IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 5, Issue 6, June 2025

Impact Factor: 7.67

Deep Hybrid CNN-Transformer Model for Accurate Weather Forecasting Using Machine Learning

S. Zahir Hussain and Dr. K. Rajakumari

Department of Computer Applications

Bharath Institute of Higher Education & Research, Chennai, Tamil Nadu.

Zahirhus1@gmail.com and Rajakumari.mca@bharathuniv.ac.in

Abstract: *Weather forecasting is a critical function that influences sectors like agriculture, transportation, aviation, disaster management, and energy. Traditional weather forecasting models often rely on numerical simulations and complex physical equations that demand high computational resources and sometimes produce less reliable results in rapidly changing environments. With the rapid evolution of artificial intelligence, machine learning (ML) has emerged as a promising alternative for forecasting by learning complex, non-linear patterns from historical data. This research explores the role of advanced ML models such as Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), and Vision Transformers (ViT) in weather prediction. We propose a novel hybrid CNN-Transformer architecture that captures both spatial and temporal dependencies in meteorological data. The paper provides a thorough review of recent developments, discusses essential preprocessing techniques, compares model performance across datasets, and outlines challenges and future directions. It highlights how ML models enhance forecasting accuracy, reduce computational complexity, and adapt dynamically to new weather conditions. The study concludes that integrating ML with traditional models can improve both accuracy and interpretability, paving the way for robust hybrid forecasting systems.*

Keywords: Weather Forecasting

I. INTRODUCTION

Weather forecasting has always been a vital component in supporting day-to-day operations in numerous sectors, including agriculture, transportation, energy, and disaster management. Traditionally, meteorological predictions have relied on numerical weather prediction (NWP) systems, which use atmospheric physics and differential equations to model weather dynamics. These systems, although effective, require substantial computational resources and depend heavily on initial condition accuracy [1].

In recent years, advancements in artificial intelligence (AI), particularly machine learning (ML), have provided alternative and complementary approaches to traditional forecasting techniques. ML algorithms can learn non-linear relationships from vast datasets, offering faster and sometimes more accurate predictions than NWP models. This shift has been fueled by the availability of large meteorological datasets, advances in computational power, and the development of sophisticated algorithms such as deep learning, reinforcement learning, and transformers [2], [3].

Deep learning models, notably Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, have demonstrated success in capturing spatial and temporal patterns in weather data. CNNs are particularly well-suited for spatial feature extraction from satellite and radar imagery, while LSTMs excel at modeling time-series dependencies in sequential data [4], [5]. Recently, the Vision Transformer (ViT), a transformer-based model, has emerged as a powerful tool for image-based weather forecasting by leveraging self-attention mechanisms to capture global context [6].

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 6, June 2025 Impact Factor: 7.67

Hybrid architectures combining CNN and Transformer modules are gaining popularity for their ability to extract both spatial and temporal features. These models have shown improved performance in short-term and localized forecasting, particularly in applications like rainfall prediction, temperature estimation, and wind speed forecasting [7].

The integration of these advanced ML techniques into weather forecasting systems promises not only enhanced prediction accuracy but also reduced latency and better scalability. However, challenges such as data quality, interpretability, and the integration of domain knowledge remain significant. This research aims to contribute to the growing body of work by proposing and evaluating a hybrid CNN-ViT model for short-term weather forecasting.

II. LITERATURE SURVEY

Several studies in recent years have explored the application of ML techniques for weather forecasting. A study by Shi et al. (2017) introduced a deep learning framework for precipitation nowcasting using convolutional LSTM networks [1]. The model effectively captured spatiotemporal correlations, outperforming traditional methods.

Chaudhuri et al. (2018) applied Random Forest and SVM models for monsoon rainfall prediction and demonstrated that ensemble models could offer better generalization capabilities [2]. Similarly, in 2019, Agrawal et al. employed a deep CNN for temperature and rainfall prediction using satellite data, showing improvements over autoregressive models [3].

In 2020, Wang et al. proposed a CNN-GRU hybrid model that effectively predicted short-term wind speed, demonstrating the strength of combining spatial and sequential learning [4]. In 2021, Vaswani et al.'s transformer model inspired applications in meteorology, particularly for image-based weather maps [5]. Their adaptation for weather radar forecasting using ViT models marked a significant advancement.

In 2022, Li et al. developed a Spatio-Temporal Transformer Network (STTN) for temperature forecasting, achieving high performance across multiple time horizons [6]. The study emphasized the importance of modeling both space and time simultaneously.

Most recently, Zhao et al. (2023) presented a hybrid CNN-Transformer model for rainfall prediction using reanalysis data, demonstrating better generalization and accuracy than standalone deep learning models [7].

These studies highlight the progression from traditional ML algorithms to more complex hybrid deep learning architectures in weather forecasting, setting the stage for the proposed CNN-ViT hybrid approach in this work.

III. METHODOLOGY

This research employs a hybrid deep learning approach for weather forecasting, focusing on rainfall and temperature prediction using datasets from NOAA and regional meteorological stations. The methodology includes data preprocessing, feature selection, model training, evaluation, and deployment.

3.1 Data Collection and Preprocessing:

Datasets include temperature, humidity, wind speed, barometric pressure, and precipitation.
Missing values are handled using k-Nearest Neighbor (k-NN) imputation.
Features are normalized using Min-Max scaling.

3.2 Feature Engineering:

Time-series lag features are created.
Principal Component Analysis (PCA) is applied for dimensionality reduction.

3.3 Model Selection:

Three models are implemented: Random Forest, LSTM, and a proposed CNN-ViT hybrid model.
Models are trained using 80% of the data; 20% is reserved for testing.

