Cyclone prediction system using Deep Learning

Project ID: 31210

B.Tech. Project Report
submitted for fulfillment of
the requirements for the
Degree of Bachelor of Technology
Under Biju Pattnaik University of Technology

Submitted By

Subrata Kumar Sahu **ROLL NO. CSE202110040**

Soumyakanta Moharana **ROLL NO. CSE202110043**

Anwesh Misra **ROLL NO. CSE202110068**

A. Sai Satyanarayana **ROLL NO. CSE202140086**

2024 – 2025

Under the guidance of

Dr. Pradeep Kumar Jena

NIST UNIVERSITY
Pallur Hills, Berhampur, Odisha – 761008, Indi

NIST UNIVERSITY

BONAFIDE CERTIFICATE

This is to certify that the B.Tech Project Report entitled **"Cyclone Prediction System using Deep Learning"** is carried out by **Subrata Kumar Sahu (Roll No. CSE202110040) , Soumya Kanta Moharana (Roll No. CSE202110043), Anwesh Misra(Roll No. CSE202110068), Allamsetti Sai Satyanarayana (Roll No. CSE202140086)** under my direct supervision and guidance.

This report is submitted in fulfillment of the requirements for the award of the degree of **Bachelors of Technology** to **NIST University** , for the academic session 2024-25.

Dr. Pradeep Kumar Jena
(Advisor)

Dr. Kunjabihari Swain
(B.Tech. Project Coordinator)

Dr. Brojo Kishore Mishra
(HoD, Dept. of Computer Science & IT)

Dr. Susmita Mahato
(Department Project Coordinator)

Institute Seal

0680 2492-421
0680 2492-422
www.nist.edu
info@nist.edu
Institute Park, Berhampur
Odisha-761008, India

Abstract

Cyclones are highly destructive weather events, posing significant threats to coastal areas due to intense winds, rain, and storm surges. Traditional forecasting models face challenges in accurately predicting cyclone formation, intensity, and trajectory due to limitations in resolution, initialization, and physical process modeling. This project aims to improve cyclone prediction by implementing both Recurrent Neural Networks (RNN), Gated Recurrent Network (GRU) and Long Short-Term Memory (LSTM) networks, which are well-suited for analyzing sequential weather data.

We developed and trained RNN, GRU and LSTM models to predict cyclone severity levels based on historical weather data. The models were tested using different time window sizes to assess their performance in forecasting cyclone intensity. By comparing RNN, GRU and LSTM, this project explores the effectiveness of each model in capturing sequential patterns and enhancing the accuracy of cyclone predictions. This work aims to improve early warning systems, providing more reliable forecasts for better disaster preparedness and response.

ACKNOWLEDGEMENT

We would like to take this opportunity to thank all those individuals whose invaluable contribution in a direct or indirect manner has gone into the making of this project a tremendous learning experience for me.

It is our proud privilege to epitomize my deepest sense of gratitude and indebtedness to my faculty guide, **Dr. Pradeep Kumar Jena** for her valuable guidance, keen and sustained interest, intuitive ideas and persistent endeavor. Her guidance and inspirations enabled me to complete my report work successfully.

We give our sincere thanks to **Dr. Kunjabihari Swain (**B. Tech Project Coordinator) for giving me the opportunity and motivating me to complete the project within stipulated period of time and providing a helping environment.

We acknowledge with immense pleasure the sustained interest, encouraging attitude and constant inspiration rendered by **Prof. (Dr.) Sukant K. Mohapatra (Chairman), Dr. Priyadarsan Patra (Vice Chancellor), Dr. Brojo Kishore Mishra (H.O.D CSE), and Dr. Susmita Mahato (Department Project Coordinator)**. Their continued drive for better quality in everything that happens at N.I.S.T. and selfless inspiration has always helped us to move ahead.

