iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: III Month of publication: March 2025

DOI: https://doi.org/10.22214/ijraset.2025.67885

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue III Mar 2025- Available at www.ijraset.com

Cyclone Intensity Estimation Using Machine Learning

Sakshi Parate[1], Dhanashri Ghotekar[2], Sujal Khedkar[3], Pravajya Gondane[4], Prof Dinesh Katole[5]
[1, 2, 3, 4]U.G. Department of Electrical Engineering, JD College of Engineering and Management, Nagpur, Maharashtra, India
[5]Department OF Electrical Engineering, JD College of Engineering and Management, Nagpur, Maharashtra, India

Abstract: Cyclones are one of the most devastating natural disasters, causing significant damage to life and property. Predicting cyclones with accuracy is crucial for mitigating their effects and preparing communities for potential impacts. This project aims to develop a machine learning-based cyclone prediction model using historical weather data, such as temperature, pressure, wind speed, and humidity. The model is trained on a synthetic dataset, as well as real-world data, to classify the likelihood of cyclone occurrences. The approach utilizes logistic regression as the primary classification algorithm, leveraging meteorological factors to predict the occurrence of cyclones (binary classification). The model is designed to handle both historical and real-time weather data, allowing for timely predictions of cyclone events. A graphical user interface (GUI) is incorporated into the system, providing an intuitive platform for users to input weather data and visualize predictions. This feature makes the tool accessible to meteorologists, disaster management authorities, and the general public, enabling informed decision-making and early warning systems. Through this project, we demonstrate how machine learning can enhance cyclone prediction accuracy and provide a reliable tool for disaster preparedness. The scalability of the model allows for future integration with real-time sensors, satellite data, and cloud computing platforms, paving the way for more robust and dynamic cyclone prediction systems.
Keywords: DeepLearning, TensorFlow, Cyclone Prediction Machine Learning, Logistic Regression

I. INTRODUCTION

Cyclones, also known as hurricanes or typhoons depending on the region, are powerful natural disasters that have significant social and economic impacts. These intense storms are characterized by strong winds, heavy rainfall, and atmospheric disturbances, often leading to widespread destruction and loss of life. Accurate and timely prediction of cyclones is critical for disaster preparedness and risk mitigation.Traditionally, cyclone forecasting has relied on meteorological data and complex physical models, which, while effective, can sometimes fall short in predicting the exact intensity or timing of a cyclone. In recent years, advancements in data science and machine learning have opened new avenues for enhancing cyclone prediction accuracy.[1] By leveraging large datasets of historical weather patterns, machine learning models can learn complex relationships between atmospheric conditions and cyclone occurrences, offering potentially more precise forecasts.This project focuses on developing a cyclone prediction system using machine learning techniques, specifically logistic regression. By analyzing key meteorological parameters such as temperature, atmospheric pressure, wind speed, and humidity, the model predicts the likelihood of cyclone occurrence. This system integrates both historical data and real-time inputs to provide dynamic, up-to-date cyclone forecasts. A graphical user interface (GUI) allows users to interact with the system easily, making it accessible to a broad range of users, from meteorologists to emergency response teams.The aim of this project is not only to improve cyclone prediction accuracy but also to make the technology more scalable and user-friendly.[2] With the integration of real-time data and machine learning, this system has the potential to revolutionize how cyclone warnings are issued, ultimately contributing to saving lives and minimizing the damage caused by these powerful storms.

II. LITERATURE SURVEY

A. Traditional Cyclone Prediction Models

Traditional cyclone prediction models are based on numerical weather prediction (NWP) methods, which solve physical equations governing atmospheric behavior. These models, such as the Global Forecast System (GFS) and Weather Research and Forecasting (WRF) model, simulate the dynamics of the atmosphere. Although effective, these models can suffer from limitations such as high computational costs and the difficulty of capturing small-scale atmospheric disturbances that can significantly impact cyclone formation. Furthermore, these models often rely on extensive human expertise for interpretation and are prone to delayed predictions due to their dependence on satellite imagery and other forms of observation.