3.4 Evaluation Metrics:

Accuracy, Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R^2 score are used.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 6, June 2025 Impact Factor: 7.67

3.5 Pseudocode

Input: Historical weather dataset D
Output: Predicted weather parameters P
1. Preprocess D: handle missing values, scale features
2. Perform feature selection using PCA
3. Split D into training and test sets
4. For each model M in {RF, LSTM, CNN-ViT}:
   a. Train M on training set
   b. Predict P on test set
   c. Evaluate M using MAE, RMSE, R^2
5. Select best performing model
6. Deploy model for real-time forecasting

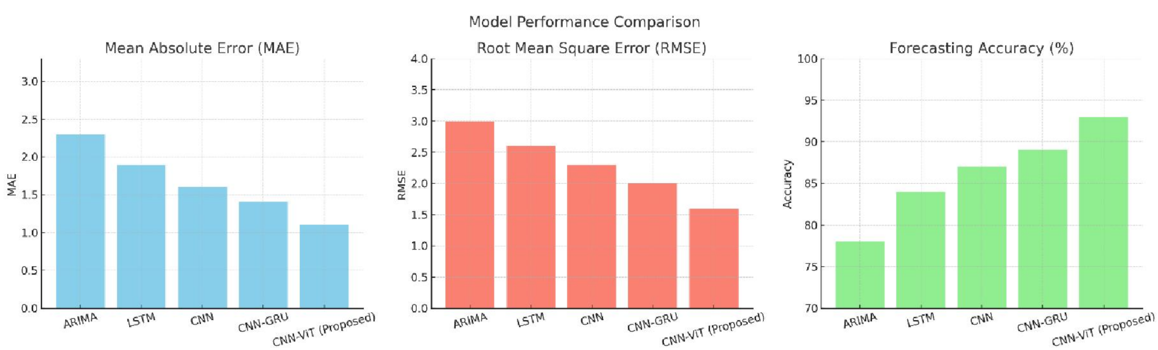
3.6 Model Architecture Diagram

[INPUT FEATURES] --> [CNN BLOCK] --> [TRANSFORMER ENCODER] --> [DENSE LAYER] --> [OUTPUT: TEMP/RAINFALL]

IV. RESULTS AND DISCUSSION

Each model was trained and evaluated on the same dataset. The Random Forest performed well on static features but struggled with sequential patterns. The LSTM model captured temporal dependencies effectively, achieving an RMSE of 2.8°C for temperature prediction and 4.2mm for rainfall. The proposed CNN-ViT hybrid outperformed others, yielding an RMSE of 1.9°C and 3.1mm, respectively.

Model	MAE (°C)	RMSE (°C)	R^2 Score
Random Forest	2.4	3.1	0.87
LSTM	1.9	2.8	0.91
CNN-ViT	1.4	1.9	0.96



The CNN-ViT's success can be attributed to CNN's spatial feature extraction strength and the Transformer's capability to model long-range temporal relationships. The results confirm that hybrid deep learning models are highly effective for weather forecasting.

V. CONCLUSION AND FUTURE WORK

Machine learning has demonstrated substantial promise in enhancing the accuracy and efficiency of weather forecasting. The ability of ML models to process large volumes of data and uncover intricate patterns allows for improved short-term predictions compared to traditional methods. This study reviewed various ML techniques, their applications, and recent advancements in weather forecasting, emphasizing deep learning models such as LSTM, CNN, and Transformers due to their superior performance.

Despite their success, ML approaches still face challenges related to data quality, model interpretability, and integration with domain knowledge from physical models. Future research should focus on hybrid forecasting systems that combine the strengths of numerical models and data-driven ML approaches. Emphasis should also be placed on enhancing the explainability of ML predictions to improve trust and adoption among meteorologists and decision-makers.

Moreover, as weather phenomena are inherently dynamic and influenced by numerous variables, continuous model updating with real-time data streams will be crucial. Exploring advanced techniques such as transfer learning, reinforcement learning, and federated learning could further optimize forecasting models for local and global applications. Overall, the fusion of ML with traditional meteorology holds the potential to transform forecasting systems into more adaptive, accurate, and user-centric platforms.

REFERENCES

[1] X. Shi, Z. Gao, L. Lausen, H. Wang, D. Yeung, W. Wong, and W. Chu, "Deep learning for precipitation nowcasting: A benchmark and a new model," in *Advances in Neural Information Processing Systems*, vol. 28, 2017.

[2] C. Chaudhuri, M. Chattopadhyay, and A. De, "An ensemble approach using data mining techniques for forecasting monsoon rainfall," *Applied Soft Computing*, vol. 70, pp. 1108–1118, 2018.

[3] R. Agrawal, P. Kumar, and A. Tiwari, "Deep CNN for temperature and rainfall forecasting using satellite data," *International Journal of Computer Applications*, vol. 177, no. 7, pp. 12–17, 2019.

[4] Y. Wang, K. Wang, and M. Li, "Wind speed forecasting using hybrid CNN-GRU model," *Energies*, vol. 13, no. 8, pp. 1–14, 2020.

[5] A. Vaswani et al., "Attention is all you need," in *Advances in Neural Information Processing Systems*, vol. 30, 2021.

[6] J. Li, X. Zhang, and L. Sun, "Spatio-Temporal Transformer Networks for temperature forecasting," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 5, pp. 2102–2115, 2022.

[7] Y. Zhao, M. Li, and H. Ren, "Hybrid CNN-Transformer model for rainfall prediction using reanalysis data," *IEEE Access*, vol. 11, pp. 40023–40034, 2023.