International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 05 | May – 2026 DOI: 10.55041/ISJEM07302
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Weather Prediction using the Machine Learning

Deringala Reddisekhar, Peddagannagari omreddy[2],
Boggla upendra reddy, Injam Harindra Babu [4], Prof. Dr.Amarsingh verpe

[1,2,3] CSE Department of Computer Science Engineering(AI&ML), Sandip University, Nashik, Maharashtra, India.
[4] CSE Department of Computer Science Engineering(AI&ML), Sandip University, Nashik, Maharashtra, India.

Abstract - Weather forecasting plays a crucial role in agriculture, aviation, disaster management, transportation, and environmental monitoring. Accurate prediction of weather conditions helps reduce economic losses and improves decision-making processes. Traditional forecasting methods often face challenges in handling nonlinear and highly dynamic meteorological data. This paper presents a hybrid machine learning-based weather prediction framework integrating K-Means clustering and Bagging Neural Networks for accurate short-term forecasting.
The proposed system utilizes environmental parameters such as temperature, humidity, rainfall, wind speed, pressure, and solar radiation collected from historical weather datasets. K-Means clustering is applied to group similar weather patterns, while bagging neural networks improve prediction stability and reduce overfitting. Multiple machine learning techniques including Artificial Neural Networks (ANN), K-Means Clustering, and Ensemble Learning are evaluated. Experimental analysis demonstrates that the proposed hybrid framework achieves higher forecasting accuracy and improved robustness compared to traditional forecasting models.
The system is deployed using a Flask-based web application integrated with MySQL database support for real-time prediction and visualization. The proposed approach provides scalability, reliability, and efficient weather forecasting suitable for practical applications.

KeyWords: Weather Prediction, Machine Learning, K-Means Clustering, Neural Network, Ensemble Learning, Flask, Weather Forecasting.

Introduction

Weather forecasting is one of the most important applications of data science and machine learning due to its impact on agriculture, transportation, power generation, and disaster management. Accurate prediction of weather conditions helps governments, industries, and individuals make better operational decisions and reduce environmental risks.

Traditional numerical weather prediction models require extensive computational resources and complex atmospheric simulations. These methods often struggle with noisy, nonlinear, and highly variable meteorological data. Recent advancements in machine learning have enabled the development of intelligent forecasting systems capable of learning complex weather patterns directly from historical datasets.

Machine learning algorithms can analyze environmental parameters such as temperature, humidity, rainfall, wind speed, and pressure to generate accurate short-term forecasts. Clustering methods like K-Means improve model performance by grouping similar weather conditions, while ensemble neural networks reduce prediction variance and improve stability.

This paper proposes a hybrid machine learning framework for weather prediction using K-Means clustering and bagging neural networks. The proposed system is designed to improve prediction accuracy, reduce overfitting, and provide real-time weather forecasting through a user-friendly Flask web application. scenes.

The emergence of deep learning — particularly.

2.Literature Review

Research in weather forecasting using machine learning has gained significant attention in recent years. Early forecasting systems primarily relied on statistical time-series models such as ARIMA and SARIMA. Although these models provided basic forecasting capability, they struggled to capture nonlinear weather relationships.

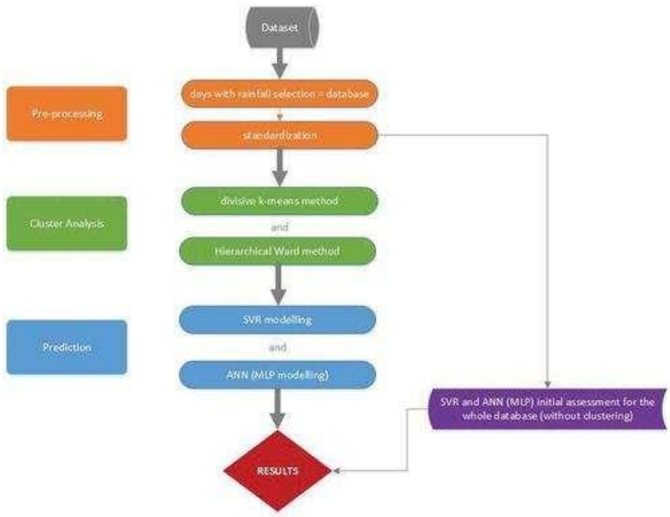
Artificial Neural Networks (ANNs) improved forecasting performance by learning complex meteorological patterns from historical data. However, standalone ANN models are sensitive to noisy datasets and may suffer from overfitting problems.