Subrata Kumar Sahu
Roll No: CSE202110040

Soumya Kanta Moharana
Roll No: CSE202110043

Anwesh Misra
Roll No: CSE202110068

A Sai Satyanarayana
Roll No: CSE202140086

Table of Contents

LIST OF FIGURES

1. Introduction

Cyclones are classified as extreme weather events that, alongside gales, rainstorms, and storm surges, can lead to significant damage in coastal regions around the globe. Over the last century, meteorologists and warning centers have focused on understanding these phenomena, achieving advancements in observational technologies, the physics of cyclone intensification, atmospheric interactions, and forecasting methods. Despite these efforts, challenges remain in predictive capabilities, particularly concerning cyclone formation, intensity, and risk assessments. The leading dynamical forecast models often exhibit low accuracy due to several factors: poor initialization of cyclone vortices, incomplete modelling of intricate physical processes, and insufficient resolution.

The difficulty in accurately forecasting cyclones continues to be a pressing issue. The inherent complexity of these systems and the limitations of current forecasting models contribute to frequent inaccuracies in predictions. Improving model resolution and refining the representation of physical processes are essential for enhancing forecast reliability. Furthermore, integrating high-quality observational data and advancing vortex initialization techniques are critical steps that could lead to substantial improvements in forecast accuracy. On going research aims to reduce the devastating effects of cyclones by providing more reliable and timely warnings to affected populations.

To solve these problems machine learning techniques are adapted to improve the prediction of these cyclones more than the traditional methods presented.

Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks are specifically tailored for handling sequential data, making them particularly effective for time-series forecasting tasks such as cyclone prediction. These models are adept at analyzing historical weather data, enabling them to uncover patterns and relationships that traditional forecasting techniques might miss. By leveraging the capabilities of RNNs and LSTMs, researchers can significantly enhance the precision of forecasts related to cyclone formation, intensity, and trajectory.

RNNs are designed to maintain a hidden state that updates as they process each element in a sequence, allowing them to retain important information from prior time steps. This feature is crucial for making predictions based on the overall context of the data. LSTMs build on standard RNNs by incorporating specialized memory cells and gating mechanisms that regulate the flow of information, which helps address challenges like the vanishing gradient problem. This enables LSTMs to effectively capture long-term dependencies within the data.

In addition to improving the accuracy and adaptability of cyclone forecasts, machine learning techniques also offer significant potential for risk assessment. By analyzing patterns in historical data, RNNs and LSTMs can help identify regions and populations that are most vulnerable to cyclone impacts. This information can inform targeted preparedness measures, enabling authorities to allocate resources more effectively and minimize the potential damage caused by these destructive storms. As research in this field continues to advance, the integration of machine learning into cyclone prediction promises to revolutionize our ability to anticipate and mitigate the devastating effects of these natural disasters.

2. Cyclone Prediction

2.1 What is Cyclone prediction?

Cyclone prediction is a critical aspect of meteorology that involves forecasting the formation, movement, and intensity of tropical cyclones, which are known as hurricanes in the Atlantic and eastern Pacific, typhoons in the western Pacific, and simply cyclones in the Indian Ocean and South Pacific. These storms pose significant threats to life and property, making accurate predictions essential for mitigating damage, preparing emergency responses, and saving lives. To make accurate forecasts, meteorologists collect vast amounts of atmospheric and oceanic data from sources such as weather satellites, Doppler radars, oceanic buoys, ships, weather balloons, and specially equipped aircraft that fly into cyclones. Once a cyclone starts forming, it is tracked using satellite imagery, radar data, and historical comparisons. Forecasting heavily relies on numerical weather prediction (NWP) models, such as the Global Forecast System (GFS), the European Centre for Medium-Range Weather Forecasts (ECMWF), the Hurricane Weather Research and Forecasting (HWRF) model, and the Global and Regional Assimilation and Prediction System (GRAPES), which use supercomputers to process vast datasets and generate forecast tracks, intensity estimates, and landfall predictions. Recent advancements in artificial intelligence (AI) and machine learning (ML) have further enhanced cyclone prediction by using neural networks, deep learning algorithms, data assimilation techniques, and predictive analytics to improve forecast accuracy. However, challenges remain, including rapid intensification, uncertainty in track prediction, unpredictable land interactions, and limited oceanic data availability. Despite these difficulties, cyclone prediction continues to improve through technological advancements and data analysis, offering hope for even more precise and timely forecasting in the future.

