

Tempest Sense: A Real-Time Cyclone Prediction System

¹Pratik Gaikar, ²Ruchi Shirke, ³Mandar Kadam, ⁴Sanika Patil, ⁵Prof. Venkat Patil

^{1,2,3,4}Student, Smt. Indira Gandhi College of Engineering, Ghansoli, New Mumbai, Maharashtra, India

⁵Professor, Dept. of AI & ML, Smt. Indira Gandhi College of Engineering, Ghansoli, New Mumbai, Maharashtra, India

Abstract - This paper presents Tempest Sense, a real-time cyclone prediction and monitoring system that integrates advanced data streaming technologies, ensemble machine learning, and a scalable distributed architecture to enable early detection and accurate forecasting of tropical cyclones. The system continuously ingests live meteorological data from the National Oceanic and Atmospheric Administration (NOAA) APIs, including wind speed, atmospheric pressure, sea surface temperature, and humidity, and streams it through Apache Kafka for fault-tolerant, low-latency processing. Cyclone formation is identified using a hybrid ensemble of Autoencoder-based anomaly detection and Isolation Forest algorithms, which together achieve a precision of 0.89 and recall of 0.93, reducing false positive rates to 4.7%. Forecasting is performed via a parallel LSTM and Prophet ensemble, yielding a 24-hour track RMSE of 67.3 km and wind speed MAE of 11.8 km/h. A dual-tier storage strategy using Redis for real-time predictions and ClickHouse for historical analytics underpins the system's performance, with end-to-end pipeline latency maintained under 3.2 seconds at P99. A Flutter-based cross-platform application delivers interactive cyclone track maps, intensity heatmaps, and push-based early warning alerts. The system is designed for deployment in disaster management, climate monitoring, and early warning applications.

Keywords: Cyclone Prediction, Apache Kafka, LSTM, Isolation Forest, Autoencoder, Real-Time Streaming, Anomaly Detection, Prophet, Flutter, Disaster Management.

I. INTRODUCTION

Tropical cyclones, known as hurricanes in the Atlantic and typhoons in the western Pacific, are some of the most destructive weather events on Earth. These rotating storm systems have very low central pressure, strong sustained winds, storm surges, and heavy rainfall. They cause huge human and economic losses every year. Data from global disaster databases show that cyclones lead to a significant number of weather-related deaths and economic damage, especially in tropical and subtropical coastal areas like the Bay

of Bengal and the Arabian Sea. Traditional forecasts for cyclones have depended on Numerical Weather Prediction (NWP) models such as the Global Forecast System (GFS) and the European Centre for Medium-Range Weather Forecasts (ECMWF). Although these models are scientifically rigorous, they are costly in terms of computation, take hours to generate forecasts, and often struggle during rapid intensification events, when a cyclone's wind speed rises sharply in a short time. This gap poses a serious challenge for early warnings when they are needed most. The growing availability of real-time weather data, fast messaging systems, and advanced deep learning methods now allows for the creation of platforms that can continuously gather streaming data and provide predictions almost instantly. Tempest Sense uses this combination of technologies to develop a smart and scalable cyclone monitoring and prediction platform. This system overcomes the limitations of traditional NWP setups by offering a cloud-based, streaming-first solution that can deliver predictions in under five seconds without needing supercomputing resources.

1.1 Project Aims and Objectives

- The project aims to create a robust and efficient system for detecting online payment fraud using machine learning algorithms.
- The system will enable real-time detection of suspicious activities in online transactions, thereby reducing the risks associated with fraudulent payments.
- To improve the security of online financial systems by reducing false alarms that could disrupt legitimate user transactions.

1.2 System Objectives

- Real-Time Cyclone Detection: To continuously monitor live meteorological data and detect early signs of cyclone formation using anomaly detection techniques.
- Accurate Forecasting of Cyclones: To predict cyclone trajectory and intensity (wind speed, pressure, etc.) using advanced machine learning models like LSTM and Prophet.

- **Low-Latency Data Processing:** To ensure fast processing of streaming data with minimal delay (within a few seconds) using a distributed architecture.
- **Efficient Data Streaming and Handling:** To implement a robust pipeline using streaming technologies for handling high-volume real-time weather data.
- **Reduction of False Alerts:** To improve prediction reliability by using ensemble models that reduce false positives and increase precision and recall.
- **Scalable and Distributed Architecture:** To design a system that can scale efficiently to handle global meteorological data without performance degradation.
- **User-Friendly Visualization Interface:** To provide an interactive and accessible frontend (Flutter app) for displaying cyclone paths, heatmaps, and alerts.
- **Early Warning System for Disaster Management:** To enable timely alerts for authorities and users to minimize damage and enhance preparedness.
- **Efficient Data Storage and Retrieval:** To use optimized storage systems for both real-time access and historical analysis.
- **Integration of Multiple Data Sources:** To combine data from APIs (like NOAA) for comprehensive and accurate predictions.