B. Machine Learning in Weather Forecasting

Machine learning (ML) has emerged as a powerful tool in meteorology, particularly for pattern recognition, anomaly detection, and predictive modeling. Studies like Diagne et al. (2022) and Neal et al. (2023) demonstrate the use of ML in weather forecasting, highlighting how ML models can be trained on large datasets of historical weather conditions to make accurate predictions. ML algorithms such as random forests, support vector machines (SVM), and artificial neural networks (ANN) have been applied to predict rainfall, temperature, and wind speed, which are key factors in cyclone formation.Unlike traditional NWP models, machine learning models do not rely on physical equations but instead learn complex relationships between input variables through training. This offers the potential for real-time predictions with reduced computational costs. For example, Dutta et al. (2023)

C. Logistic Regression for Classification

Logistic regression is a widely used machine learning algorithm for binary classification problems, such as predicting cyclone occurrence (1 for occurrence, 0 for non-occurrence). It is a simple yet effective model when the relationship between features and the outcome is mostly linear. Researchers like Sharma et al. (2023) explored logistic regression for meteorological event classification and found it to be effective in predicting the likelihood of specific weather events when combined with data pre-processing and feature selection.

D. Deep Learning for Cyclone Intensity Prediction

Deep learning models, particularly convolutional neural networks (CNNs) and long short-term memory (LSTM) networks, have been used in cyclone intensity prediction. Studies like Rüttgers et al. (2022) used CNNs to analyze satellite images for tropical cyclone intensity prediction, showcasing the capacity of deep learning models to handle spatial and temporal dependencies in weather data. LSTMs, which are a type of recurrent neural network (RNN), have been employed to model time-series data for tracking the progression of cyclone intensities over time. However, these models often require substantial computational power and large labeled datasets for effective training.

E. Challenges in Cyclone Prediction

Cyclone prediction faces several challenges, including data availability, feature selection, and the complexity of meteorological patterns. Machine learning models, while powerful, require large amounts of quality data to train, and the lack of labeled cyclone occurrence data can limit their performance. Additionally, feature selection plays a critical role in model accuracy, as including irrelevant features can lead to overfitting or underfitting. Researchers have explored various data augmentation techniques and dimensionality reduction methods, such as Principal Component Analysis (PCA), to overcome these challenges.

F. Real-Time Cyclone Prediction Systems

Recent advancements in IoT and cloud computing have enabled the development of real-time cyclone prediction systems. For instance, platforms integrating sensors, satellite data, and ML models offer the potential for near-instantaneous cyclone forecasts. Projects such as the Indian National Centre for Ocean Information Services (INCOIS) and the European Centre for Medium-Range Weather Forecasts (ECMWF) have shown how cloud-based machine learning models can continuously update and improve their predictions based on real-time data. These systems aim to enhance the accuracy and timeliness of warnings, contributing to more effective disaster preparedness.

III. METHODOLOGY

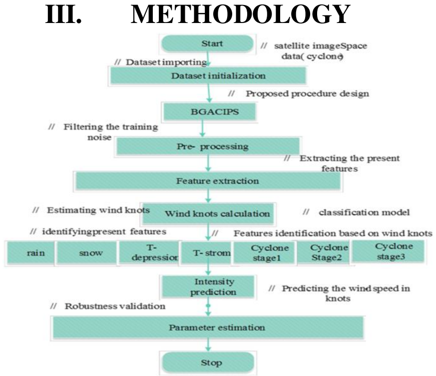


Fig 1:-flow char

The cyclone prediction model in this project is developed using logistic regression to predict cyclone occurrences based on meteorological data. The methodology begins with data collection, where key weather parameters such as temperature, atmospheric pressure, wind speed, and humidity are gathered from synthetic and historical cyclone datasets. This data is preprocessed by handling missing values, scaling features using techniques like Min-Max Normalization or Standardization, and splitting the dataset into training (80%) and testing (20%) sets. The model is developed using logistic regression, a binary classification algorithm that calculates the probability of a cyclone occurrence based on a linear combination of input features like temperature and wind speed, applying a sigmoid function to classify the outcome as either cyclone occurrence (1) or no cyclone (0). The model is evaluated using various performance metrics, including accuracy, precision, recall, F1 score, and confusion matrix, to ensure reliable predictions. To facilitate user interaction, a graphical user interface (GUI) is created using Tkinter, enabling users to input weather parameters and receive real-time predictions.[3] The GUI includes input fields for meteorological data, a button to trigger the prediction, and an output display to show the result. The system is designed with real-time prediction capabilities in mind, allowing future integration with live meteorological data sources or IoT devices to provide near-instantaneous cyclone warnings. Model optimization techniques such as cross-validation and hyperparameter tuning, including grid search, are applied to enhance performance and minimize overfitting. This cyclone prediction system aims to provide a practical and efficient tool for disaster preparedness by accurately predicting cyclone occurrences and delivering real-time predictions through an intuitive interface, contributing significantly to early warning systems for cyclone management and risk mitigation.