2.2 Cyclone Prediction Process

Key aspects of cyclone prediction include:

- **Formation and Development**: Predicting whether a tropical disturbance will develop into a full-fledged cyclone and its potential intensity.
- **Track Prediction**: Forecasting the path a cyclone will take, including its direction and speed of movement.

- **Intensity Forecasting**: Predicting the changes in a cyclone's wind speed and pressure over time.

2.3 Cyclone Intensity Forecasting

Cyclone intensity forecasting is crucial for disaster preparedness, early warning systems, and mitigating the devastating effects of tropical cyclones. Accurate forecasting helps governments, disaster management teams, and the general public take necessary precautions. The methods used for cyclone prediction involve a combination of meteorological data collection, computational modeling, and statistical analysis.

Cyclone Prediction Methods

1. **Meteorological Data Collection** Meteorological data is crucial for cyclone prediction, providing real-time and historical atmospheric and oceanic conditions. Key data sources include:

   a. **Satellites**

   Satellites help monitor cyclones from space, providing real-time imagery and meteorological parameters:

   - **Infrared & Visible Imagery** – Detects cloud patterns and cyclone structure.
   - **Microwave Sensors** – Measures precipitation and internal storm dynamics.
   - **Scatterometers** – Estimate surface wind speeds over oceans.

   Popular meteorological satellites for cyclone tracking:

   - **GOES (Geostationary Operational Environmental Satellites)** – Used by NOAA.
   - **INSAT (Indian National Satellite System)** – Used by the India Meteorological Department (IMD).
   - **Himawari-8** – Used by the Japan Meteorological Agency.

   b. **Weather Stations**

   Ground-based weather stations provide localized atmospheric data, including:

   - Temperature and humidity
   - Barometric pressure changes
   - Wind speed and direction
   - Rainfall measurements

   c. **Aircraft Reconnaissance**

Specialized aircraft, such as Hurricane Hunters, gather real-time cyclone data on:

- Wind speeds at different altitudes
- Air pressure within the storm's eye
- Temperature and humidity variations

d. Buoys and Ships

Weather buoys and ships in the ocean collect critical data such as:

- Sea surface temperature (SST)
- Wave heights
- Oceanic conditions affecting cyclone formation and intensification

2. **Computer Models for Cyclone Prediction**

Advanced computational models use mathematical equations to simulate cyclone behaviour based on atmospheric and oceanic interactions.

a. Numerical Weather Prediction (NWP) Models

NWP models use complex algorithms based on fluid dynamics and thermodynamics to predict weather conditions in a three-dimensional grid.

Popular NWP models for cyclone prediction:

- **Global Forecast System (GFS)** – NOAA’s model for medium to long-term forecasts.
- **European Centre for Medium-Range Weather Forecasts (ECMWF)** – Known for high accuracy in tracking cyclone paths.
- **Weather Research and Forecasting (WRF) Model** – Used for regional cyclone simulations.

b. Ensemble Modelling

Ensemble models run multiple simulations with slightly varied initial conditions to reduce prediction uncertainty.

Examples:

- **GEFS (Global Ensemble Forecast System)**
- **ECMWF Ensemble Prediction System (EPS)**

c. Machine Learning and AI-Based Models

AI techniques analyze vast meteorological datasets to recognize cyclone behaviour patterns.

Common AI techniques used:

- **Deep Learning Neural Networks** – Improve intensity predictions based on historical data.
- **Support Vector Machines (SVM)** – Classify cyclone patterns.
- **Recurrent Neural Networks (RNNs)** – Capture sequential cyclone behaviour for long-term forecasting.

