RESEARCH ARTICLE OPEN ACCESS

INTELLIGENT CLIMATE PREDICTOR

Ponnaganti Sreenivas Rao[1], Nagulapally Vardhan[2], J. Likitha Varshini[3], P. Sujith Chary[4], Rudra Narayan Rout[5]

[1]Assistant Professor, Department of CSE (AI & ML), SIET, Hyderabad, India-500086
[2,3,4,5]UG Student, Department of CSE (AI & ML), SIET, Hyderabad, India-500086

Abstract

Climate change has emerged as one of the greatest problems in the world today, causing global warming, extreme weather conditions, rising sea levels, and environmental deterioration[2]. The fulfilment of the sustainable development goal 13 (climate action) would require accurate prediction, timely warning, and evidence-driven decision-making. This study recommends a climate-focused solution which would use machine learning algorithms and deep learning methods to analyse climate information and take pre-emptive actions[5].

The system will receive input from historical and current information on various parameters including temperature, precipitation levels, carbon dioxide emission, air pollution index, rate of deforestation, and weather. Machine learning algorithms like linear regression, random forest, and support vector machines would be used to identify patterns and forecast climate information[10]. Long short-term memory (LSTM) would be applied in deep learning to predict climate events including temperature increase, precipitation variability, and climate extremes.
The system will make predictions about future climate changes, detect anomalies, and assess environmental risks.
Keywords: Climate Change, Global Warming, Sustainable Development Goal 13 (Climate Action), Machine Learning, Deep Learning, Climate Prediction.

1. Introduction

Climate is a fundamental factor that supports and regulates life on earth. In recent years climate patterns have became unpredictable due to Global Warming , Solar Radiation , Volcanic Eruptions ,Ocean Currents , Deforestation ,Urbanization and Industrial Activities. Sudden changes in factors like rainfall, temperature and humidity makes the traditional adapt to changing patterns and continuously improve their performance.
We propose an Intelligent forecasting methods less reliable and lower accuracy.
Accurate climate prediction is very important to reduce the risks that are associated with the weather conditions like floods, droughts and heat waves[4]. Present weather forecasting methods mostly relay on statistical models and manual analysis which fail to capture complex data in climate data. As a result, there is need of intelligent systems that can analyze large volumes of historical data and provide accurate predictions.
Machine Learning(ML) provides various tools for identifying patterns in the large datasets[5]. By algorithms that are provided in the Machine learning such as regression models, decision trees, and time series forecasting techniques, Machine Learning can learn from past climate data and predict future conditions with high precision and accuracy. These models can
Climate Predictor using Machine Learning, which will forecast future climatic conditions based on the historical weather data. The system uses the Machine Learning features like data preprocessing techniques, feature selection, and predictive modelling to give the accurate climate predictions. The proposed model is service based model which mainly focuses to assist farmers and environmental agencies in making informed decisions and helps in disaster prevention by giving accurate results
The uniqueness of our project lies in integrating machine learning techniques with the climate data analysis to create an accurate, scalable and efficient prediction system. The proposed system mainly focuses on automation, adaptability and improved prediction accuracy compared to the traditional methods, which is

making it suitable for real world applications such as agriculture planning and disaster prevention.
This system is specially designed for region-based climate prediction with adaptability to localized datasets.

2. Proposed System

The proposed system, Intelligent Climate Predictor is a machine learning-based project that is designed to predict the future climatic conditions using historical weather data.
Traditional weather forecasting systems rely on numerical models and manual interpretation which makes it less precise. In case of proposed system it provides data-driven techniques to provide more scalable, adaptive and intelligent predictions.
The main aim of this system is to provide climate parameters such as temperature, humidity, rainfall and atmospheric conditions based on the previous patterns. By utilizing machine learning algorithms, the systems can base the historical datasets and generate accurate predictions for future.
This proposed system focuses particularly on Indian climatic conditions, ensuring accurate predictions in the particular regions.

Objectives of the Proposed System

The main objectives of the proposed system are:

- To develop a project that predicts the future climate conditions.
- To give accurate and real time climate predictions
- To analyze the historical weather data efficiently and accurately
- To design a user friendly interface for easy interactions
- To create a service based project that can help in the agricultural departments
- To predict the future natural disasters

System Overview

The Intelligent Climate predictor system consists of:

1. Data Collection Model
2. Data Preprocessing Module
3. Machine Learning Model
4. Prediction Engine
5. User Interface(UI)

These all components work together to form a complete pipeline, transforming raw data into meaningful predictions.

System Architecture

The proposed system follows a structured architecture:

Data Collection

Data collection is the foundational step in the Intelligent Climate Predictor system, where the climate data is collected from multiple reliable sources. The system collects real-time
information such as temperature, humidity, rainfall and also speed of the wind using online weather services like OpenWeatherMap and WeatherAPI. The historical climate records are also obtained from the organizations like Indian Meteorological Department and also from the Kaggle datasets. The combination of real-time and past data helps the system understand patterns behaviour. The collected data is usually stored in structured formats.