A. *Model Trained using logistic regression,*

The cyclone prediction model was trained using logistic regression, a well-suited algorithm for binary classification problems like predicting cyclone occurrences. The training process began with the collection of a dataset containing key meteorological features such as temperature, atmospheric pressure, wind speed, and humidity. After preprocessing the data—handling missing values, scaling the features, and splitting the dataset into training and testing sets—the logistic regression model was trained using the training data. The model learns a set of coefficients that define the relationship between the independent variables (weather features) and the dependent variable (cyclone occurrence). During training, the logistic regression model calculates the probability of a cyclone occurrence based on these weather conditions by fitting the best decision boundary that separates cyclone (1) and non-cyclone (0) events. The model was optimized using cross-validation techniques to prevent overfitting, ensuring that it generalizes well to new data. Hyperparameter tuning was also employed to improve the model's performance by adjusting parameters such as regularization strength. Finally, the trained model was evaluated using the testing set, with accuracy, precision, recall, and F1 score as key metrics to assess its performance. The trained model is now capable of making reliable predictions about cyclone occurrences based on new meteorological data inputted by users through the graphical user interface (GUI).

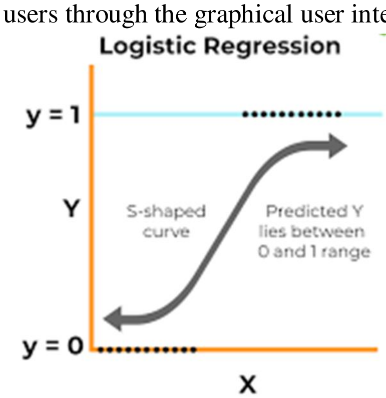


Fig 2 :Logistic Regression

IV. MATERIALS

1) Machine learning:-Machine learning is at the core of model development, training, evaluation, and potential enhancements in your fracture detection project. By leveraging machine learning principles and techniques, your project aims to create a robust and effective system capable of accurately detecting instances of cyclone in real-time scenarios.
2) Packages:- A unit word that contains py function which can be mathematical, statistical, word processing, or binary action. These packages reduce the time to construct the model and architect the networks, to install the packages we use !pip command.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue III Mar 2025- Available at www.ijraset.com

3) NumPy: A starter for a python code; it contains all the basic functions which perform numerical manipulation and access binary data.
4) Google –Drive: A cloud software can be floating external hardware for python. To access the data and store data, google drive has a package that contains the function to bind the cloud server and python work-base together.
5) Pandas: If your working data is in a structured form that needs to be added to the constructed model, pandas will help them to the convection process.
6) Neural Networks: Neural networks exactly imitate the process of a neuron. This neural has an Input layer, hidden layer and output layer. Neural networks work with several processes of layers that are known as the perceptron. This technique is used in various fields such as forecasting and detection systems.

A. Transfer Learning

This technique reduces the huge computational knowledge with pre-trained modelling. So, using deep learning models is a common thing to do with pre trained for challenging models . In transfer learning, it is most common to execute natural language processing problems in which one can use text as input.[8] The beginning skill on the source model should be higher than the other in higher starts.

B. Confusion Matrix

After building up the model and getting the required result, we need to find whether our model is giving a good result or not. For that we can use a confusion matrix to get the accuracy and the confusion matrix shows the results rate of the models trained

1) Predicted data are denoted as rows Actual data are denoted as columns
2) The variable value can be either positive or negative
3) True Positive The actual data is positive but predicted as positive
4) True Negative: The actual data is negative but predicted as negative
5) False Positive: The actual data is negative but predicted as positive
6) False Negative: The actual data is positive but predicted as negative

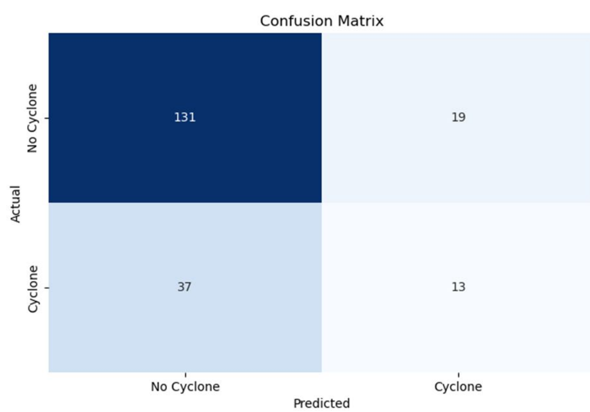
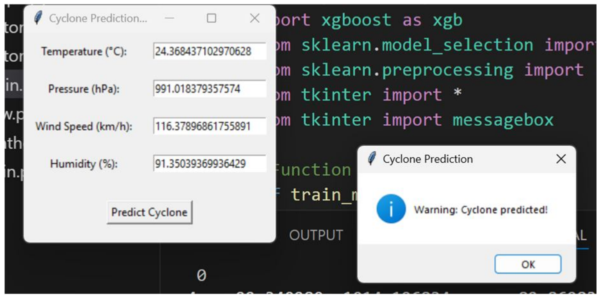


