting accuracy.

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 05, Issue 03, March 2025, pp : 958-966
www.ijprems.com
editor@ijprems.com
e-ISSN : 2583-1062
Impact Factor : 7.001

6. CONCLUSION

The LSTM-based temperature prediction model performs exceptionally well, with high R^2 scores, low MSE values, and minimal deviation in predictions. The scatter plot confirms strong predictive accuracy with no significant bias. The GEV distribution effectively models extreme climate events, suggesting a finite upper limit to severity. The correlation matrix reveals key meteorological relationships, such as strong positive correlations among temperature-related variables and negative correlations between temperature and humidity/pressure. These insights enhance forecasting accuracy and support data-driven climate analysis.

7. REFERENCES

[1] Ibrahim SK, Ziedan IE, Ahmed A. Study of climate change detection in North-East Africa using machine learning and satellite data. IEEE Journal of Selected Topics in Applied EarthObservations and Remote Sensing. 2021 Oct 19;14:11080-94.[1]

[2] Sahil K, Mehta P, Bhardwaj SK, Dhaliwal LK. Development of mitigation strategies for climate change using artificial intelligence to attain sustainability. InVisualization techniques for climate change with machine learning and artificial intelligence 2023 Jan 1 (pp. 421-448). Elsevier.[2]

[2] Milojevic-Dupont N, Creutzig F. Machine learning for geographically differentiated climate change mitigation in urban areas. Sustainable Cities and Society. 2021 Jan 1;64:102526.[3]

[3] Montillet JP, Kermarrec G, Forootan E, Haberreiter M, He X, Finsterle W, Fernandes R, Shum CK. A review on how Big Data can help to monitor the environment and to mitigate risks due to climate change.[4]

[4] Ikegwu AC, Nweke HF, Mkpojiogu E, Anikwe CV, Igwe SA, Alo UR. Recently emerging trends in big data analytic methods for modeling and combating climate change effects. EnergyInformatics. 2024 Feb 7;7(1):6.[5]

[5] Gryshova I, Balian A, Antonik I, Miniailo V, Nehodenko V, Nyzhnychenko Y. Artificial intelligence in climate smart in agriculture: toward a sustainable farming future. Access to science, business, innovation in the digital economy, ACCESS Press. 2024;5(1):125-40.[6]