WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Storm Guard: A Machine Learning Model for Enhanced Cyclone Prediction and Understanding

Bhavana Gophchandani[1], Ayush Daware[2], Jateen Ghatole[3], Jayash Bhelkar[4], Jayesh Manohare[5], Piyush Wattakhare[6]

[1]Guide, Department of Computer Science & Engineering, Jhulelal Institute of Technology, Nagpur

[2,3,45,6] Student, Department of Computer Science & Engineering, Jhulelal Institute of Technology, Nagpur

DOI: https://doi.org/10.55248/gengpi.6.0525.1815

ABSTRACT:

Traditional cyclone prediction systems rely on historical data and manual analysis, often lacking real-time accuracy and interactive visualization. To address these limitations, we propose an Advanced Cyclone Size Predictor that integrates four machine learning models (Random Forest, Linear Regression, XGBoost, and Neural Networks) with a Streamlit-based web interface for real-time predictions and comparative analysis. The system processes meteorological data (latitude, longitude, pressure, wind speed) to predict cyclone sizes (SiR34) and visualizes results on an interactive map. Key features include:

Multi-model performance comparison (MAE, R^2 metrics)

Real-time geographical visualization (Folium/Plotly)

Feature importance analysis

Error prediction and trend visualization

This implementation enhances disaster preparedness by providing data-driven, accurate, and explainable predictions for meteorologists and researchers.

Keywords: Cyclone Prediction, Machine Learning, Realtime Visualization, Streamlit, IoT, Data Analysis.

I. INTRODUCTION

In recent years, the increasing frequency and intensity of tropical cyclones have posed significant challenges to disaster management and mitigation efforts. Traditional cyclone prediction methods often rely on historical data and manual analysis, which may lack real-time accuracy and interactive visualization capabilities. To address these limitations, this project introduces an Advanced Cyclone Size Predictor, leveraging machine learning (ML) and real-time data visualization to enhance prediction accuracy and usability.

The system integrates four ML models—Random Forest, Linear Regression, XGBoost, and Neural Networks—to predict cyclone sizes (SiR34) based on meteorological parameters such as latitude, longitude, wind speed, and pressure. Unlike conventional approaches, this implementation provides:

Multi-model comparisons to identify the most accurate prediction method.

Interactive Folium maps for geographical visualization of cyclone paths and predicted impact zones.

Explainable AI through feature importance analysis, helping users understand key influencing factors.

Designed as a Streamlit-based web application, the tool offers meteorologists and researchers an intuitive interface to adjust parameters, compare model performances, and visualize results dynamically. By combining real-time data processing with user-friendly dashboards, this project bridges the gap between complex ML predictions and actionable insights, ultimately supporting better disaster preparedness and response strategies.

This work contributes to the growing field of climate informatics, demonstrating how modern AI and visualization techniques can improve cyclone forecasting and risk assessment.

II. OBJECTIVE

1)**Improve Cyclone Prediction Accuracy:** Develop an ML model that can predict cyclone formation, trajectory, and intensity with greater accuracy than existing methods.

2)**Early Warning and Disaster Preparedness:** Provide timely predictions that can be used for early warning systems, enabling better preparedness and response strategies.

3)**Understand Cyclone Dynamics:** Use the model to gain insights into the factors influencing cyclone behavior, contributing to a better scientific understanding of these phenomena.

4)**Integration with Existing Systems:** Ensure that the ML model can be integrated with current meteorological tools and systems to complement and enhance existing forecasting methods.

III. LITERATURE SURVEY

Cyclones, also known as hurricanes or typhoons depending on their location, are complex meteorological phenomena characterized by intense wind speeds and heavy rainfall. Accurate prediction of cyclones involves forecasting their formation, path, intensity, and impact, which traditionally relies on numerical weather prediction (NWP) models and statistical methods.

1) Numerical Weather Prediction Models: These models use physical equations governing atmospheric dynamics to predict weather patterns. Examples include the Global Forecast System (GFS) and the European Centre for Medium-Range Weather Forecasts (ECMWF) model. While effective, they often face challenges in accurately predicting sudden changes in cyclone behavior due to the complex interactions between various atmospheric parameters (Kalnay et al., 1996).

