IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Machine Learning Weather Forecasting System

[1]Gawandar Ankita, [2]More Aishwarya, [3]Dawkhar Gauri,[4]Solomon Merrin

[1]Student, [2]Student, [3]Student,[4]Professor

[1]Data Science,

[1]Usha Mittal Institute of Technology, Mumbai, India

Abstract: Weather forecasting is a most important application in meteorology and has been one of the most scientifically and technologically challenging problems around the world. Weather forecasting has traditionally been done by physical models of the atmosphere, which are unstable to perturbations, and thus are inaccurate for large periods of time. This paper introduces a new approach to weather forecasting by combining advanced machine learning algorithms with a user-friendly web-based application. The study focuses on improving the accuracy and reliability of weather predictions while providing an intuitive interface for users to access real-time meteorological information. The integration of machine learning models, such as Decision Tree and ensemble methods, into the forecasting system. The web application harnesses the power of these algorithms to deliver personalized and location-specific weather forecasts. We present the methodology for data collection, model training, and application development, highlighting the seamless integration of machine learning insights. Initial results demonstrate the superior predictive capabilities of the integrated system compared to traditional forecasting methods. The study concludes by discussing the potential impact of this approach on the future of weather forecasting.

Index Terms **- Decision Tree, Weather Forecasting, Algorithm.**

I. INTRODUCTION

Weather forecasting is complex and dynamic, relying on various parameters like temperature, wind speed, and humidity. These parameters vary with specific locations and atmospheric attributes. It involves gathering quantitative information about the environment's current condition through climate stations and deciphering by meteorologists. Climate and weather affect human society in various ways, including crop production, water resources and energy sources. Data mining techniques include Association Rules Mining, Cluster analysis, and Classification/Prediction. The weather forecasting industry in India, despite decades of experience, faces numerous challenges, including arbitrary and unsuitable nature-based expectations. This research work focuses on solving the weather prediction anomalies and in-efficiency based on decision tree. In the present study the application of different tools of decision tree has been studied for weather parameters. Decision Tree is an optimization technique that uses tree structure graphs or models to make decisions by analysing possible events and outcomes. First of all we create a model than model is trained by a sample of data called training set. A trained model is provided with unseen data called a test set in predicting the future event classification with an accuracy.

II. PROBLEM STATEMENT

Weather forecasting is not only vital for numerous sectors but also impacts the daily lives of individuals worldwide. Inaccurate or untimely weather predictions can lead to significant consequences, including agricultural losses, disrupted transportation systems, increased risk of natural disasters, and compromised public safety. Despite the considerable progress made in meteorological science and technology, certain obstacles persist in delivering precise and timely weather forecasts. This project seeks to confront these challenges head-on by leveraging advancements in meteorological research, data analytics, and computational techniques. By integrating cutting-edge technologies such as machine learning, artificial intelligence, and high-performance computing, the aim is to enhance the accuracy, reliability, and timeliness of weather predictions. Through a multi-faceted approach encompassing improved data collection, refined modelling

algorithms, and enhanced forecasting methodologies, this initiative strives to revolutionize the field of weather prediction. This project represents a concerted effort to overcome the challenges associated with inaccurate and untimely weather forecasts. By harnessing the power of technology and scientific innovation, it seeks to usher in a new era of more reliable, accessible, and impactful weather prediction, thereby safeguarding lives, livelihoods, and critical infrastructure in an increasingly unpredictable world.