Statistical Techniques for Cyclone Prediction

Statistical methods analyze historical cyclone data to identify trends and correlations for forecasting.

a. **Climatology and Persistence Model**

- Assumes future cyclones will behave similarly to past ones with similar initial conditions.
- Useful for short-term forecasts but less effective for long-term predictions.

b. **Regression Models**

Regression techniques analyze past cyclone data to create predictive equations for storm intensity and movement.

Factors considered:

- Sea surface temperature anomalies
- Wind shear variations
- Atmospheric pressure changes

c. **Analog Techniques**

Compares current cyclone conditions with past cyclones that had similar tracks and intensities. Provides insights into potential storm behaviour.

d. **Bayesian Networks**

Uses probability distributions to estimate the likelihood of cyclone intensification based on multiple influencing factors.

2.4 Importance of Cyclone Prediction:

Accurate cyclone prediction is crucial for:

- **Early Warning Systems:** Issuing timely warnings to coastal communities and shipping vessels, allowing for evacuation and preparation.

- **Disaster Management:** Planning and coordinating emergency response efforts, including resource allocation and deployment.
- **Economic Impact:** Mitigating the economic losses caused by cyclones, such as damage to infrastructure and disruption of transportation and trade.
- **Research and Development:** Improving our understanding of cyclone dynamics and refining prediction models for future advancements.

2.5 Challenges in Cyclone Prediction:

- **Complex Atmospheric Processes:**

  Cyclones are influenced by a wide range of atmospheric and oceanic factors, including sea surface temperature, humidity, wind shear, atmospheric pressure variations, and interactions with other weather systems. These factors are highly dynamic and can change rapidly, making precise forecasting difficult. Additionally, cyclones undergo processes such as rapid intensification, eyewall replacement cycles, and interactions with jet streams, which add layers of unpredictability. Even small variations in initial atmospheric conditions can lead to significant differences in the predicted cyclone track and intensity, making accurate long-term forecasts extremely challenging.
- **Data Limitations:** Accurate cyclone prediction heavily relies on real-time observational data collected from satellites, weather buoys, aircraft reconnaissance, and radar systems. However, a significant challenge arises due to limited data availability, especially over remote oceanic regions where cyclones typically originate and intensify. Satellite imagery provides broad coverage but may lack the precision needed for detailed analysis, while aircraft reconnaissance is costly and limited to specific regions. Additionally, gaps in historical cyclone data hinder the ability to train machine learning models effectively, impacting forecast accuracy. The lack of high-resolution data can result in uncertainties regarding cyclone formation, intensity changes, and landfall predictions.
- **Model Uncertainty:** Numerical weather prediction models are fundamental in cyclone forecasting, but they come with inherent uncertainties. These models rely

on complex mathematical equations that simulate atmospheric processes, and their accuracy depends on several factors, including the quality of input data, model resolution, and assumptions made in parameterization. Different models may yield varying predictions for the same cyclone due to differences in how they handle key variables such as convection, ocean-atmosphere interactions, and energy exchanges. Additionally, small errors in initial conditions can amplify over time, leading to significant discrepancies in projected cyclone paths and intensities. As a result, forecasters must often rely on multiple models and ensemble forecasting techniques to reduce uncertainty, yet absolute accuracy remains elusive.

2.6 Advancements in Cyclone Prediction:

Despite the numerous challenges associated with cyclone prediction, significant progress has been made over the years due to advancements in meteorological technology, data collection methods, and forecasting models. These improvements have enhanced the accuracy, reliability, and lead times of cyclone forecasts, allowing authorities to take timely action to mitigate damage and save lives. Some of the key advancements include:

1. Improved Satellite Technology

Satellites play a crucial role in cyclone monitoring by providing real-time data on cloud formations, sea surface temperatures, wind speeds, and atmospheric moisture levels. Modern geostationary and polar-orbiting satellites, such as the GOES (Geostationary Operational Environmental Satellite) and Himawari series, offer high-resolution imagery and continuous monitoring of tropical storm systems. The use of microwave and infrared sensors enables better detection of storm intensification, while advanced algorithms help track the development of cyclonic systems with greater precision.