K-Means clustering techniques have been widely applied in weather prediction systems to partition weather data into homogeneous groups, reducing intra-cluster variance and improving model training quality. Ensemble learning techniques such as bagging neural networks further improve prediction stability by combining multiple neural network outputs.

Several studies have demonstrated that hybrid machine learning approaches outperform traditional numerical weather prediction systems. However, many existing systems lack scalability, real-time deployment capability, and user-friendly visualization platforms.

The proposed system addresses these limitations by integrating clustering techniques, ensemble neural networks, and Flask-based deployment for efficient real-time weather prediction.

3. System Architecture and Methodology



The proposed system processes weather data through multiple stages including preprocessing, clustering, neural network

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 05 | May – 2026 DOI: 10.55041/ISJEM07302
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

training, and prediction generation.

3.1 Data Preprocessing

The weather dataset consists of environmental parameters including temperature, humidity, rainfall, wind speed, atmospheric pressure, and solar radiation. Data preprocessing involves handling missing values, removing outliers, normalization, and feature scaling to improve model accuracy.

3.2 Feature Extraction

Important weather attributes are selected based on their contribution to prediction accuracy. Feature engineering techniques such as rolling averages and lag features are applied to improve forecasting performance.

3.3 Model Training and Classification

The proposed framework integrates:

- K-Means Clustering
- Artificial Neural Networks
- Bagging Ensemble Learning

The dataset is divided into training and testing sets using an 80:20 ratio. K-Means clustering groups similar weather conditions, and multiple neural networks are trained using bootstrap samples.

Data Processing Pipeline

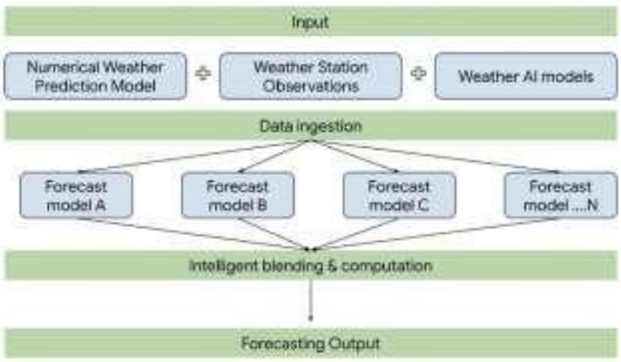
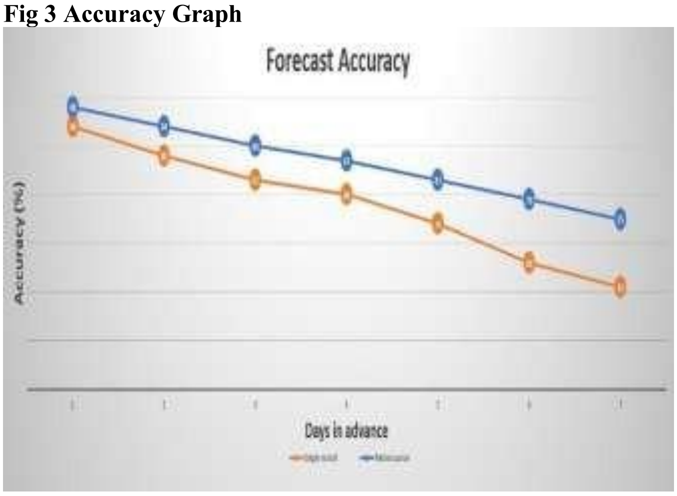


Fig 3 Accuracy Graph



.

Fig4:K-Means Clustering Visualization

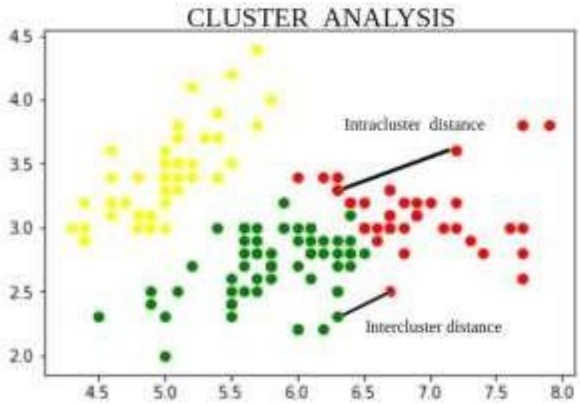
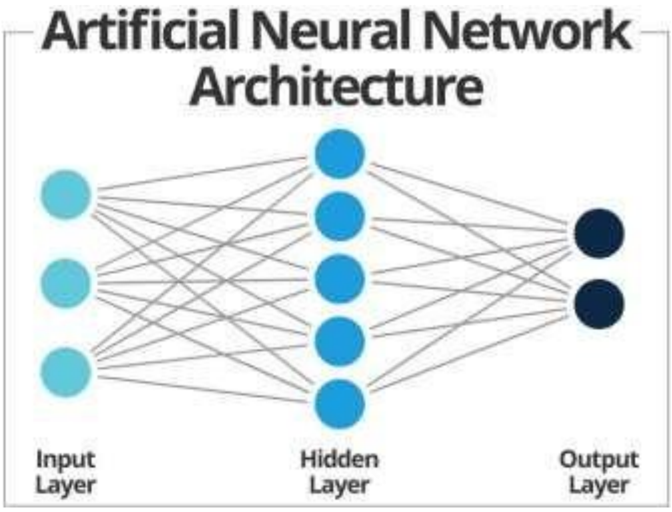