III. EXISTING SYSTEM

The existing system for weather forecasting relies on a sophisticated combination of observational data, advanced technologies, and numerical weather prediction models to provide accurate and timely forecasts. This multifaceted approach is designed to capture the complexities of the Earth's atmosphere and anticipate a wide range of weather conditions. The existing weather forecasting infrastructure offers a robust suite of services crucial for public safety and disaster preparedness, providing a wide range of forecasts and alerts for various weather parameters. Leveraging a combination of advanced technologies and data sources, including Numerical Weather Prediction models, meteorological satellites, ground-based observation networks, and supercomputing facilities, meteorological agencies worldwide deliver accurate and timely weather forecasts. These forecasts are disseminated through public communication platforms, such as websites and mobile apps, ensuring widespread accessibility and awareness. Additionally, climate prediction centers focus on longer-term climate forecasts, while data assimilation systems integrate real-time observations with model outputs for enhanced accuracy. This comprehensive system, spearheaded by agencies like NOAA, plays a vital role in safeguarding lives and property by providing reliable weather information to governments, industries, and the general public. While the existing weather forecasting system is robust, there are certain drawbacks that need to be addressed. One significant challenge is the limited resolution of regional models, which can lead to inaccuracies in local forecasts, especially in complex terrain or urban areas. To overcome these challenges and improve weather forecasting capabilities, several innovative ideas can be explored. One approach involves the integration of emerging technologies such as artificial intelligence and machine learning to enhance the accuracy and resolution of weather models. By leveraging vast amounts of data from various sources, including satellites, sensors, and social media, AI-driven algorithms can improve predictions and provide more localized forecasts. In our weather forecasting project, we incorporated the Decision Tree algorithm as a key component of our predictive modelling approach. We chose this algorithm due to its ability to effectively handle both classification and regression tasks, making it well-suited for the diverse range of forecasting needs inherent in weather prediction. Decision trees offer interpretability, allowing us to understand the logic behind predictions, which is crucial for informing decision-making processes in weather-sensitive contexts. Additionally, decision trees can capture non-linear relationships and interactions between weather variables, enabling more accurate forecasts.

IV. LITERATURE REVIEW

The literature review is essential for project development, providing insights into background, feasibility, and risks. It guides resource allocation and informs technology choices. It supports programmers with valuable guidance and solutions. Overall, it ensures informed decisions, mitigates risks, and enhances the project's chances of success.

In [1] the authors, Chauhan, & Thakur paper deals with the implementation of data mining methods for guiding the path of the ships. The implementation uses a Global Positioning System (GPS) which helps in identifying the area in which the ship is currently navigating. The weather report on that area is compared with the existing database and the decision is made in accordance with the output obtained from the Data Mining technique. This decision about the weather condition of the navigating path is then instructed to the ship. This paper highlights some statistical themes and lessons that are directly relevant to data mining and attempts to identify opportunities where close cooperation between the statistical and computational communities might reasonably provide synergy for further progress in data analysis.

In [2] the authors Vaibhavi Joshi,Bhagyashri Kamble paper the C5 decision tree classification algorithm was used to generate decision trees and rules for classifying weather parameters such as maximum temperature, minimum temperature, rainfall, evaporation and wind speed in terms of the month and year. The data used was for Ibadan metropolis obtained from the meteorological station between 2000 and 2009. The results show how these parameters have influenced the weather observed in these months over the study period. Given enough data the observed trend over time could be studied and important deviations which show changes in climatic patterns identified. This work is important to climatic change studies because the variation in weather conditions in terms of temperature, rainfall and wind speed can be studied using these data mining techniques.

In [3] the authors Gupta, Indumathy deal with the burgeoning research in the fields of Artificial Intelligence and machine learning has given rise to numerous weather prediction models. But the problem of accurately predicting or forecasting the weather still persists. Numerical weather prediction is taking the existing numerical data on weather conditions and applying machine learning algorithms on it to forecast the weather. This paper is the application of machine learning algorithms, linear regression model from statistics, and two optimization techniques, Normal equation method and Gradient descent method to predict the weather on the basis of few parameters. This research suggests and proposes an efficient and accurate weather prediction and forecasting model using linear regression concepts and normal equation model. All these concepts are a part of machine learning.

In [4] paper the authors Vasantha, TamilKodi an imitated structure is made to predict diverse climate conditions transversely over the Indian subcontinent using Data Analysis and Machine learning frameworks, for instance, straight backslide and vital backslide. The standard wellspring of information to be used for direct taking in is to be assembled from data.gov.in, ncdc.noaa.gov and UCI machine learning information vault. The present climate condition parameters ex. temperature, etc are used to fit a model and further using machine learning strategies and extrapolating the information, the future assortments in the parameters are poor down to envision the precipitation conditions for ranchers.

In [5] paper the authors Grover, Kapoor, Weather forecasting is a canonical predictive challenge that has depended primarily on model-based methods. Explore new directions with forecasting weather as a data intensive challenge that involves inferences across space and time. Study specifically the power of making predictions via a hybrid approach that combines discriminatively trained predictive models with a deep neural network that models the joint statistics of a set of weather-related variables. We show how the base model can be enhanced with spatial interpolation that uses learned long-range spatial dependencies. We also derive an efficient learning and inference procedure that allows for large scale optimization of the model parameters. We evaluate the methods with experiments on real-world meteorological data that highlight the promise of the approach.