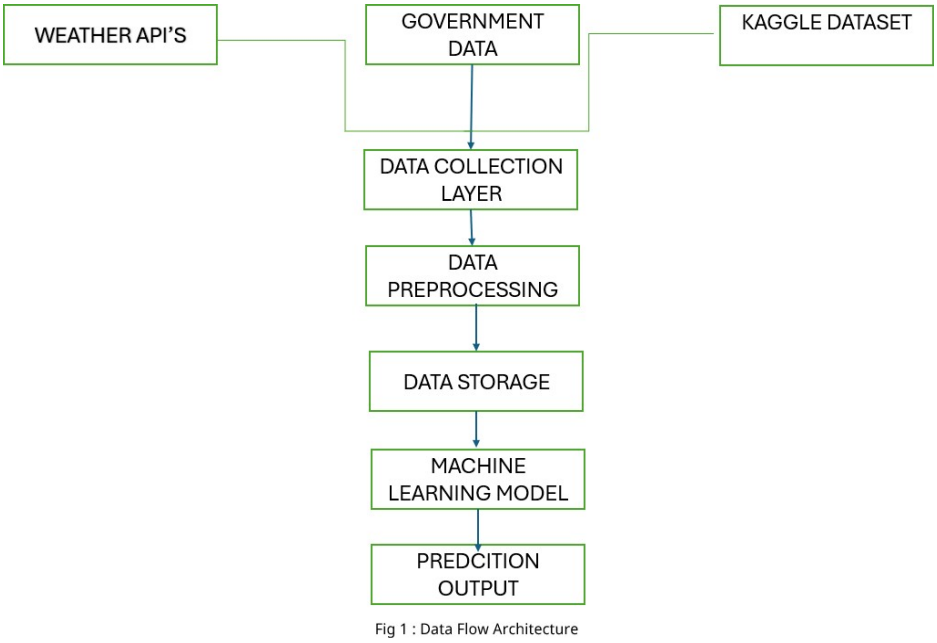

fig 1. Data Flow Architecture

Data Preprocessing

Data processing is an important step in Intelligent Climate Predictor system, it transforms the raw and unstructured historical climate data into a clean usable format for Machine Learning models[7][13]. This would include handling of missing data, removal of duplicate entries and inconsistent data, and converting data to correct formats. Standardization is used to achieve uniformity within the data range. There are certain other attributes which would also have to be considered for achieving redundancy elimination and increased predictive accuracy .Feature engineering techniques are also used to extract useful information like seasonal patterns from the data attributes. The dataset that is collected is further divided into training and the testing sets to enable effective model training and evaluation. This stage improves the reliability of the system and gives the accurate and precise climate predictions.

Machine learning Model

The ML Model can be referred to as a cornerstone of the Intelligent Climate Predictor System, because it assists in translating past weather observation into the forecast of the future weather condition. The development of the ML Model involves the employment of supervised machine learning algorithms, which enable the use of weather data history featuring inputs and outputs. As a result, patterns linked with certain weather occurrences can be detected with the help of this algorithm. The examples of algorithms that were used for creating the ML model can be linear regression, decision tree, and random forest algorithms. The random forest regressor algorithm seems to be the best algorithm among the others, because it is created on the basis of the use of many decision trees, which makes it more accurate. When considering input data features, one needs to pay attention to past temperature, humidity, precipitation, wind speed, month, and year. Data pre-processing is needed prior to starting training of the algorithm. Following this, the splitting process of the dataset in the ratio of 80% for the training sample and 20% for the testing sample is performed. While the training sample serves as the basis of developing the predictive model, its performance is estimated on the basis of the testing sample. It is noteworthy that during the learning process, the model learns how to detect the complicated interconnection between the variables and adjust the parameters to minimize prediction errors. The performance of the model is assessed in terms of the three metrics: MAE (the metric that estimates the average error), MSE (the metric punishing high errors), and R^2 score (the measure of how well the model fits the dataset). Once the learning phase is finished, the model is integrated into the prediction engine of the

system and works by getting inputs from users (date or weather conditions) providing outputs regarding temperature, precipitation and humidity

Prediction

The prediction is the final stage of the process, where the machine learning algorithm predicts the future value of the climate based on the inputs provided by the user. After successful training of the machine learning algorithm, the first step is to predict the climate parameter values, including temperature, rainfall, and humidity, according to any criterion. For example, the criterion may be in days or in terms of anything else. In the first place, the inputs of the parameters will be provided to the machine by the user in days or in terms of any other environmental factors. The inputs are then processed by the machine, and the output comes out on the basis of the trained pattern.

