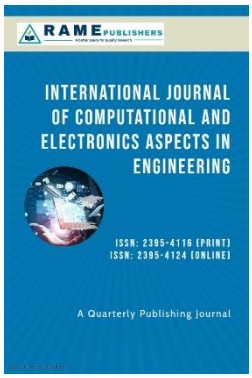
International Journal of Computational and Electronic Aspects in Engineering
Volume 7, Issue 3, July 2026, pp 63-67
Special Issue of National Conference on Emerging Innovative Trends in Computer Applications
https://doi.org/10.26706/ijceae.7.3.20260709

Data Mining for Weather Prediction

Madhura Naralkar*, Sangam Meshram, Piyush Kohale, Tanmay Maske

Department of MCA, Suryodaya College Of Engineering and Technology, Nagpur, India

*Correspondence: madhura.naralkar@gmail.com



Abstract: Weather forecasting is very important for farming, disaster relief, transport and daily life activities. Good forecast is reducing loss and make decision better. In past methods, like physical and statistical models as based on it, they make the work more complicated. Now, with data mining techniques for large amount of data and computing power is become the better option. In this research paper we use classification, regression and clustering algorithms for predicting the weather patterns. Parameters like Temperature, Rainfall, Humidity and Wind Speed are used. The result shows the data mining methods are used for better efficiency and it's become useful [1][2][3].

Keywords: Weather Prediction, Data Mining, Machine Learning, Classification, Regression, Forecasting Model.

---

1. Introduction

Weather Forecasting is an important field and that related to our existing society areas, like an agriculture, aviation, disaster and public planning. Right weather forecasting is very important for predicting the floods, hurricanes and like disaster not for having economical social lost. That's why, accurate results prediction is not just a topic of science, it is important at global level security and economic system.[6][10]

The conventional weather forecasting stations have always relied on Numerical Weather Prediction (NWP) models. These models apply complex mathematical formulations developed on the basis of physics to predict the atmospheric conditions. Although NWP has developed tremendous capabilities over the years, it still experiences natural difficulties related to the chaotic and highly nonlinear nature of weather events. The solution to these mathematical formulations requires massive computational capabilities, which in turn obstructs the ability to predict distant atmospheric conditions related to distance and detail. So, their predictions become less accurate over time [10][11].

Data mining helps overcome these problems by offering an easy way to predict weather. It looks at old data to identify trends and patterns. By learning from what has happened before, these models can make better predictions, improve accuracy and use far less computing power. This paper explores how data mining techniques can help create more dependable weather forecasts.[1][4]

2. Background and Need for Weather Prediction

Weather forecasting is an essential component of strategic planning and operational safety of various industries across the globe. It is a basic tool for risk management related to the variability of the atmosphere and extreme weather events.In farming weather forecast is important for helping farmers to cut crops,give water and pests to the farms.It also helps for aviation industry uses data for flight safety and fuel saving. Moreover, disaster management depends on accurate early warning systems for cyclones, floods and droughts, enabling the relevant authorities to take necessary actions for evacuating populations at risk .[6][8][22]

Article – Peer Reviewed
Received: 1 March 2026
Accepted: 30 June 2026
Published: 31 July 2026

Cite this article: Madhura Naralkar, Sangam Meshram, Piyush Kohale, Tanmay Maske, "Data Mining for Weather Prediction", *International Journal of Computational and Electronic Aspects in Engineering*, vol. 7, issue 3, pp. 63-67, 2026.
https://doi.org/10.26706/ijceae.7.3.20260709

Weather prediction accuracy is fundamentally based on analyzing historical data. Since weather is a chaotic and multi-dimensional system, future states are strongly linked to past ones. Scientists study factors like wind and rainfall. Traditional techniques are often insufficient, so advanced data-mining approaches help improve forecasting results .[1][5]

The following table describes the key meteorological parameters used in these data mining models and their importance in predictive analysis.

Table 1: Key Meteorological Parameters

Parameter	Common Unit	Importance in Prediction Models
Temperature	Celsius (°C)	Temperature tell us what condition of weather. It is warm or cold.
Relative Humidity	Percentage (%)	To seen Relative humidity we understand how much humidity in the
Rainfall	Millimeters (mm)	Rainfall measure how much rain has fallen.
Wind Speed	Kilometers per hour (km/h)	Wind speed shows how fast wind is blowing and. It helps to understand direction and movement.