2. Enhanced Data Collection Methods

Accurate cyclone prediction relies heavily on data from various sources, including weather buoys, drones, aircraft reconnaissance, and automated weather stations. Specialized aircraft, such as the NOAA Hurricane Hunters, fly directly into cyclones to gather critical data on wind speed, pressure, and humidity. Unmanned aerial systems and underwater drones are also being deployed to collect oceanic and atmospheric data, providing a more comprehensive understanding of storm dynamics.

The expansion of global observation networks has significantly improved the availability of high-resolution data for forecasting.

3. Advanced Numerical Weather Prediction (NWP) Models

The development of sophisticated numerical weather prediction models has greatly enhanced cyclone forecasting capabilities. High-resolution models, such as the Global Forecast System (GFS), the European Centre for Medium-Range Weather Forecasts (ECMWF), and the Hurricane Weather Research and Forecasting (HWRF) model, incorporate vast amounts of observational data and use complex mathematical equations to simulate atmospheric behavior. These models have improved track forecasting, reducing errors in landfall predictions and enabling better preparedness. Additionally, ensemble forecasting techniques, which run multiple simulations with slight variations in initial conditions, help assess uncertainties and improve forecast confidence.

4. Machine Learning and Artificial Intelligence (AI) Applications

Recent advancements in machine learning and artificial intelligence are revolutionizing cyclone prediction. AI-based models can analyze historical cyclone data and identify patterns that traditional methods may overlook. Deep learning algorithms are being used to improve intensity predictions and track forecasting by processing large datasets from satellites, radars, and climate models. AI-driven models are also assisting in the early detection of cyclone formation, increasing the lead time for warnings.

5. High-Performance Computing (HPC) and Big Data Analytics

The integration of high-performance computing and big data analytics has enabled meteorologists to run complex simulations with higher speed and accuracy. Supercomputers can process vast amounts of weather data in real time, allowing for rapid updates to forecasts and better risk assessments. The ability to analyze multiple variables simultaneously has contributed to more precise cyclone intensity and impact predictions.

6. Improved Early Warning Systems and Communication Strategies

Advancements in communication technologies have enhanced the dissemination of cyclone warnings to the public and emergency response agencies. Automated alert systems, mobile applications, and satellite-based communication networks ensure that cyclone warnings reach vulnerable communities in a timely manner. Additionally, better coordination between meteorological agencies, government bodies, and disaster management organizations has strengthened preparedness and response efforts, minimizing casualties and property damage.

7. Research and Development in Climate Science

Ongoing research in climate science is further improving the understanding of cyclone behavior and long-term trends. Studies on the influence of climate change on cyclone frequency, intensity, and patterns help refine prediction models and develop more effective mitigation strategies. The continuous collaboration between meteorologists, climate scientists, and data scientists is paving the way for even more accurate and reliable cyclone forecasts in the future.

3. Methodology

Methodology for Cyclone Prediction Using Time Series Data

The methodology for cyclone prediction using time series data involves multiple stages, including dataset preparation, data pre-processing, splitting the data, model training, performance evaluation, and result analysis. These steps are crucial in developing an accurate machine learning model capable of predicting cyclone patterns using historical weather data.

a. **Dataset Preparation**

The dataset consists of **7,961 rows** and **12 columns**, providing detailed information about cyclonic systems. Each entry includes a **Serial Number**, which serves as a unique identifier for a system within a given year. The **Basin of Origin** column indicates the oceanic region where the system originated. The **Name** column, which contains the cyclone's name (if assigned), has missing values since not all cyclones are named. The **Date** column is recorded in the **DD-MM-YYYY** format but is stored as an object, which may require conversion to a proper date-time format for analysis. The **Time (UTC)** column records the time in **Coordinated Universal Time (UTC)**. Cyclone intensity is represented by the **CI No (or T. No)** column. The **Estimated Central Pressure (hPa) [E.C.P]** indicates the system's pressure, while the **Maximum Sustained Surface Wind (kt)** records the highest wind speed in knots. The dataset also classifies cyclones using the **Grade** column, which includes categories like D, DD, and CS. Another key variable is the **Outermost Closed Isobar (hPa)**, representing the pressure of the outermost closed isobar. The **Year** column specifies the event year. Notably, the dataset contains a duplicate column for **Estimated Central Pressure (hPa) [E.C.P]**, suggesting data redundancy. These characteristics highlight potential areas for data cleaning, particularly in handling missing values, converting date formats, and addressing duplicate columns.