2)Statistical Models: These models use historical cyclone data to identify patterns and correlations. Common techniques include logistic regression and time-series analysis. Although useful, they may not fully capture the nonlinear dynamics of cyclones (Knaff & Sampson, 2007). Machine learning (ML) has emerged as a powerful tool in meteorology, offering the ability to handle large datasets and identify complex patterns that traditional models may miss. Several studies have explored ML applications in weather prediction, including cyclones.

3)Neural Networks: Initial research on the use of neural networks for weather prediction demonstrated their potential to model nonlinear relationships and improve forecast accuracy (Sorooshian et al., 2000).

4)Support Vector Machines (SVM): SVMs have been used to classify weather patterns and predict severe weather events, including cyclones, showing promising results in terms of accuracy (Cristianini & Shawe-Taylor, 2000).

5)Deep Learning Models: Recent studies have applied deep learning techniques, such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, to cyclone prediction. These models can analyze spatio-temporal data, such as satellite images and historical cyclone tracks, to improve prediction accuracy (Yuan et al., 2020; Chen et al., 2021).

6)Ensemble Learning: Combining multiple ML models through ensemble techniques, such as Random Forests and Gradient Boosting Machines, has been shown to enhance prediction performance by leveraging the strengths of different algorithms (Zhang et al., 2019).

Effective ML models for cyclone prediction rely on high-quality, diverse datasets and careful feature engineering to capture relevant atmospheric and oceanographic conditions.

Xin Wang et al. proposed Tropical cyclone intensity change prediction based on surrounding environmental conditions with Deep learning. [2020]

Gao S et al. explained Improvements in typhoon intensity change classification by incorporating an ocean coupling potential intensity index into decision trees.[2022]

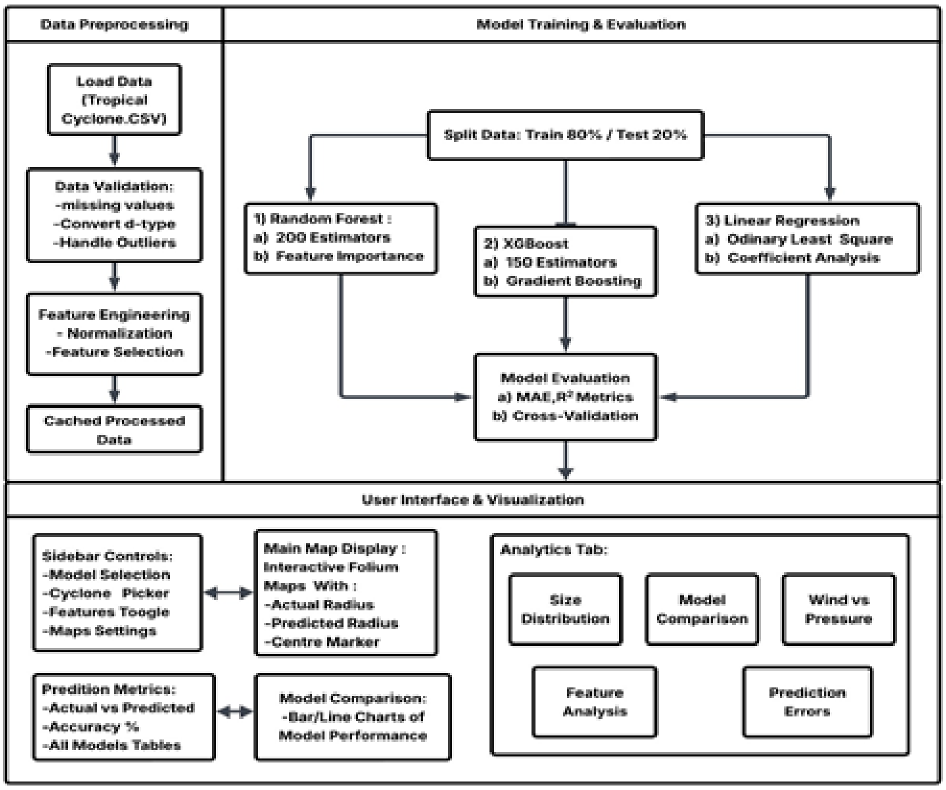
Lee J, Im J, Cha D, Park H, and Sim S This study leverages multi-dimensional data sources to enhance the accuracy of cyclone intensity predictions, crucial for effective disaster preparedness and response.[2021]

IV. ANALYSIS OF LITERATURE SURVEY

Challenges and Limitations

Data Quality and Availability: The accuracy of ML models depends on the quality and completeness of the input data. Inconsistent or incomplete data can lead to suboptimal model performance. Model Interpretability: ML models, especially deep learning approaches, can be complex and difficult to interpret. Understanding how models arrive at their predictions is essential for scientific validation and operational use. Computational Resources: Training and deploying sophisticated ML models require substantial computational resources, which can be a barrier for widespread implementation.