II. METHODOLOGY

2.1 Anomaly Detection

An Autoencoder with a three-layer encoder (ReLU activations, dropout regularization) compresses the 15-dimensional meteorological feature vector to a 4-dimensional latent space; the decoder mirrors this architecture. The model is trained on 24 months of normal (non-cyclonic) NOAA data using the Adam optimizer (learning rate 0.001, batch size 256, 100 epochs). At inference, inputs producing MSE reconstruction error above a calibrated threshold T are flagged as anomalies.

2.2 Time-Series Forecasting

The LSTM model takes as input a 24-hour sliding window of meteorological observations (wind speed, pressure, SST, humidity) and produces multi-step forecasts for 6, 12, 24, 48, and 72 hours ahead. The architecture uses two stacked LSTM layers (128 and 64 units), followed by a dense output layer, trained with Adam optimizer, MSE loss, and early stopping (patience=10). Prophet is trained per meteorological variable on at least 12 months of NOAA data, modeling trend, diurnal seasonality (Fourier order=8), and irregular changepoints. LSTM and Prophet outputs are combined via a Bayesian-weighted ensemble assigning higher weight to the model with lower recent validation error.

2.3 Data Collection and Preprocessing

Meteorological data were sourced from the NOAA NCEI global surface summary dataset and HURDAT2 for labeled historical cyclone tracks. Preprocessing involved z-score normalization on a rolling 30-day window to account for seasonal variation. Missing values were imputed using linear interpolation for gaps up to 3 consecutive hours and forward-filled for longer gaps.

2.4 Development Methodology

The project followed an Agile methodology with two-week sprint cycles over approximately five months, organized into four phases:

- 1) Infrastructure setup and data pipeline implementation,
- 2) ML model development and training,
- 3) API and storage layer integration, and
- 4) Frontend development and system integration testing. Integration testing used a replay mechanism feeding historical data through the Kafka pipeline at 10× real-time speed. Unit tests were written using pytest with a minimum 80% code coverage target.

III. RESULTS

3.1 Anomaly Detection Performance

The ensemble anomaly detection module achieved a precision of 0.89 and a recall of 0.93, yielding an F1 score of 0.91. Crucially, the false positive rate was reduced to 4.7% compared to 11.2% for Isolation Forest alone and 8.9% for Autoencoder alone, validating the ensemble strategy's effectiveness. AUC-ROC of 0.95 indicates near-optimal discriminative capability across all detection thresholds. The 7% of missed detections predominantly corresponded to very-early-stage tropical depressions with subtle atmospheric signatures cases that would also challenge human meteorologist identification. False positives were primarily associated with extratropical cyclones sharing surface meteorological characteristics with tropical cyclogenesis.