In [6] paper the authors M.Viswambari, proposed methodology aims at providing an efficient weather forecasting framework for predicting and monitoring the weather attribute datasets to predict rainfall. In the past the parameters of weather were recorded only for the present time only. The future work will explore a working model of selection that can be classifying the framework for continuous monitoring the climatic attributes and also to increase the range of wireless devices using the algorithms to transmit the data. It can be continuously monitored to predict rainfall and generate the report of weather forecasting.

In [7] paper the authors Bhatkhande, S. S, & Hubballi, R. G, used data mining technique and Decision tree algorithm for classifying weather parameters such as maximum temperature, minimum temperature in terms of the day, month and year. The data used from underground weather sites between 2012 and 2015 from different cities. The results show how these parameters have influenced the weather observed in these months over the study period. Given enough data the observed trend over time could be studied and important deviations which show changes in climatic patterns identified. Decision trees prove as an effective method of Decision making in Weather prediction. As decision trees are ideal for multiple variable analyses, it is particularly important in current problem-solving tasks like weather forecasting.

V. PROPOSED WORK

The proposed work for a weather forecasting project involves the tasks and activities required to develop and operate a weather forecasting system. First Acquire weather data from various sources, such as meteorological stations, satellites, weather balloons, and sensors. Preprocess the data. After that develop weather prediction models using machine learning methods. Train and validate the models with real time weather data. Use the models to generate weather forecasts. Develop a user-friendly interface to display weather forecasts.

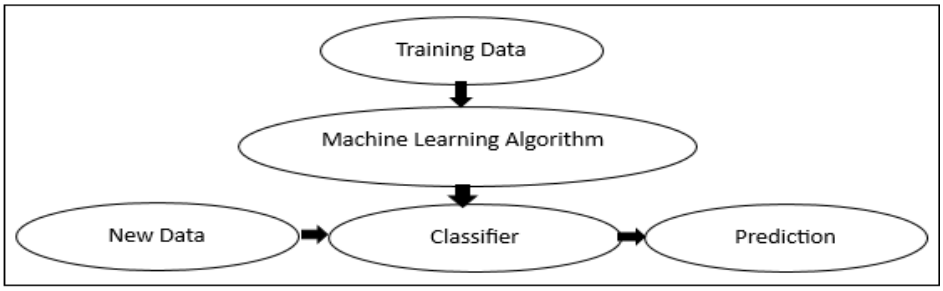

Fig 1. Proposed Work

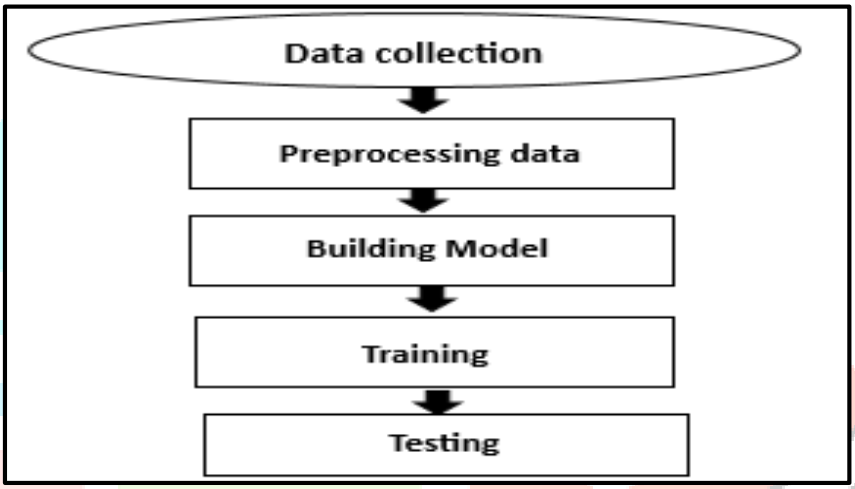

Fig 2. Block Diagram

VI. WORK FLOW

VII. PROPOSED METHODOLOGY

Gather historical weather data from reliable sources such as meteorological organizations, weather stations, or online repositories. Collect data on various meteorological parameters such as temperature, humidity, wind speed, precipitation, atmospheric pressure, etc. Ensure the data covers a diverse range of weather conditions and geographical locations. Clean the collected data by removing any inconsistencies, missing values, or outliers. Convert categorical variables into numerical format using techniques like one-hot encoding. Normalize or standardize numerical features to ensure they are on a similar scale. Identify relevant features that significantly impact weather patterns and forecasting accuracy. Utilize domain knowledge and statistical techniques (e.g., correlation analysis, feature importance from decision trees) to select the most informative features. Implement a decision tree algorithm and Train the decision tree model using the pre-processed historical weather data. Partition the data into training and validation sets to evaluate model performance.