Visualization

The visualization step is dedicated to visualizing climate data so that users understand what trends and changes have occurred in the climate data over time through visual representation. In other words, instead of providing users with numerical values regarding temperature, rainfall, and humidity, they will be provided with graphical representation including charts and lines and other ways to visualize the data. Consequently, there is no need for users to have expertise in the area of climatology to make sense of the prediction results. Visualization can be done by using the libraries Matplotlib and Seaborn. There are several types of visualizations, including line charts, bar charts, and others. Line charts are going to be used to show how temperatures have changed over time. Bar charts will illustrate how rainfall varied from one month to another.

Working of Proposed System

The functioning of the Intelligent Climate Predictor system is based on a sequence of steps that take the raw historical data through a series of processes to yield useful climate predictions. To begin with, the system gathers weather information, including temperature, humidity, rainfall, and wind speed, from reliable sources. Next, the gathered information undergoes a pre-processing stage, which entails addressing missing values and noise within the information, after which it is arranged in a form appropriate for further processing. After the pre-processing stage, key attributes such as the date (year, month, day) are derived to enable the system to distinguish seasonal trends. Prediction follows training as the next stage. Once the system receives the required training, it is able to predict some outcomes right away. As soon as users provide input regarding any specific dates or environmental conditions, the input is processed automatically by the application and then provided to the trained model. The prediction about different aspects such as temperature, rainfall, and humidity will be done using this input from the model and then provided to the users in the form of graphical representation. Hence, the entire process can be considered as an automated series comprising data collection, processing, machine learning, and visualization.

Technologies Used

The Intelligent Climate Predictor software has been designed using different programming languages, libraries and tools that make it easy to do data manipulations and machine learning visualizations. The programming language that has been used to develop the software is Python because it is easy to learn but very efficient in data science and machine learning applications. In order to be able to manipulate and pre-process data, different libraries have been used including Pandas and NumPy which help to manage big data sets and conduct all kinds of calculations and data clean-ups. When it comes to data visualization, programs such as Matplotlib and Seaborn can be used for producing graphs which will highlight the trends and forecasts of climate data. In order to build a user interface for the project, Streamlit can be used since it is an open-source platform which helps create a web application quickly. In addition to this, the sources from where data will be collected can be platforms such as Kaggle or other meteorological databases.

3. Experimental Setup

Experimental process in the design of Intelligent Climate Predictor describes the approach in which the machine learning algorithm is designed, trained, and evaluated from the real dataset. The initial process for the design of the experiment is getting details related to previous climatic conditions by authentic sources like Kaggle or any other data source that includes parameters such as temperature, humidity, rainfall, and wind speed. The collected dataset is normalized along with the handling of missing values in it. The cleaned data set is divided into the training data set and the testing data set in a ratio of 80:20, respectively, whereby the training data set is employed in the training process, while the testing data set is used in the evaluation process. The random forest is selected to be trained because it has the highest accuracy compared to other algorithms. The random forest uses the training data set in its learning process by understanding the relationship between the input and the output variables. After the training of the model is completed, then it will be tested based on the test data in order to analyze its efficiency. Some of the measures that may be used to do this include mean absolute error (MAE), mean squared error (MSE), and the R2 Score. This experiment may be performed using the Python programming language in environments such as Jupyter notebook and Visual Studio code utilizing different libraries in Python including Pandas, NumPy, Scikit-learn, Matplotlib, and Seaborn.

4. Results

Analysis of Historical Climate Data

The historical dataset provides the foundational understanding of climatic conditions in India over the past decade. Based on the graphical and tabular data provided below which are collected from Kaggle Dataset , clear patterns of rainfall, summer temperatures, and winter temperatures.