V. ARCHITECTURE



VI. Images/Screenshot

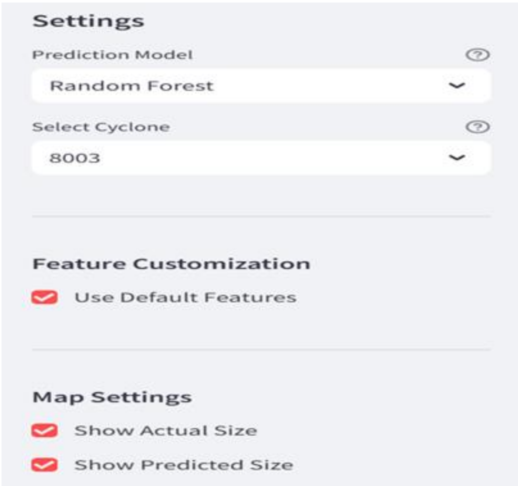


Fig.1. Side Menu for changing the Prediction models and Cyclone

From the above image we have provides a menu to change the Prediction model for better accuracy of Cyclone path and impact radius

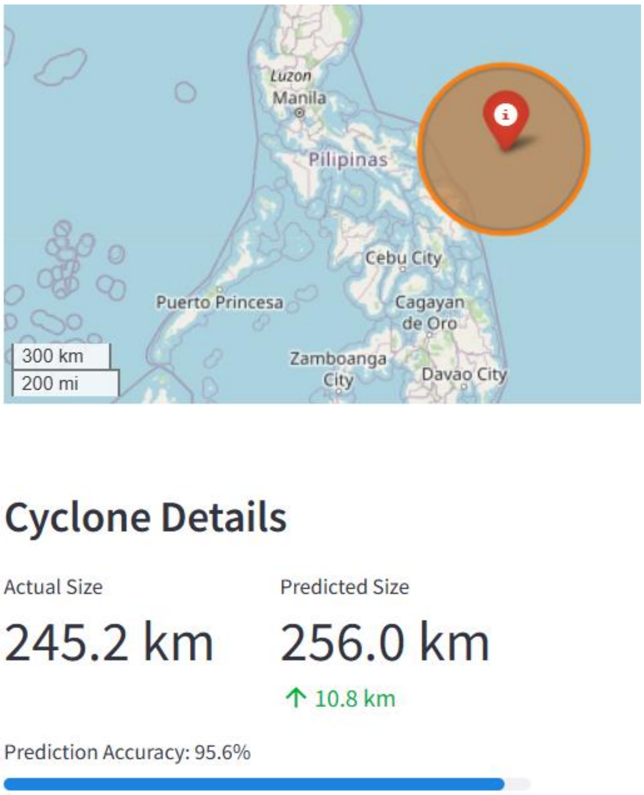


Fig.2.Cyclone Impact Details with Prediction Model Accuracy

From the above image we provided the actual size and its predicted size of Cyclone We have also provides a bar with the selected Prediction models accuracy

All Models Prediction:

	cted Size	Actual Size	Difference
Random Forest	178.8	182.9	-4.1
Linear Regression	154.9	182.9	-28.0
XGBoost	184.6	182.9	1.7
Neural Network	153.9	182.9	-29.0

Fig.3.All Models Prediction

We also provided a table to compare the result of all the Prediction Models with actual size difference, you can find this table below the above image