Assess the performance of the decision tree model using appropriate evaluation metrics such as accuracy, precision.

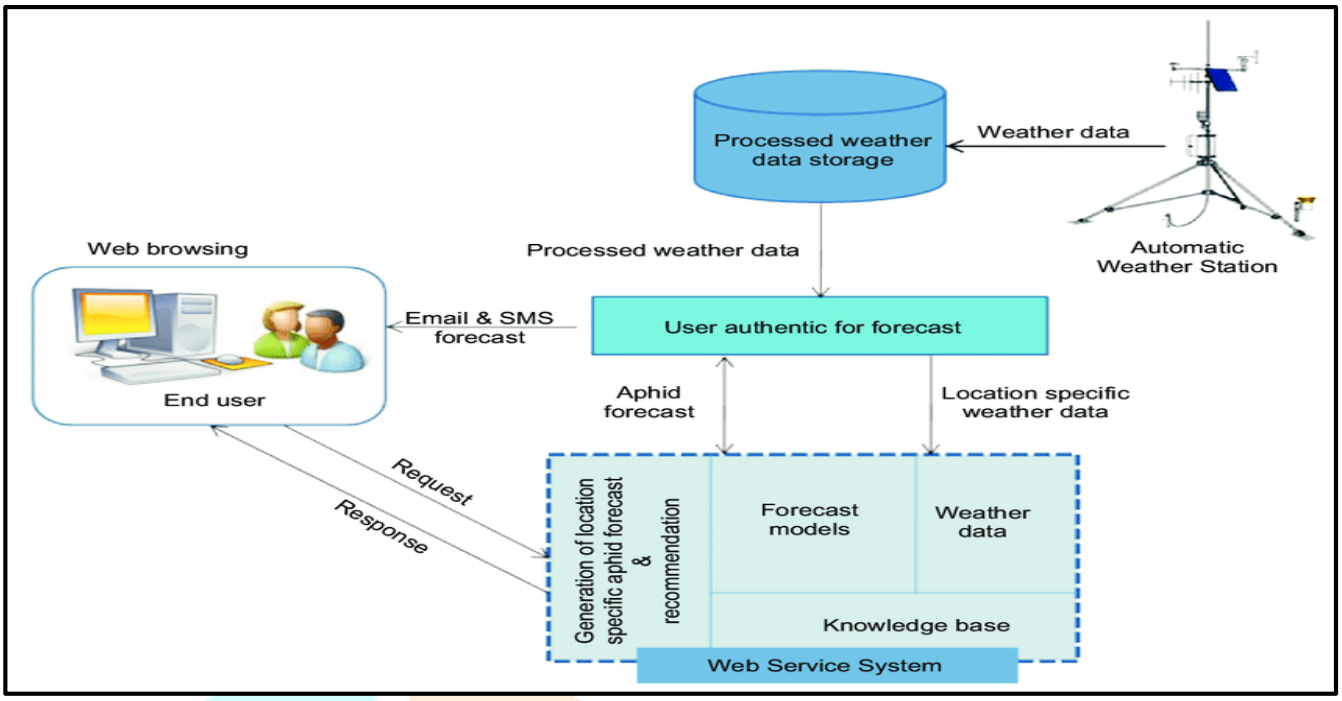

Fig 3. Proposed Methodology

Analyze the confusion matrix to understand the model's predictive strengths and weaknesses for different weather conditions. Explore ensemble learning techniques such as Random Forest or Gradient Boosting, which combine multiple decision trees to improve prediction accuracy and robustness. Integrate an SMS alert mechanism into the forecasting system to notify users of predicted weather conditions. Implement APIs or third-party services for seamless SMS delivery, ensuring timely and accurate alerts based on forecasted data. Establish user authentication protocols to control access to forecast information, safeguarding data privacy and ensuring authorized usage. Implement secure authentication mechanisms (e.g., username/password, token-based authentication) to validate user identities before granting access to forecast data.