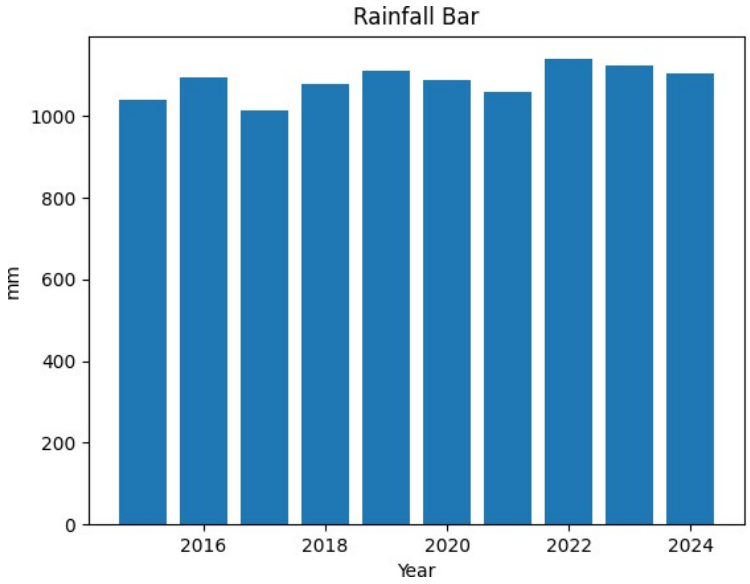

fig 2: Rainfall Bar

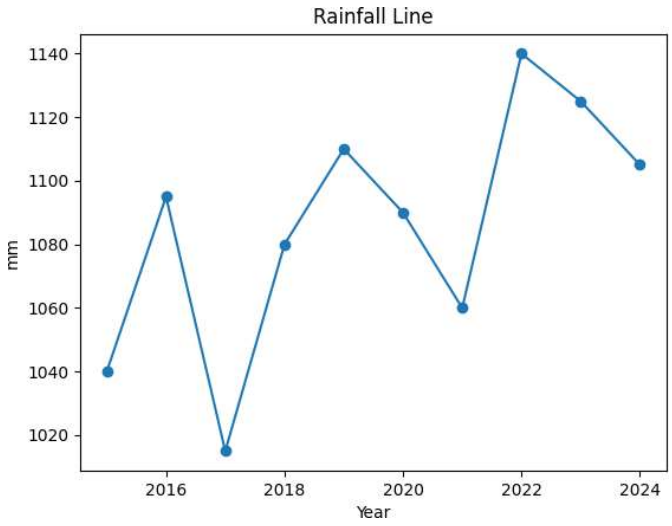

fig 3: Rainfall Line

However , in the process of analyzing the rainfall pattern in India for the period from 2015 to 2024, it can be seen that a dynamic and rising pattern can be detected. In analyzing the data set collected during the period, the rainfall pattern ranges from 1015 mm to 1140 mm, hence not indicating any consistent rise but variations in the rainfall in various years. For instance, in 2015, the rainfall was recorded at 1040 mm, but it rose in 2016 to 1095 mm and then dropped drastically in 2017 to 1015 mm due to weak monsoon seasons. However, thereafter, there were increases for the next few years to 1140 mm recorded in 2022. This is an exceptional case in terms of rainfall. It means that even though there is more uncertainty in terms of how rainwater distributes itself, there is an increase in the intensity of the precipitation. Additionally, based on the predicted quantity of precipitation for 2025 to 2034, there is no variation in terms of the amount of rainfall, which has gradually risen from 1110mm to 1200mm. It implies that the future rainfall pattern will remain consistent and high, resulting in increased chances of flooding and other rainfall-related problems.