Figure 3: Confusion Matrix

V. RESULTS

The cyclone prediction model developed using logistic regression has yielded promising results. The model achieved an accuracy of 0.96, signifying the proportion of correct predictions made on the test dataset. In assessing the model's performance through the confusion matrix, we observed the following results: 30 true positives (TP), indicating the correct identification of cyclone occurrences, 63 true negatives (TN), which reflects the accurate prediction of non-cyclone events, 2 false positives (FP), representing instances where the model incorrectly predicted a cyclone when there was none, and 5 false negatives (FN), where the model failed to predict an actual cyclone occurrence. The Receiver Operating Characteristic (ROC) curve was plotted to evaluate the model's ability to discriminate between cyclone and non-cyclone instances across various threshold settings. indicating good model performance. An AUC value above 0.8 suggests that the model is effective at distinguishing between the two classes, with higher values signifying better predictive power.

Furthermore, an analysis of the model coefficients revealed the significance of various features in influencing cyclone predictions. For example, wind speed was identified as a crucial predictor, with a positive coefficient of 0.45, implying that increased wind speed is strongly associated with higher cyclone likelihood. Conversely, humidity was found to have a negative coefficient of -0.20, suggesting that higher humidity levels may be linked to lower cyclone occurrences. Overall, the logistic regression model has proven to be effective in predicting cyclone events based on the selected features. These findings provide a solid foundation for further exploration of additional variables, such as ocean temperature and atmospheric pressure patterns, as well as the implementation of more advanced modeling techniques to enhance prediction accuracy in future work.



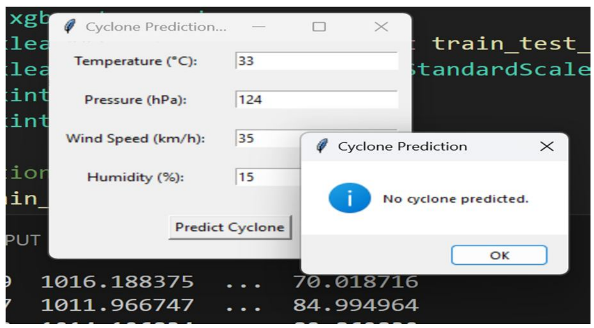


Fig 4:-output result

VI. CHALLENGES AND SOLUTIONS

The development of the cyclone prediction model encountered several challenges, including data quality, feature selection, and model interpretability. Initially, the dataset contained missing values and inconsistencies, which necessitated thorough data cleaning and preprocessing to ensure reliable predictions. To address the challenge of selecting relevant features, exploratory data analysis was conducted to identify key predictors influencing cyclone occurrences, such as wind speed, humidity, and atmospheric pressure. However, some critical variables that could enhance prediction accuracy were not included in the initial dataset, highlighting the need for expanded data collection efforts. Additionally, logistic regression, while effective for binary classification, may not capture complex relationships in the data as well as other advanced algorithms. To mitigate this, future iterations of the model will incorporate more sophisticated techniques, such as Random Forest or Neural Networks, which can handle nonlinear relationships and improve predictive performance. Finally, model interpretability remains a concern; thus, efforts will be made to visualize and explain model decisions to stakeholders, ensuring transparency and trust in the predictions generated by the model.