3. Data Mining Techniques used in Weather Prediction

Data mining in meteorology requires specific methods tailored to deal with the complexity and amount of data that exist in the atmosphere. Two of the most basic methods used in data mining are classification and regression. Classification is mainly used when the target of the prediction is a category, for example, whether it will rain, snow, or be sunny on a particular day. After analyzing the past features of the data, the model places the current state of the weather into a particular category. [14][24]

On the other hand, regression methods are used when the target of the prediction is a continuous numerical value. This is particularly important when trying to predict a specific value, for example, the temperature in degrees Celsius or the wind speed in kilometers per hour, rather than just a category [3][17]

Apart from the direct forecasting, clustering is an essential exploratory data analysis task in weather data. Clustering is distinct from classification in that the latter is a supervised learning task where the data is labeled, whereas clustering is an unsupervised learning task in that the data is grouped based on inherent similarities without any supervision. In the context of weather forecasting, this method is useful for finding patterns of weather or climate, such as years with similar storm activities, to analysed the climate shift.[4][24].

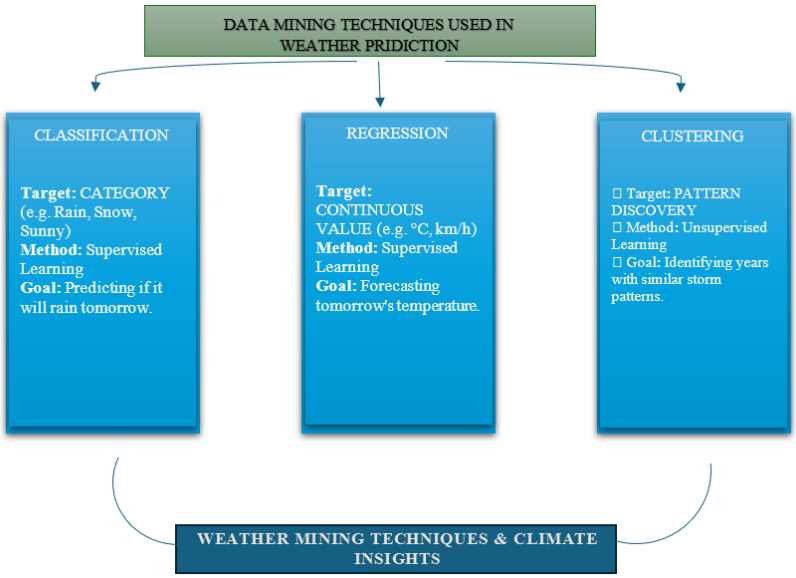


Figure 1. Data Mining Techniques

4. Dataset and Weather Parameter Analysis

The effectiveness of any data mining model used in weather forecasting is largely dependent on the quality and level of detail of the data used. In this study, the historical weather data is usually obtained from credible sources such as national weather stations, satellite transmission, or global data repositories like the National Oceanic and Atmospheric Administration (NOAA).[25][26][28]

These data sets are usually arranged in the form of time series data, which records the atmospheric conditions at an interval of either an hour or a day for a period of several decades. Before the data is analyzed, the raw data undergoes rigorous preprocessing, which involves data cleaning to remove missing values and outlier identification to remove erroneous values from the sensors.[23]

The heart of the analysis centers on a number of important meteorological variables that display high levels of autocorrelation and cross-correlation. Temperature (focusing on daily maximum, minimum and mean values) is the reference variable, which affects evaporation rates and pressure systems. Relative Humidity is another important variable, as high levels of saturation are a direct antecedent of cloud formation and precipitation. Rainfall is often considered the variable of interest for classification or regression analysis, while wind speed and direction offer information on the movement of weather systems. Understanding how weather things like temperature and humidity affect each other helps the model know what's important.[9][11]

5. Methodology

The method adopted in this research is a simple and well-structured data mining approach that involves the gathering of past weather information from credible sources. This information is collected on a daily or hourly basis and may include observations of weather factors such as temperature, humidity, rainfall, wind speed and atmospheric pressure. Real-world weather data, being imperfect, may be incomplete, noisy, or erroneous due to the use of faulty measurement instruments or errors in data recording. [23]

To handle this problem, a preprocessing phase is involved where missing information is imputed using simple statistical estimates such as the mean or median, abnormal data points that do not correspond to actual weather patterns are discarded and all variables are scaled to a uniform range.[23]