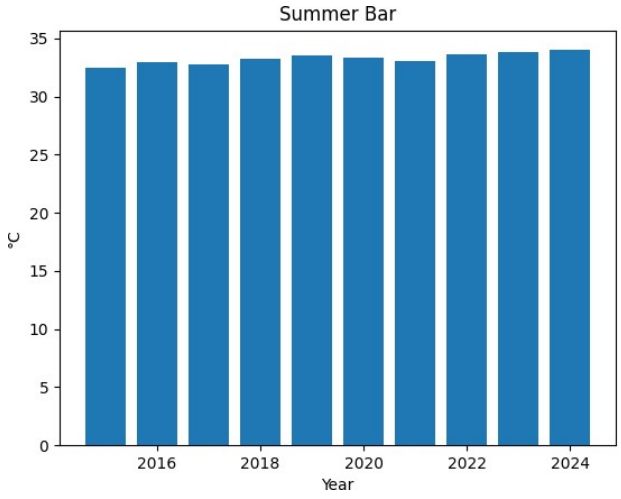

fig 4 : Summer Bar

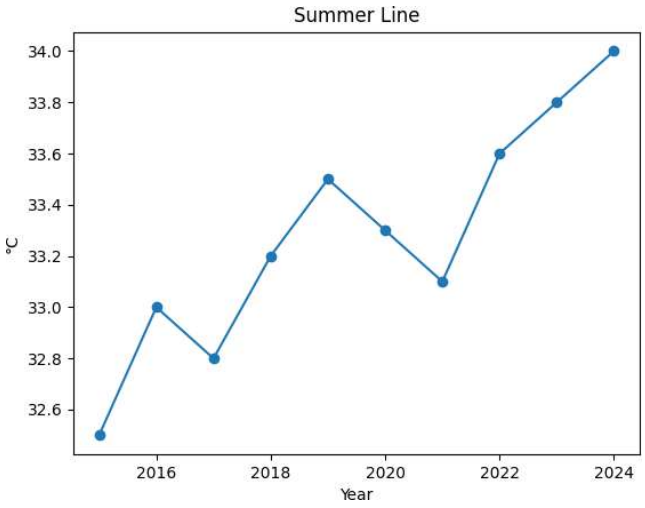

fig 5 : Summer line

Trend analysis of the average summer temperatures in the period from 2015 to 2024 demonstrates the presence of the positive trend with certain seasonal deviations. Thus, beginning from the value of 32.5 degrees centigrade in 2015, the trend attains its highest value of 33.0 degrees centigrade in 2016, following which it starts showing a temporary fall reaching
32.8 degrees centigrade in 2017. From that point and up until 2021, there is a rather stable growth in the level of temperatures, raising them to 33.2 degrees centigrade in 2018 and continuing the process of growth to reach 33.5 degrees centigrade in 2019. There is another drop in the temperature during the next two years reaching the levels of 33.3 and 33.1 degrees centigrade correspondingly, which suggests some temporary weather anomalies. Nevertheless, since 2022, the temperature level started to grow sharply again, reaching values of 33.6, 33.8, and eventually 34.0 degrees centigrade in 2024. Overall, it is a total increase by 1.5 degrees in ten years, which demonstrates the overall increasing temperature in summer.

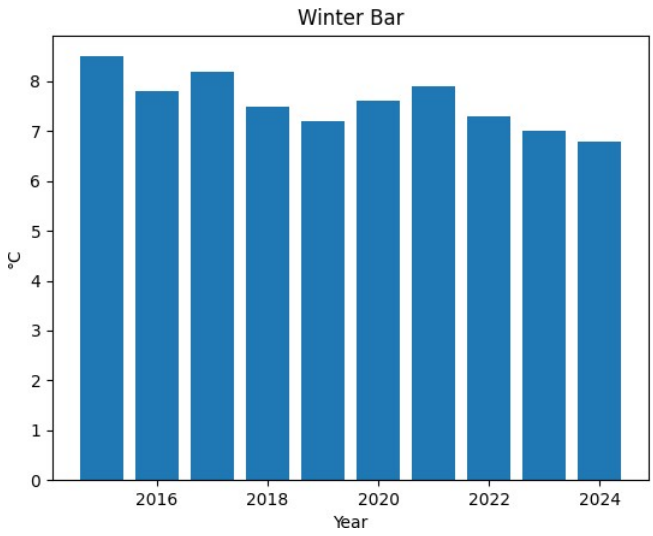

fig 6 : Winter Bar

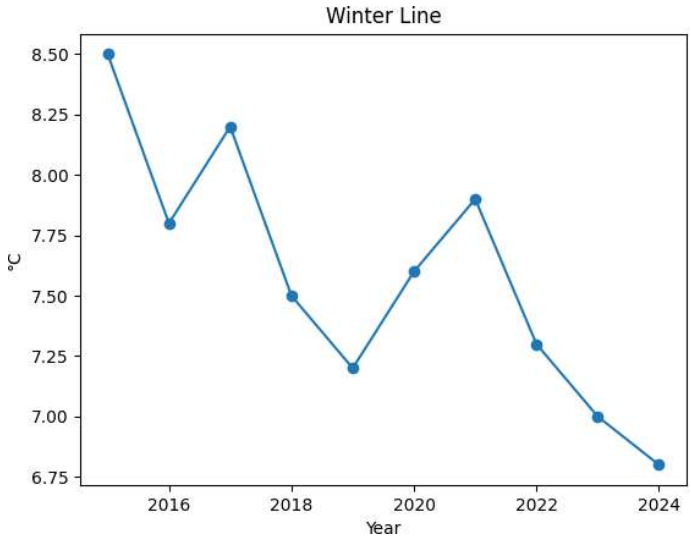

fig 7 : Winter Line

From the graph on the temperature trends during winter months from 2015 through 2024, we observe that there is a general downward trend with some ups and downs. The temperature starts at 8.5°C in 2015 before plummeting to 7.8°C in 2016 and rising slightly to 8.2°C in 2017. Thereafter, the temperature gradually falls to reach 7.5°C in 2018 and 7.2°C in 2019. However, there is a sudden rise in the temperature from 2020 to 2021 to be 7.6°C and 7.9°C, respectively, showing the irregularities in climate trends. Nonetheless, since 2022, the trend of temperatures keeps falling downwards until the end of the period. As can be deduced from the statistics provided, the drop is estimated at about 1.7°C over the span of ten years. Factoring in the possible fluctuations in the temperature levels, the overall trend is evident – winter temperatures are constantly falling. Thus, there is a noticeable difference between the seasons, with winters being increasingly colder and summers hotter. The graphical representation provided clearly proves this statement.