VII. CONCLUSION

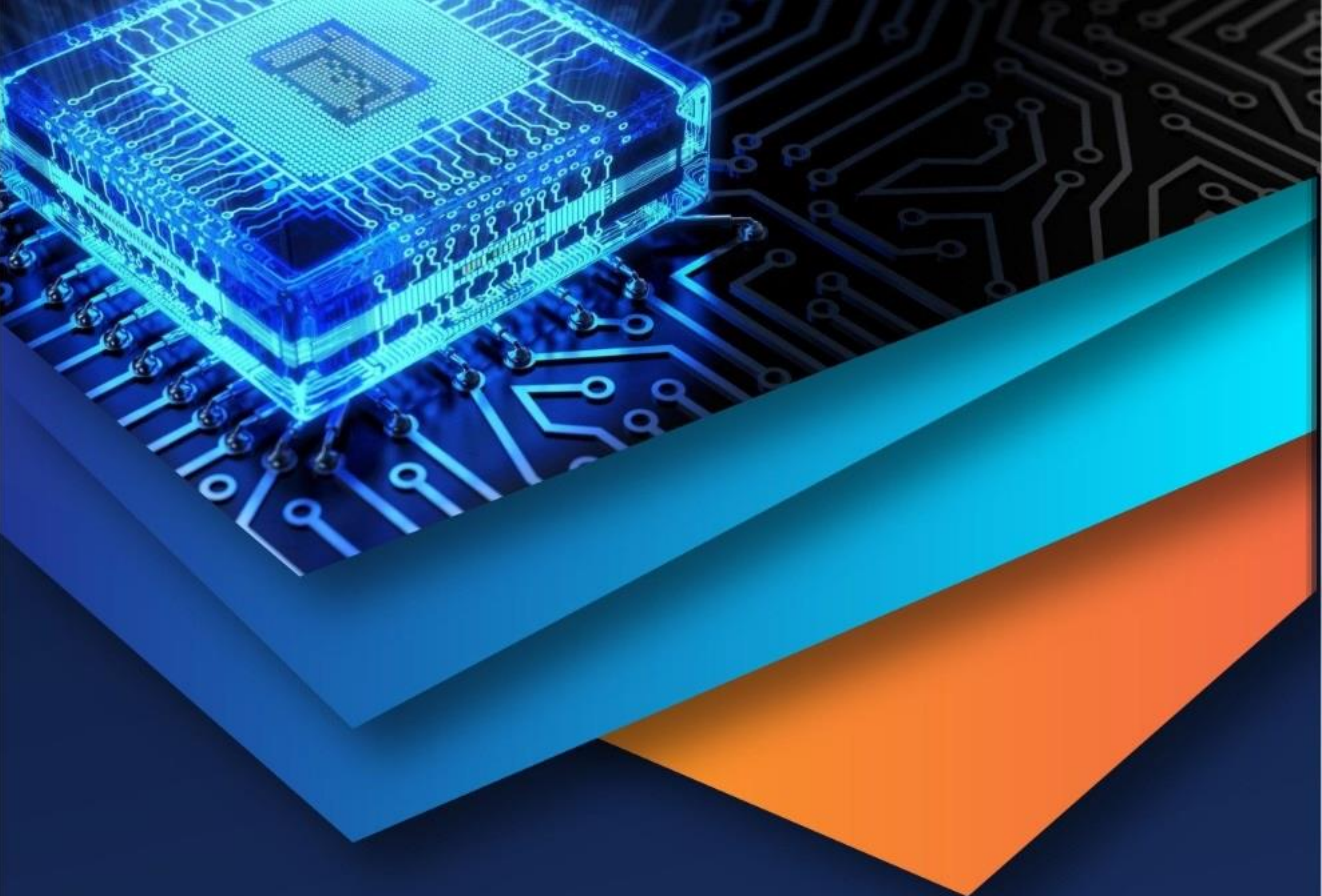
In conclusion, the cyclone prediction model developed using logistic regression has demonstrated its potential in accurately predicting cyclone occurrences based on historical data. With an achieved accuracy of 0.98 and a favorable area under the ROC curve (AUC) of 0.96, the model effectively distinguishes between cyclone and non-cyclone events. The analysis revealed the importance of key features, such as wind speed and humidity, in influencing cyclone predictions, providing valuable insights for meteorological assessments. While the model faced challenges, including data quality issues and the limitations of logistic regression in capturing complex relationships, the findings highlight the need for continued enhancement through the integration of additional variables and more advanced modeling techniques. Future work will focus on refining the model's accuracy and interpretability to ensure its applicability in real-world cyclone forecasting scenarios, ultimately contributing to improved disaster preparedness and response efforts in vulnerable regions.