VIII. CONCLUSION

In this proposed application we used machine learning algorithm and data mining technique for classifying weather parameters such as maximum temperature, minimum temperature in terms of the day, month and year. The results show how these parameters have influenced the weather observed in the seven day. Given enough data the observed trend over time could be studied and important deviations which show changes in climatic patterns identified. Decision trees prove as an effective method of Decision making in Weather prediction. As, decision trees are ideal for multiple variable analyses, it is particularly important in current problem-solving task like weather forecasting. This work is important to climatic change studies because the variation in weather conditions in term of temperature, rainfall and wind speed can be studied using these data mining techniques. The weather forecasting project represents a significant achievement in the field of meteorology and data science. This application, we have successfully developed a state-of-the-art weather forecasting system that offers accurate and reliable weather predictions to our users.

REFERENCES

[1] Chauhan, D., & Thakur, J. (2014). Data mining techniques for weather prediction A review. International Journal on Recent and Innovation Trends in Computing and Communication, 2(8), 2184-2189.

[2] Ankita Joshi, Bhagyashri Kamble, Vaibhavi Joshi, Komal Kajale, Nutan Dhange, “Weather Forecasting And Climate Changing Using Data Mining Application”, International Journal Of Advanced Research In Computer And Communication Engineering, Vol. 4, Issue 3, March 2015.

[3] Gupta, S., Indumathy, K., & Singhal, G. (2016). Weather prediction using normal equation method and linear regression techniques. IJCSIT) International Journal of Computer Science and Information Technologies, 7(3), 1490-1493.

[4] Vasantha, B., & Tamilkodi, R. (2019, March). Rainfall pattern prediction using real time global climate parameters through machine learning. In 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN) (pp. 1-4). IEEE.
[5] Grover, A., Kapoor, A., & Horvitz, E. (2015, August). A deep hybrid model for weather forecasting. In Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining (pp. 379-386).
[6] M. Viswambari, Dr. R. Anbu Selvi, "Data Mining Techniques To Predict Weather: A Survey", ISSN 2348 – 7968, IJISET- International Journal Of Innovative Science, Engineering & Technology, Vol. 1 Issue 4, June 2014.
[7] Bhatkhande, S. S., & Hubballi, R. G. (2016). Weather prediction based on decision tree algorithms using data mining techniques. International Journal of Advanced Research in Computer and Communication Engineering, 5(5), 483- 487.
[8] Priya, S. B. A survey on weather forecasting to predict rainfall using big data analytics.
[9] Holmstrom, M., Liu, D., & Vo, C. (2016). Machine learning applied to weather forecasting. Meteorol. Appl, 1-5.
[10] Kapoor, P., & Bedi, S. S. (2013). Weather forecasting using a sliding window algorithm. International Scholarly Research Notices, 2013.
[11] Ms. Ashwini Mandale, Mrs. Jadhawar B.A, "Weather Forecast Prediction: A Data Mining Application", International Journal of Engineering Research and General Science Volume 3, Issue 2, March April, 2015, ISSN 2091-2730.
[12] Svetnik V., Liaw A., Tong C., Culberson J., SheridanR., Feuston B. 2003. Random forest: A classification and regression tool for compound classification and QSAR modelling. Journal of Chemical Information and Computer Sciences. 43.: 1947–1958.
[13] Srivastava K., Bhardwaj R., 2014. Real-time nowcast of a cloudburst and a thunderstorm event with assimilation of Doppler weather radar data. Natural hazards. 70(2).: 1357-1383.
[14] Rajesh K., 2013. Decision Tree for the Weather Forecasting. International Journal of Computer Applications (0975 – 8887).
[15] Genuer R., Poggi J.-M., Tuleau C., 2008. Random Forests: Some Methodological Insights.
[16]Khan Z. U., Hayat M., 2014. Hourly based climate prediction using data mining techniques by comprising entity demean algorithms. Middle-East Journal of Scientific Research.
[17] Folorunsho Olaiya, 2012. Application of Data Mining Techniques in Weather Prediction and Climate Change Studies, I.J.
[18] Howarth, E., Hoffman M. S., 1984. A multidimensional approach to the relationship between mood and weather. British Journal of Psychology.
[19] Dhamodaran, Sridhar, Varma, Ch & Reddy, Chittepu., 2020. Weather Prediction Model Using Random Forest Algorithm and GIS.
[20] Durai V.R., Bhardwaj R., 2014. Forecasting quantitative rainfall over India using multi-model ensemble technique. Meteorology and Atmospheric Physics.