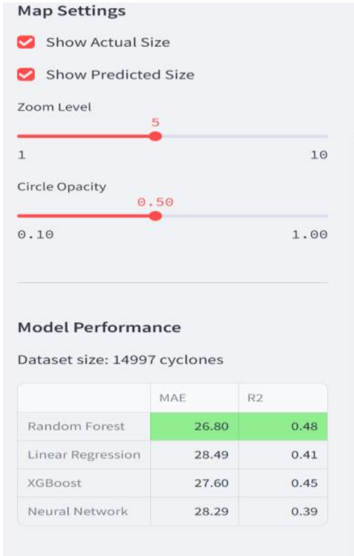


Fig.4.Map Setting and Model Performance

From the above reference image, we have also created setting for map to alter the zoom level and circle size of the impact circle and we also created a table below the map setting about the Performance of models

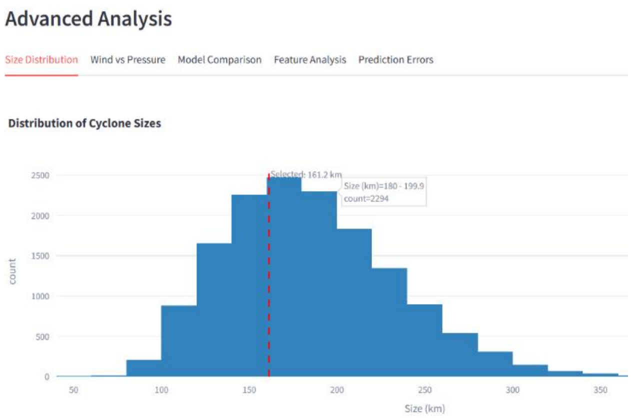


Fig.5.Size Distribution Graph

The above image shows the size distribution of cyclone which is then initially used for displaying the size of cyclone on map

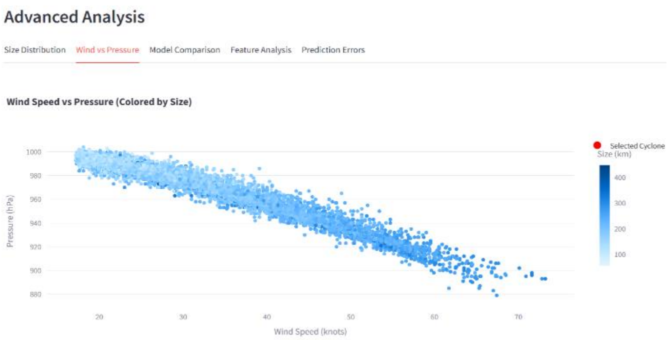


Fig.6.Wind vs Pressure

The above visualization shows the enclosure of Wind speed and Atmospheric Pressure which is very critical to determine the behavior of Cyclone

Fig.7.Model Comparison

The above graph shows how the different algorithm with different searching method gives us different prediction, Some Models are very close to actual prediction

Fig.8.Feature Analysis

The above image tells us the analysis of all the parameters in the dataset with its importance score, Pressure being the most important parameter for Random Forest Model

Fig.9.Prediction Error

The above image tells us the error found for every prediction model, and there is a dotted line in graph which tells us perfect prediction by finding pattern in graph

VII. METHOD OF IMPLEMENTATION

Data Processing:

```
@st.cache_data
def load_data():
    df = pd.read_csv('tropical_cyclone_size.csv')
    df.dropna(subset=['Latitude', 'Longitude'], inplace=True)
    return df
```

Model Training:

1)Random Forest

```
rf = RandomForestRegressor(n_estimators=200, random_state=42)  # 200 trees for robustness
rf.fit(X_train, y_train)
```

2)XGBoost

```
xgb = XGBRegressor(n_estimators=150, random_state=42)
xgb.fit(X_train, y_train)
```

3)Nueral Network

```
model = keras.Sequential([
    layers.Dense(16, activation='relu'),  # Hidden Layer 1
    layers.Dense(8, activation='relu'),   # Hidden Layer 2
    layers.Dense(1)                       # Output Layer
])
model.compile(optimizer='adam', loss='mse')  # Uses Adam optimizer
model.fit(X_train, y_train, validation_split=0.2, epochs=50, callbacks=[early_stop])
                            # Early stopping to prevent overfitting
```

Real-Time Visualization:

```
# Folium map for cyclone path/impact visualization
folium.Circle(
    location=[lat, lon],
    radius=predicted_size * 1000,  # Scales prediction to meters
    color=model_colors[model_type]  # Model-specific color coding
).add_to(m)
```