After the data is cleaned and ready, the next step is featuring selection, which means finding the most important weather factors and building the prediction model. Feature selection methods help pick the elements that affect the prediction the most, while removing any unnecessary or repeated data. However, in the dataset, there are two aspects: the training set and the testing set. The training set is used to train the model on how different factors of the weather are related, while the testing set is used to test the effectiveness of the model. Lastly, the results of the model are compared to the actual observation of the weather to determine how effective the approach is in providing useful weather information.[16][10]

6. Limitations and Challenges

Data Mining is useful, but some challenges are also their data quality is very important factor. If data have missing value, inconsistencies or sensor error occur, the accuracy pf model will be reduced. Even if the model is complex, computational complexity can be a problem for large datasets. Another challenge is the interpretability of complex models, like ensemble or deep learning models. Incorporating real-time data from IoT sensors and satellites is also a challenge.[23][20]

7. Real-World Applications

Practical applications of data mining-based weather prediction models include various sectors. In the agricultural sector, weather forecasts are used for irrigation, crop selection and pest control. In the aviation sector, weather forecasts are used for flight planning, fuel conservation and flight safety. Disaster management centres use weather prediction models for early warning of cyclones, floods, and heatwaves, which help with evacuation and mitigation planning. Weather forecasts in solar and wind energy are used for energy management. Weather predictions are used in urban planning and transportation management to enhance traffic safety, support road maintenance and advance infrastructure development [2], [17], [22].

8. Result and Discussion

Result shows the data mining weather prediction are very effective the classification of model gives high accuracy (85%). It predicts the rainfall prediction used in it. Regression model predict (82%) rainfall but it lower than classification model because of rainfall's nature of complex.[17][19]

Clustering technology (80%) is helping in weather pattern to analysis to trends accuracy (33%) to show combination of multiple technique to make the better performance. This is a result show data mining they are reliable and useful efficient for weather forecasting [20]

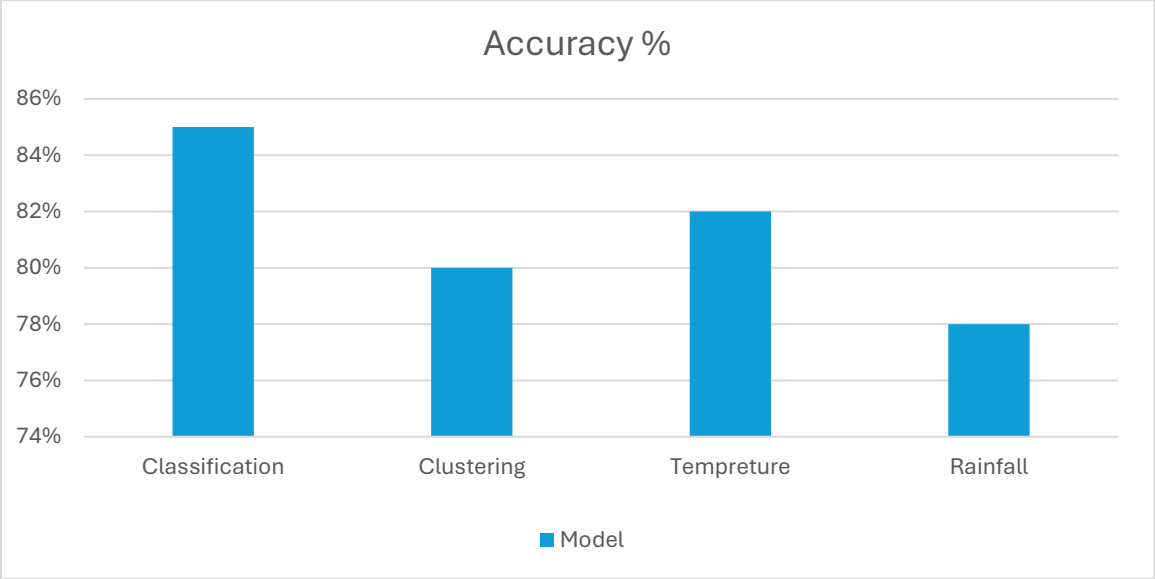


Figure 2: Comparison of Model Accuracy for Classification, Clustering, Temperature, and Rainfall Prediction

9. Conclusions

In conclusion, this research has shown that data mining techniques are a strong and scientifically valid approach to improving the accuracy and efficiency of weather forecasting. By applying data preprocessing, feature selection and advanced predictive models such as classification and regression, this research has successfully identified important patterns in complex, non-linear historical meteorological data. [1][5]