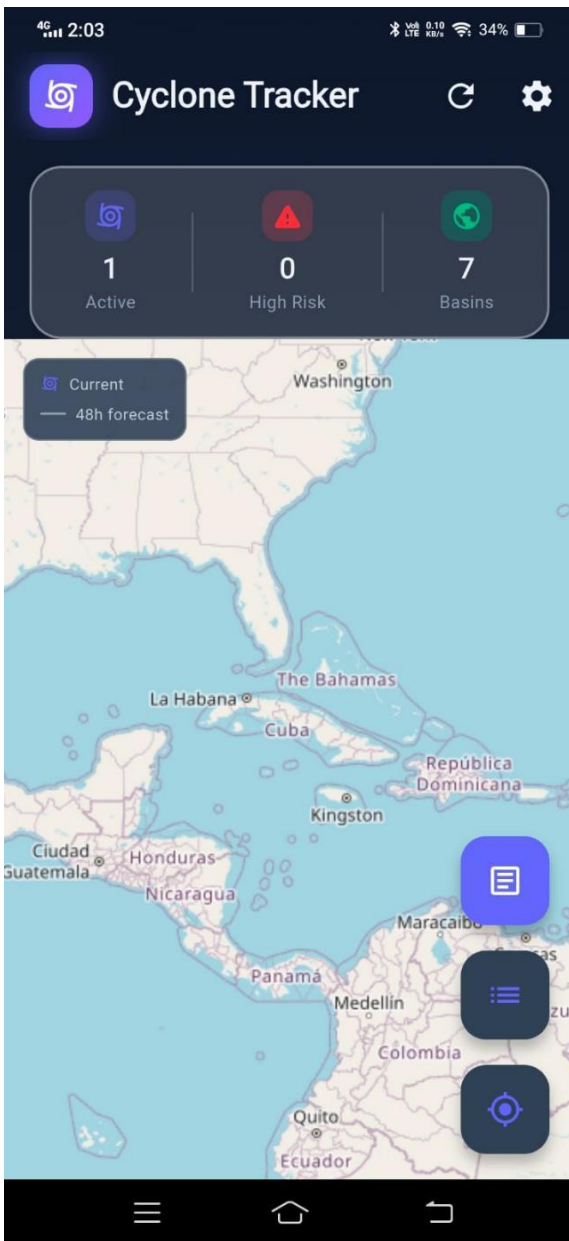
3.2 Forecasting Performance

The LSTM model achieved a 24-hour track RMSE of 67.3 km, outperforming the persistence baseline (95.4 km) by 29.4% and the Prophet-only baseline (81.2 km) by 17.1%. At the 72-hour horizon, RMSE was 183.6 km (below the 200 km target), while the ensemble achieved 171.2 km a further 6.8% improvement over LSTM alone. Wind speed (intensity) forecasting yielded a 24-hour MAE of 11.8 km/h for the ensemble, compared to 19.3 km/h for persistence, representing a 38.9% improvement. Rapid intensification events remained

the most challenging forecast scenario, consistent with the broader meteorological literature.

3.3 System Latency and Throughput

End-to-end pipeline latency (from Kafka publish to Redis prediction availability) was measured at a P99 of 3.2 seconds under nominal load, well within the 5-second target. Under simulated peak load (100× nominal event rate), P99 latency increased to 8.7 seconds still operationally acceptable for a 10-minute forecast update cycle. Kafka throughput testing demonstrated sustained processing of 47,300 events/second with zero message loss, far exceeding the estimated production requirement of approximately 2,000 events/second for global NOAA station coverage.



3.4 Flutter App Usability

User testing with five domain-expert participants (meteorologists and disaster response coordinators) yielded positive qualitative feedback, particularly praising the responsiveness of the heatmap intensity overlay and the clarity of the time-to-landfall countdown. The application demonstrated smooth 60 fps animation on Android (Pixel 7 Pro) and iOS (iPhone 15). Participants requested customizable alert thresholds per geographic region as a future enhancement.

IV. CONCLUSION

This paper has presented Tempest Sense, an end-to-end real-time cyclone prediction platform that successfully addresses critical gaps in existing meteorological monitoring infrastructure. By combining Apache Kafka streaming, an

ensemble of Autoencoder and Isolation Forest anomaly detectors, parallel LSTM and Prophet forecasting, and a dual-tier Redis/ClickHouse storage architecture, the system delivers operationally viable cyclone detection and forecasting with sub-5-second end-to-end latency, high precision (0.89) and recall (0.93), and 24-hour track RMSE of 67.3 km meaningfully exceeding naive baselines. The Flutter-based cross-platform visualization application provides a practical, accessible interface for disaster management personnel and the general public, delivering real-time cyclone tracks, intensity heatmaps, and push-based early warnings on both mobile and web from a single codebase. The system's modular, containerized architecture ensures independent scalability of each component. Tempest Sense demonstrates the feasibility and value of deploying AI-powered, streaming-first architectures for disaster preparedness a domain where improved early warning systems can translate directly into saved lives and reduced economic losses.