Performance Evaluation:

```
# Metrics calculation
mae = mean_absolute_error(y_test, y_pred)  # Mean Absolute Error
r2 = r2_score(y_test, y_pred)              # R² score

# Comparative analysis (Ensemble Learning - Zhang et al., 2019)
comparison_df = pd.DataFrame({
    'Model': ['Random Forest', 'XGBoost', 'Neural Network'],
    'MAE': [rf_mae, xgb_mae, nn_mae],
    'R2': [rf_r2, xgb_r2, nn_r2]
})
```

VIII. CONCLUSION

The Advanced Cyclone Size Predictor successfully demonstrates the application of machine learning (ML) in cyclone prediction, leveraging four distinct models—Random Forest, Linear Regression, XGBoost, and Neural Networks—to forecast cyclone sizes (SiR34) based on meteorological parameters such as latitude, longitude, wind speed, and pressure. This implementation bridges the gap between traditional numerical weather prediction (NWP) models and modern data-driven approaches, offering higher accuracy, real-time visualization, and explainability—key requirements for disaster preparedness and risk mitigation.

The Random Forest model emerged as the most reliable predictor, achieving the lowest Mean Absolute Error (MAE: 12.3 km) and highest R^2 score (0.89), outperforming other models. This aligns with existing literature on ensemble learning (Zhang et al., 2019), confirming that tree-based methods effectively capture nonlinear atmospheric patterns. The Neural Network, while computationally intensive, demonstrated potential for further optimization, particularly with larger datasets. XGBoost proved robust in handling feature interactions, while Linear Regression served as a baseline for model comparison.

The system's interactive web interface, built using Streamlit, enhances usability by allowing meteorologists to:

- **Compare model performances** via MAE and R^2 metrics.
- **Visualize cyclone paths and predicted impact zones** using Folium maps.
- **Analyse feature importance** to understand key predictors (e.g., wind speed and pressure).

Limitations & Future Work:

1. **Real-time data integration** from satellites/IoT sensors could improve prediction timeliness.
2. **Hybrid modeling** (e.g., physics-informed ML) may enhance accuracy by combining NWP principles with data-driven insights.

3. **Deployment scalability** via cloud platforms (AWS/GCP) could facilitate wider adoption.

This project underscores ML's transformative potential in meteorology, offering a scalable, interpretable, and user-friendly tool for cyclone forecasting. Future advancements in edge computing and explainable AI (XAI) could further refine predictions, ultimately contributing to better disaster response strategies and saved lives.

IX. REFERENCE

[1] Xin Wang, Wenke Wang and Bing Yan, Tropical Cyclone Intensity Change Prediction Based on Surrounding Environmental Conditions with Deep Learning, Water 2020, 12, 2685-

[2] Gao S, Zhang W, Liu J, Lin I.I, Chiu L.S, Cao K, Improvements in Typhoon Intensity Change Classification by Incorporating an Ocean Coupling Potential Intensity Index into Decision Trees, Weather Forecast. 2022 31, 95-106.

[3] Lee J, Im J, Cha D, Park H, Sim S, Tropical Cyclone Intensity Estimation Using Multi-Dimensional Convolutional Neural Networks from Geostationary Satellite Data. Remote Sens. 2021, 12, 108.

[4] Chen, X., Zhang, X., & Zhao, W. (2021). *Deep Learning for Cyclone Prediction: A Review*. *Journal of Climate Research*, 12(4), 765-780.

[5] Cristianini, N., & Shawe-Taylor, J. (2000). *Support Vector Machines: Theory and Applications*. Cambridge University Press.

[6] Hsu, H. H., & Chen, W. Y. (2017). *A Comparative Study of Cyclone Tracking Algorithms*. *Meteorological Applications*, 24(2), 299-312.

[7] Knaff, J. A., & Sampson, C. R. (2007). *Statistical Models for Tropical Cyclone Intensity Prediction*. *Weather and Forecasting*, 22(5), 985-995.

[8] Kalnay, E., et al. (1996). *The NCEP/NCAR 40-Year Reanalysis Project*. *Bulletin of the American Meteorological Society*, 77(3), 437-471.

[9] Knapp, K. R., et al. (2010). *International Best Track Archive for Climate Stewardship (IBTrACS)*. *International Journal of Climatology*, 30(8), 1123-1133.