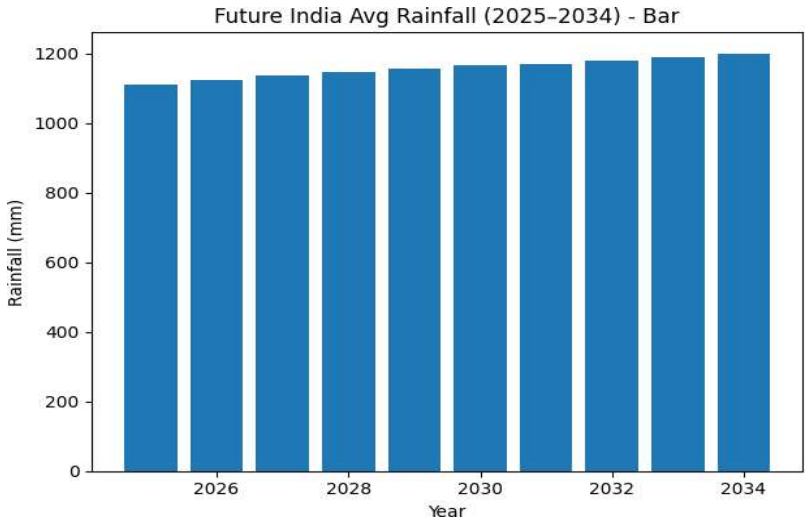

fig 8 : Predicted India Rainfall Trend

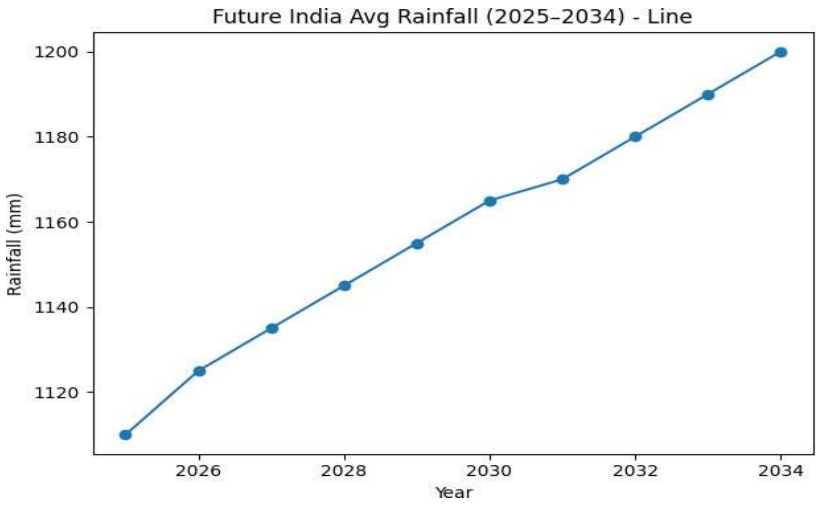

fig 9 : Predicted Indian Rainfall Trend(Line graph)

The projected rainfall for India between 2025 and 2034 illustrates a positive trend with no interruptions at all, as illustrated on the bar graph below. The initial rainfall value in 2025 is roughly 1110 mm, and increases every year to 1200 mm by 2034. Unlike the historical data that included ups and downs, the projected rainfall indicates continuity with no interruptions. The rate of growth per year in terms of rainfall volume is fairly even, and averages at about 8-10 mm per year, illustrating the existence of a long-term trend in the model instead of short-term fluctuations. The reason for an increased rate of rainfall lies in climate changes, leading to greater moisture levels due to higher temperatures. The last important point is the absence of sharp falls or spikes within the data forecasts. This means that the machine learning model was able to pick up on the trend in past years' data while ignoring any anomalies, making the resulting data more linear compared to rainfalls in previous years. From an environmental perspective, increased rainfall will definitely result in serious consequences. Although large amounts of rain can be beneficial for agriculture and water supply in some regions, too much rainfall may lead to floods, erosion, and waterlogging. Moreover, constant rain may influence monsoon seasons, consequently disrupting agriculture and water management.

Overall, the graph discussed above depicts the increase in rainfall in India for the next decade. It also shows that the rain pattern in the coming years will change because of the rise in rainfall, and it might have some benefits, along with some difficulties as well.