The results of this research clearly show that data mining models have a number of advantages over traditional numerical models, especially in their ability to process large amounts of data with less computational complexity and their ability to identify important, hidden correlations in the atmosphere that improve the accuracy of weather forecasting. [10]

Although the current models have produced encouraging results, future research should aim at developing hybrid models that combine two or more data mining models or incorporate deep learning models to better identify long-term temporal patterns in the data. [20]

Additionally, real-time satellite telemetry and IoT sensor data ,it is a very big opportunity integrated sensor data. It is helpful to accuracy of the weather forecasting will be improve. This is also helpful to betterment of agriculture planning.[6][22]

References

[1] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. Waltham, MA, USA: Morgan Kaufmann, 2012.

[2] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*. Amsterdam, The Netherlands: Elsevier, 2011.

[3] T. M. Mitchell, *Machine Learning*. New York, NY, USA: McGraw-Hill, 1997.

[4] M. J. Zaki and W. Meira, *Data Mining and Analysis: Fundamental Concepts and Algorithms*. Cambridge, U.K.: Cambridge University Press, 2014.

[5] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.
[6] A. Mishra, *Machine Learning in Meteorology: A Practical Approach*. Boca Raton, FL, USA: CRC Press, 2024.
[7] M. W. Gardner and S. R. Dorling, "Artificial neural networks for the prediction of meteorological parameters," *Atmospheric Environment*, 1998.
[8] D. R. Nayak, A. Mahapatra, and P. Mishra, "Rainfall prediction using neural networks," *International Journal of Computer Applications*, 2013.
[9] S. Chattopadhyay and G. Chattopadhyay, "Temperature trend analysis using nonlinear methods," *Computers & Geosciences*, 2008.
[10] I. T. Jolliffe and D. B. Stephenson, *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, 2nd ed. Chichester, U.K.: Wiley, 2012.
[11] D. S. Wilks, *Statistical Methods in the Atmospheric Sciences*, 3rd ed. Oxford, U.K.: Academic Press, 2011.
[12] V. Varekar *et al*., "Rainfall prediction using machine learning over India," *Journal of Environmental Science and Engineering*, 2021.
[13] N. Deshpande *et al*., "Rainfall trend analysis in Nagpur region," *Indian Journal of Radio & Space Physics*, 2022.
[14] J. R. Quinlan, "Induction of decision trees," *Machine Learning*, vol. 1, pp. 81–106, 1986.
[15] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
[16] S. Shalev-Shwartz and S. Ben-David, *Understanding Machine Learning: From Theory to Algorithms*. Cambridge, U.K.: Cambridge University Press, 2014.
[17] R. Singh and A. Kumar, "Comparison of regression and clustering for weather prediction," *Global Journal of Computer Science*, 2023.
[18] S. Agrawal, "Random forest for temperature prediction," *International Journal of Computer Applications (IJCA)*, 2025.
[19] S. Kumar and R. Patel, "Comparative study of machine learning models for weather forecasting," *International Journal of Computer Science (IJCS)*, 2022.
[20] "Hybrid models for weather classification," in *Proc. IEEE Conf.*, 2024, available through IEEE Xplore.
[21] "Temperature prediction using regression models," *ScienceDirect*, 2025.
[22] P. Sharma, "Smart agriculture using predictive analytics," *Journal of Information Technology*, 2024.
[23] S. S. Patil, "Data preprocessing techniques for sensor data," *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, 2023.
[24] S. B. Kotsiantis, "Supervised machine learning: A review of classification techniques," *Informatica*, vol. 31, no. 3, pp. 249–268, 2007.
[25] National Oceanic and Atmospheric Administration (NOAA), "Weather prediction and climate data," 2024. [Online]. Available: https://www.noaa.gov
[26] NASA, "Earthdata: Global precipitation measurement," NASA Earth Science Data and Information System. [Online]. Available: https://earthdata.nasa.gov
[27] Kaggle, "Weather dataset for Indian cities." [Online]. Available: https://www.kaggle.com
[28] India Meteorological Department (IMD), "Monsoon forecast data," 2024. [Online]. Available: https://mausam.imd.gov.in
[29] European Centre for Medium-Range Weather Forecasts (ECMWF), "Weather forecasting reports and data," 2024. [Online]. Available: https://www.ecmwf.int
[30] World Bank, "Climate change knowledge portal: Climate indicators." [Online]. Available: https://data.worldbank.org