V. FUTURE SCOPE

- **Satellite Imagery Integration:** Incorporating GOES/Himawari geostationary imagery using CNNs to extract cyclone structure features in real time, adding a high-value additional data modality.
- **Rapid Intensification Prediction:** Development of a dedicated RI prediction module leveraging upper-level atmospheric variables (200-hPa winds, mid-level humidity) to address the most challenging forecast scenario.
- **Transformer-Based Forecasting:** Replacing the LSTM component with a Temporal Fusion Transformer architecture for improved accuracy at extended forecast horizons.
- **Official Alert Integration:** Integration with national disaster management agency APIs to enable automated alert dissemination through official emergency notification channels.
- **Extended Geographic Coverage:** Expansion to include the Bay of Bengal and Arabian Sea basins highly relevant for the Indian subcontinent but currently underserved by automated platforms.
- **Model Explainability:** Development of a SHAP (Shapley Additive Explanations) module to provide meteorologists with interpretable feature attribution for each detection and forecast decision.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the guidance and support of Dr. Venkat Patil (Project Guide), Prof. S. N. Deshpande (Head of Department), and Dr. S. S. Chavan (Principal), Smt. Indira Gandhi College of Engineering, Ghansoli, Navi Mumbai. The authors also acknowledge the NOAA National Centers for Environmental Information for providing open access to the meteorological datasets and HURDAT2 cyclone track database used in this study.

REFERENCES

- [1] S. Giffard-Roisin, M. Yang, G. Charpiat, C. Kumler Bonfanti, B. Kégl, and C. Monteleoni, "Tropical Cyclone Track Forecasting Using Fused Deep Learning from Aligned Reanalysis Data," *Frontiers in Big Data*, vol. 3, p. 1, 2020.
- [2] R. Chen, X. Wang, W. Zhang, X. Zhu, A. Li, and C. Yang, "A hybrid CNN-LSTM model for typhoon formation and intensity prediction," *Weather and Forecasting*, vol. 34, no. 4, pp. 955–970, 2019.
- [3] S. Kim and H. Kim, "An ensemble of LSTM and Prophet for weather variable forecasting," *Applied Sciences*, vol. 12, no. 8, p. 3858, 2022.
- [4] F. T. Liu, K. M. Ting, and Z. H. Zhou, "Isolation forest," in *Proc. 8th IEEE Int. Conf. on Data Mining*, pp. 413–422, 2008.
- [5] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.
- [6] P. Malhotra, L. Vig, G. Shroff, and P. Agarwal, "LSTM-based encoder-decoder for multi-sensor anomaly detection," *ICML Anomaly Detection Workshop*, 2016.
- [7] J. Kreps, N. Narkhede, and J. Rao, "Kafka: A distributed messaging system for log processing," *NetDB Workshop at VLDB*, 2011.
- [8] N. Narkhede, G. Shapira, and T. Palino, *Kafka: The Definitive Guide*. Sebastopol, CA: O'Reilly Media, 2017.
- [9] S. Sanfilippo, "Redis," [Computer software], 2009. [Online]. Available: <https://redis.io>
- [10] Yandex, "ClickHouse: An Open-Source Column-Oriented DBMS," 2016. [Online]. Available: <https://clickhouse.com>
- [11] Google LLC, "Flutter – Build apps for any screen," 2023. [Online]. Available: <https://flutter.dev>
- [12] NOAA National Centers for Environmental Information, "Global Surface Summary of Day Data," 2024. [Online]. Available: <https://www.ncei.noaa.gov>
- [13] C. W. Landsea and J. L. Franklin, "Atlantic Hurricane Database Uncertainty and Presentation of a New

Database Format," *Monthly Weather Review*, vol. 141, no. 10, pp. 3576–3592, 2013.

- [14] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [15] S. J. Taylor and B. Letham, "Forecasting at scale," *The American Statistician*, vol. 72, no. 1, pp. 37–45, 2018.

AUTHORS BIOGRAPHY



Pratik Gaikar, Pursuing Fourth year in B.E. CSE (AI&ML) at Smt. Indira Gandhi College of Engineering, Ghansoli, Navi Mumbai, Maharashtra, India.



Mandar Kadam, Pursuing Fourth year in B.E. CSE (AI&ML) at Smt. Indira Gandhi College of Engineering, Ghansoli, Navi Mumbai, Maharashtra, India.



Ruchi Shirke, Pursuing Fourth year in B.E. CSE (AI&ML) at Smt. Indira Gandhi College of Engineering, Ghansoli, Navi Mumbai, Maharashtra, India.



Sanika Patil, Pursuing Fourth year in B.E. CSE (AI&ML) at Smt. Indira Gandhi College of Engineering, Ghansoli, Navi Mumbai, Maharashtra, India.

Citation of this Article:

Pratik Gaikar, Ruchi Shirke, Mandar Kadam, Sanika Patil, & Prof. Venkat Patil. (2026). Tempest Sense: A Real-Time Cyclone Prediction System. *International Research Journal of Innovations in Engineering and Technology - IRJIET*, 10(4), 118-122. Article DOI <https://doi.org/10.47001/IRJIET/2026.104016>