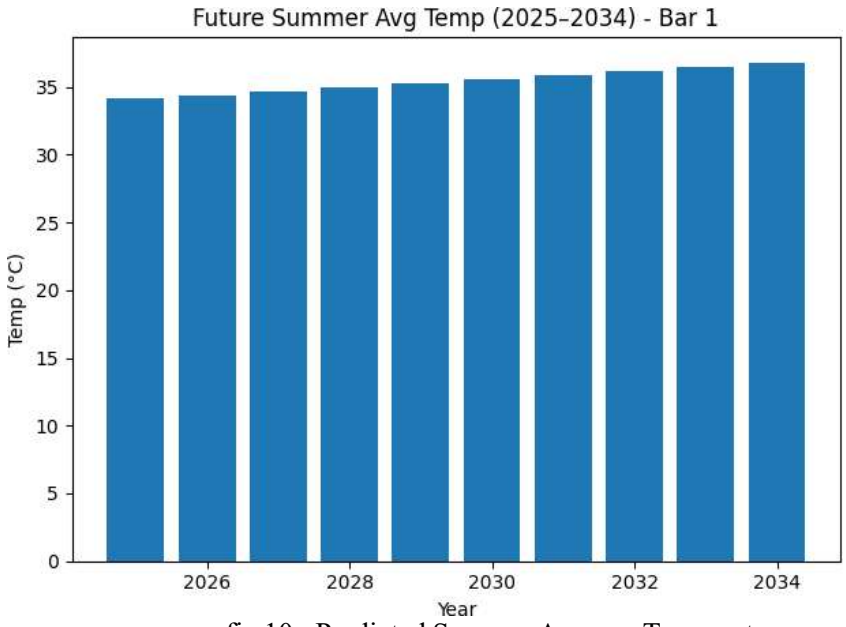

fig 10 : Predicted Summer Average Temperature

From the years 2025 to 2034, it can be noted from the bar graph provided above that there is a clear increase in the average temperature in India during the summer period. From this, it can be seen that the average temperature in 2025 will be approximately 34.2 degrees Celsius, and from here on, it increases to 36.8 degrees Celsius within a period of ten years, resulting in an increase of about 2.6 degrees in the temperature. There is no fluctuation in the temperature level but an increase, which indicates that the model applied has been able to establish an increase in temperature levels. The temperature is rising at the rate of 0.2–0.3°C per year, meaning that in the coming years, the warming will increase because of factors such as greenhouse gases.

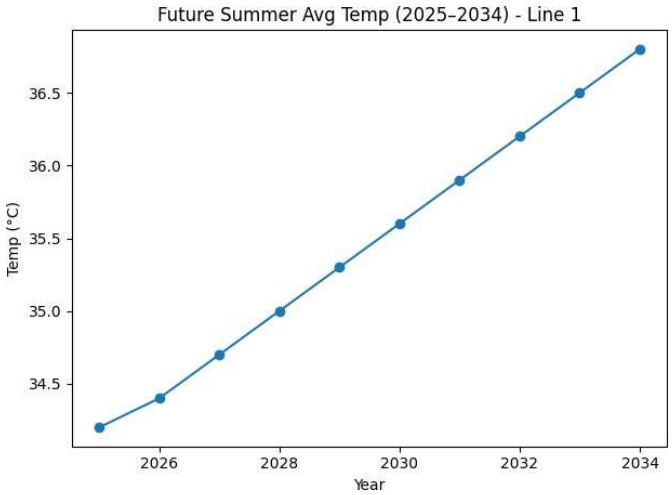

fig 11 : Predicted Summer Average Temperature(line)

It can also be observed that the predicted values always lie above the value of 34°C for the duration of ten years. In other words, the threshold temperature has changed as a result of

which the past temperature becomes the new temperature now. With respect to the environmental issues, higher summer temperatures might have several implications. One of such implications is a higher number of occurrences of heat waves, which means higher demand for water resources. Another implication is the loss of moisture from the soil, affecting agriculture in India. Finally, another consequence of higher temperatures can be heat strokes and other related diseases among humans.In conclusion, the graph shows the tendency towards rising temperatures in the coming ten years. As it can be seen from the above analysis, the summer temperatures in India will keep increasing for ten years at least.

It is evident that the outcomes of this study indicate the changes in climate properties in India. According to the historical data analysis for the period between 2015 and 2024, it could be stated that rainfall showed minor variability with slight growth, whereas the temperatures during summer grew, and those during winter decreased. Thus, the increase in climate variability and extreme climatic seasons was observed.Furthermore, as for the results that were predicted using machine learning algorithms, it is possible to note the predicted increase in climate variability, including the rise in rainfall, rapid growth in summer temperatures, and further reduction in winter temperatures for the years from 2025 to 2034.As one can see, the climate change process is evident, and it might become even more intense than now. Moreover, the extreme weather and rainfall changes will influence many spheres of life in India; therefore, people should adapt and take measures to reduce their impacts on the society and environment. It is important to note that the outcomes of this research prove that machine learning tools are quite effective for climate predictions.

5. Applications

The results obtained from this research are highly enlightening regarding the alterations going on in our climatic environment and can have various applications in the outside world. As we all know, because the impacts of climatic change have become very apparent with each passing day, it would be imperative to apply these results in practical life situations. As far as the uses of these results are concerned, it can be said that one of the most valuable uses of these results would be realized in the domain of agriculture. The farmer depends upon his knowledge about the trends in seasons in order to sow seeds and then harvest the crop. Since there is going to be a rise in temperature and rainfall, the farmer will be able to take the necessary decisions. Another crucial use lies in disaster management.