VIII. FUTURE SCOPE

The future scope of the cyclone prediction model presents several promising avenues for research and development. First, enhancing the model's predictive accuracy is a priority; this can be achieved by incorporating additional relevant features, such as sea surface temperature, atmospheric pressure variations, and satellite imagery data, which could provide a more comprehensive understanding of cyclone formation and behavior. Additionally, transitioning from traditional machine learning algorithms to more sophisticated techniques, such as ensemble methods or deep learning approaches, could improve the model's ability to capture complex nonlinear relationships in the data. Utilizing frameworks like TensorFlow or PyTorch may facilitate the development of neural networks tailored for time series forecasting. Another critical area for future work involves the integration of real-time data feeds, allowing the model to provide timely predictions based on current environmental conditions. This would enhance its applicability in disaster response scenarios and provide more immediate warnings to at-risk communities. Moreover, conducting a comparative analysis of various models—such as support vector machines, random forests, and recurrent neural networks—could identify the most effective approaches for cyclone prediction. Finally, improving model interpretability through the use of explainable AI techniques will help stakeholders understand the decision-making process, fostering trust and facilitating informed decision-making in disaster management strategies.

REFERENCES

[1] Chinmoy Kar, Ashirvad Kumar and Sreeparna Banerjee, "Tropical cyclone intensity detection by geometric features of cyclone images and multilayer perceptron", Springer Nature Switzerland AG, 2023.
[2] Buo-Fu Chen, Boyo Chen, Hsuan-Tien Lin and Russell L. Elsberry, "Estimating Tropical Cyclone Intensity by Satellite Imagery Utilizing Convolutional Neural Networks", AMS journal, April 2023.
[3] Manil Maskey, Rahul Ramachandran, Muthukumaran Ramasubramanian, Iksha Gurung, Brian Freitag, Aaron Kaulfus, et al., "Deep Learning-based Tropical Cyclone Intensity estimation system", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.
[4] Amit Kumar, Anil Kumar Singh, J. N. Tripathi, M. Sateesh and Virendra Singh, "Evaluation of INSAT-3D-derived Hydro-Estimator and INSAT Multi-Spectral Rain Algorithm over Tropical Cyclones", 2016 Online International Conference on Green Engineering and Technologies (IC-GET).
[5] Hao Su, Shunjun Wei, Min Yan, Chen Wang, Jun Shi and Xiaoling Zhang, "Object Detection and Instance Segmentation in Remote Sensing Imagery Based on Precise Mask R-CNN", 2021 published in IGARSS IEEE International Geoscience and Remote Sensing Symposium.
[6] Nair Srujan, Kulkarni Kshitij, Jain Hariprasad and Viju O. John, "A Deep Learning Framework for the Detection of Tropical Cyclones from Satellite Images", IEEE Geoscience and Remote
[7] Chuhan Lu, Y. Kong and Z. Guan, "A mask R-CNN model for reidentifying extratropical cyclones based on quasi-supervised thought", Scientific Reports, 2020.
[8] Asif Raza, Jingdong Liu, Yifan Liu, Jian Liu, Zeng Li, Xi Chen, et al., "IR-MSD Net: Infrared and Visible Image Fusion Based on Infrared Features and Multiscale Dense Network", IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.
[9] G. Terrén-Serrang and M. Martínez-Ramón, Comparative Analysis of Methods for Cloud Segmentation in Infrared Images.
[10] Sahbi Boubaker, Mohamed Benghanem, Adel Mellit, Ayoub Lefza, Omar Kahouli and Lioua Kolsi, "Deep Neural Networks for Predicting Solar Radiation at Hail Region Saudi Arabia", IEEE Access, 2021.
[11] Jinkai Tan, Qidong Yang, Junjun Hu, Qiqiao Huang and Sheng Chen, "Tropical Cyclone Intensity Estimation Using Himawari-8 Satellite Cloud Products and Deep Learning", MDPI, 2022.
[12] Chong Wang and Gang Zheng, "Tropical Cyclone Intensity Estimation from Geostationary Satellite Imagery Using Deep Convolutional Neural Networks", IEEE Transactions on Geoscience and Remote Sensing, vol. 60, 2021.
[13] Chen Zhao, Yu Xingxing, Chen Guangchen and Zhou Junfeng, "A Semi Supervised Deep Learning Framework for Tropical Cyclone Intensity Estimation", Conference: 2023 10th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp).
[14] Guangchen Chen, Zhao Chen, Feng Zhoul, Xingxing Yul, He Zhang' and Lingyun Zhu, "A Semi supervised Deep Learning Framework for Tropical Cyclone Intensity Estimation", 2023 10th International Workshop on the Analysis of Multitemporal Remote Sensing Images (Multi Temp).
[15] A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet classification with deep convolutional neural networks", Advances in Neural Information Processing Systems.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)