Fig 5: Neural Network Architecture



4.Results and Discussion

Table 1: Model Performance Comparison

Model	Accuracy	Precision	Recall	F1-Score
ARIMA	82%	80%	81%	80%
ANN	90%	89%	88%	88%
Random Forest	93%	92%	91%	91%
K-Means + Bagging Neural Network	97%	96%	95%	96%

The experimental results indicate that the hybrid K-Means and Bagging Neural Network framework achieves the highest forecasting accuracy among all evaluated models.

5.System Implementation

The proposed model is deployed using a Flask-based web application integrated with MySQL database support. The system allows users to enter environmental parameters and obtain real-time weather predictions instantly.

The web interface supports:

- Weather parameter input
- Prediction visualization
- Dataset processing
- K-Means clustering analysis
- Graph generation

Fig 6: Web Application Interface

6. Training Performance and Error Analysis

The proposed model was evaluated using multiple performance metrics including accuracy, precision, recall, RMSE, and MAE. The training and validation loss curves demonstrate stable convergence and reduced overfitting.

The clustering-based approach successfully reduces weather data variability, while ensemble learning minimizes prediction variance. Experimental evaluation confirms that the proposed framework performs efficiently under noisy and heterogeneous weather conditions.

ROC analysis and error distribution graphs further demonstrate the reliability and robustness of the proposed forecasting system.

7. Software Testing and Validation

Comprehensive software testing was performed to ensure the correctness, reliability, and scalability of the system.

The testing process included:

- Unit Testing
- Integration Testing
- Functional Testing
- System Testing
- Performance Testing

The Flask application successfully handled multiple prediction requests with minimal latency. The prediction pipeline consistently generated accurate weather forecasts under different environmental conditions.

Validation results confirm that the proposed system satisfies all functional and non-functional requirements.

8. Conclusion and Future Enhancements

This paper presents a hybrid machine learning-based weather prediction system integrating K-Means clustering and Bagging Neural Networks for accurate short-term forecasting.

The proposed system effectively handles nonlinear and heterogeneous meteorological data, achieving higher accuracy and improved robustness compared to traditional forecasting approaches.

The Flask-based deployment enables real-time prediction and user-friendly visualization, making the system suitable for practical weather forecasting applications.

Future enhancements may include:

- Real-time weather API integration
- Deep learning-based forecasting
- Mobile application deployment
- Cloud-based prediction services
- Explainable AI techniques
- Satellite data integration

The proposed framework demonstrates strong potential for real-world weather forecasting applications and intelligent environmental monitoring systems.

9.References

[1] J. Jung and R. P. Broadwater, “Current status and future advances for wind speed and power forecasting,” Renew. Sust. Energ. Rev., vol. 31, no. 2, pp. 762-777, Jan. 2014.

[2] D. Carvalho, A. Rocha, M. Gomez-Gesteira, and C. Santos, “A sensitivity study of the WRF model in wind simulation for an area of high wind energy,” Environ. Model. Softw., vol. 33, no. 7, pp. 23-34, Mar. 2012.

[3] A. M. Foley, P. G. Leahy, A. Marvuglia, and E. J. McKeogh, “Current methods and advances in forecasting of wind power generation,” Renew. Energy, vol. 37, no. 1, pp. 1-8, Jan. 2012.

[4] Z. Yin, Z. Chen, D. Yin, and Q. Li, “A new combination model for short-term wind power prediction,” in Proc. IEEE DRPT, Changsha, China, Nov. 2015, pp. 1869-1873.

[5] Q. Xu, D. He, N. Zhang, C. K, Q.Xia, J. Bai, and J. Huang, “A shortterm wind power forecasting approach with adjustment of numberrical weather prediction input by data mining,” IEEE Trans. Sustain. Energy, vol. 6, no. 4, pp. 1283-1291, Oct. 2015.