The increasing number of rains in the forthcoming years indicates that the chances of floods will be increased. Disaster management companies can use these predictions to come up with effective early warning systems as well as plans for prevention of these natural disasters. The prediction that there will be more warm summers implies that heat waves will occur in the forthcoming years. There are numerous other applications of this research in various fields. There must be adequate provisions for the storage of water because of the predicted increase in the rainfalls. This research helps a great deal in urban planning which involves construction of water reservoirs because of the increasing rainfall. Another area that has gained significantly from this research is water resource management. With the increase in rainfall, it is very important that there must be plans for utilizing water.

In connection with technologies, one should state that this project serves as another proof of the effectiveness of machine learning technologies application for climate studies. Indeed, it becomes obvious that the algorithm operates correctly allowing making predictions regarding climate changes. This aspect can be used when dealing with more data and creating more complicated models.As far as the interpretation of results is concerned, several important conclusions can be made. First of all, it is important to say that historical data is quite inconsistent due to climatic natural fluctuations. However, predicted data seems to follow a regular tendency line as a machine learning algorithm always concentrates on long-run tendencies. While such approach provides for making more reliable predictions, the problem is that some unexpected changes may not be tracked in time.

To conclude, one should emphasize that the findings prove the continuing trend of climate change. As it can be seen from the figures above, temperature extremes and rainfall levels increase.

6. Conclusion

The current research has been able to accurately predict and analyze climatic behavior patterns in India

through the use of machine learning algorithms based on past data. The findings from the research have demonstrated that there is a shift in climatic patterns towards increased temperatures during summers, reduced temperatures during winters, and increased precipitation levels. From the predictions provided, it can be said that this trend is expected to continue in the coming years. The continued increase in both temperatures and precipitation levels implies that there will be an escalation in environmental problems like heat waves and floods. Moreover, through this project, it is evident that machine learning can be used effectively to analyse information relating to the environment. This is essential for decision-making in sectors of agriculture, disaster, and urban development. Conclusion Through this project, one gets an excellent insight into current climate changes as well as an insight into the future to prepare humanity for the future. It is through collective effort that society will be able to cope with all these changes.

References

[1] India Meteorological Department. (n.d.). *Climate data services*. Government of India. Retrieved from https://www.imd.gov.in

[2] Intergovernmental Panel on Climate Change. (2021). *Climate change 2021: The physical science basis*. Cambridge University Press. https://doi.org/10.1017/9781009157896

[3] National Aeronautics and Space Administration. (n.d.). *Global climate change: Vital signs of the planet*. Retrieved from https://climate.nasa.gov

[4] World Meteorological Organization. (2023). *State of the global climate 2023*. Retrieved from https://public.wmo.int

[5] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12*, 2825–2830.

[6] Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering, 9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

[7] McKinney, W. (2010). Data structures for statistical computing in Python. In S. van der Walt & J. Millman (Eds.), *Proceedings of the 9th Python in Science Conference* (pp. 51–56).

[8] Python Software Foundation. (n.d.). *Python documentation*. Retrieved from https://www.python.org

[9] Allen, M. R., Dube, O. P., Solecki, W., Aragón-Durand, F., Cramer, W., Humphreys, S., Kainuma, M., Kala, J., Mahowald, N., Mulugetta, Y., Perez, R., Wairiu, M., & Zickfeld, K. (2018). Framing and context. In *Global warming of 1.5°C*. Intergovernmental Panel on Climate Change.

[10] Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

[11] Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics, 29*(5), 1189–1232.

[12] Hansen, J., Sato, M., Ruedy, R., Lo, K., Lea, D. W., & Medina-Elizade, M. (2006). Global temperature change. *Proceedings of the National Academy of Sciences, 103*(39), 14288–14293.

[13] Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer.

[14] Intergovernmental Panel on Climate Change. (2021). *Climate change 2021: The physical science basis*. Cambridge University Press.

[15] Kumar, K. R., Sahai, A. K., Kumar, K. K., Patwardhan, S. K., Mishra, P. K., Revadekar, J. V., Kamala, K., & Pant, G. B. (2006). High-resolution climate change scenarios for India. *Current Science, 90*(3), 334–345.

[16] McKinney, W. (2010). Data structures for statistical computing in Python. In *Proceedings of the 9th*

Python in Science Conference (pp. 51–56).
[17] Pedregosa, F., et al. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12*, 2825–2830.
[18] Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning, 1*(1), 81–106.
[19] Rasmussen, C. E., & Williams, C. K. I. (2006). *Gaussian processes for machine learning*. MIT Press.
[20] Shukla, J., Nobre, C., & Sellers, P. (1990). Amazon deforestation and climate change. *Science, 247*(4948), 1322–1325.
[21] Stocker, T. F., Qin, D., Plattner, G. K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., & Midgley, P. M. (2013). *Climate change 2013: The physical science basis*. Cambridge University Press.
[22] World Meteorological Organization. (2023). *State of the global climate 2023*.