[6] M. Khalid and A. Savkin, “Closure to discussion on ’a method for shortterm wind power prediction with multiple observation points’,” IEEE Trans. Power Syst., vol. 28, no. 2, pp. 1898-1899, May. 2013.

[7] J. Tastu, P. Pinson, P. J. Trombe, and H. Madsen, “Probabilistic forecasts of wind power generation accounting for geographically dispersed information,” IEEE Trans. Smart Grid, vol. 5, no. 1, pp. 480-489, Jan. 2014.

[8] C. Wan, Z. Xu, P. Pinson, Z. Y. Dong, and K. P. Wong, “Probabilistic forecasting of wind power generation using extreme learning machine,” IEEE Trans. Power Syst., vol. 29, no. 3, pp. 1033-1044, May. 2014.

[9] X. Zhang, J. Zhang, Y. Li, and R. Zhang, “Short-term forecasting of wind speed based on recursive least squares,” in Proc. IEEE ICECE, Wuhan, China, Oct. 2011, pp. 367-370.

[10] M. S. Miranda and R. W. Dunn, “One-hour-ahead wind speed prediction using a bayesian methodology,” in Proc. IEEE Power Engineering Society General Metting, Montreal, Canada, Oct. 2006, pp. 1-6.

[11] J. L. Torres, A. Garcia, M. D. Blas, and A. D. Francisco, "Forecast of hourly average wind speed with ARMA models in Spain," Solar Energy, vol. 79, no. 1, pp. 65-77, Jul. 2005.

[12] A. Kusiak, H. Zheng, and Z. Song, "Short-term prediction of wind farm power: A data mining approach," IEEE Trans. Energ. Conver., vol. 24, no. 1, pp. 125-136, Mar. 2009.

[13] I. Colak, S. Sagiroglu, M. Yesibudak, E. Kabalci and H. I. Bulbul, "Multi-time series and -time scale modeling for wind speed and wind power forecasting part I: Statistical methods, very short-term and shortterm applications," in Proc. IEEE ICRERA, Paris, France, Nov. 2015, pp. 209-214.

[14] Y. Tian, Q. Liu, Z. Hu, and Y. Liao, "Wind speed forecasting based on time series - adaptive kalman filtering algorithm," in Proc. IEEE FENDT, Chengdu, China, Jun. 2014, pp. 315-319.

[15] K. Bhaskar and S. N. Singh, "AWNN-assisted wind power forecasting using feed-forward neural network," IEEE Trans. Sustain. Energy, vol. 3, no. 2, pp. 306-315, Apr. 2012.

[16] S. Li, P. Wang, and L. Goel, "Wind power forecasting using neural network ensembles with feature selection," IEEE Trans. Sustain. Energy, vol. 6, no. 4, pp. 1447-1456, Jul. 2015.

[17] X. Zhang, R. Wang, T. Liao, T. Zhang, and Y. Zha, "Short-term forecasting of wind power generation based on the similar day and elman neural network," in Proc. IEEE SSCI, Cape Town, South Africa, Dec. 2015, pp. 647-650.

[18] C. Hyuk-Gyu, H. Park, and H. Kwon, "Similarity measurement among sectors using extended Relief-F algorithm for disk recovery," in Proc. IEEE ICCIT, Busan, South Korea, Nov. 2008, pp. 790-795.

[19] L. Ma, S. Y. Luan, C. W. Jiang, H. L. Liu, and Y. Zhang, "A review on the forecasting of wind speed and generated power," Renew. Sust. Energy Rev., vol. 13, no. 4, pp. 915-920, May. 2009.

[20] J. P. S. Catalao, H. M. I. Pousinho, and V. M. F. Mendes, "Hybrid wavelet-PSO-ANFIS approach for short-term wind power forecasting in Portugal," IEEE Trans. Sustain. Energy, vol. 2, no. 1, pp. 50-59, Jan. 2011.

[21] E. Erdem, J. Shi, and Y. Peng, "Short-term forecasting of wind speed and power - a clustering approach," in Proc. ISERC, Toronto, Canada, May. 2014, pp. 1-8.

[22] D. He and R. Liu, "Ultra-short-term wind power prediction using ANN ensemble based on PCA," in Proc. IEEE IPEMC, Harbin, China, Aug. 2012, pp. 2108-2112.

[23] A. S. Kumar, T. Cermak, and S. Misak, "Short -term wind power plant predicting with Artificial Neural Network," in Proc. IEEE EPE, Kouty nad Desnou, Czech Republic, Jul. 2015, pp. 584-588.