b. **Pre-Processing Data**

Raw weather data collected from the API often contains inconsistencies and missing values. Pre-processing is necessary to clean and format the data for effective model training. The key steps include:

- **Handling Missing Values:** Filling gaps using interpolation techniques or removing incomplete records.

- **Removing Outliers:** Detecting and eliminating extreme values that may affect model accuracy.
- **Normalization:** Scaling the data to a uniform range (e.g., 0 to 1) to ensure that all features contribute equally to model learning.
- **Formatting Data for Model Training:** Converting time series data into structured input suitable for deep learning models like LSTM, RNN, and GRU.

Proper pre-processing ensures that the dataset is optimized for accurate cyclone prediction.

c. Splitting Data

Once the data is pre-processed, it is divided into two subsets:

- **Training Data:** Used for training the machine learning models. Typically, 70-80% of the dataset is allocated for training.
- **Testing Data:** Used to evaluate model performance. The remaining 20-30% of the data is set aside for testing.

Splitting the dataset prevents overfitting and ensures that the models generalize well to unseen data.

d. Model Training

To analyze time series weather data effectively, three deep learning models are trained:

I. Long Short-Term Memory (LSTM) Model

LSTM is a specialized type of Recurrent Neural Network (RNN) designed for sequential data. It is highly effective for cyclone prediction because it can:

- Capture long-term dependencies in time series data.
- Avoid the problem of vanishing gradients, which affects standard RNNs.
- Retain information over extended periods, making it ideal for forecasting weather conditions.

II. Recurrent Neural Network (RNN) Model

RNNs are widely used for processing sequential data. Unlike traditional neural networks, RNNs have memory units that store previous time steps' information. While effective for short-term dependencies, RNNs struggle with long-term dependencies due to vanishing gradients, making LSTM a better alternative for cyclone prediction.

III. Gated Recurrent Unit (GRU) Model

GRU is another variation of RNN, similar to LSTM but with a simplified architecture. It uses fewer parameters than LSTM, making it computationally efficient while still handling sequential data effectively. GRUs are particularly useful for:

- Reducing training time while maintaining high accuracy.
- Handling sequential dependencies efficiently with their gating mechanism.

4. Hyperparameter Tuning:

To optimize model performance, hyperparameters such as learning rate, batch size, number of layers, and activation functions are fine-tuned. This ensures the best possible cyclone prediction accuracy.

a. Performance Evaluation

Once the models are trained, they are evaluated on the testing dataset using various performance metrics:

- **Accuracy:** Measures how often the model makes correct predictions.
- **Precision:** Evaluates the proportion of true positive predictions out of all predicted positives.
- **Recall (Sensitivity):** Measures the model's ability to detect actual cyclones.
- **F1-Score:** Balances precision and recall, providing a comprehensive performance metric.

Comparing the performance of **LSTM, RNN, and GRU** helps determine which model is best suited for cyclone prediction.

b. Results and Insights

The final step involves analyzing and interpreting the model outputs:

- **Model Comparison:** Evaluating which model performs best based on accuracy, recall, and F1-score.
- **Error Analysis:** Identifying potential reasons for incorrect predictions and refining the models accordingly.

- **Decision-Making Capability:** Assessing whether the model can make informed cyclone predictions for real-world applications.

The insights gained from this study can help improve **early warning systems**, enhance **disaster management**, and contribute to **research on cyclone dynamics**.

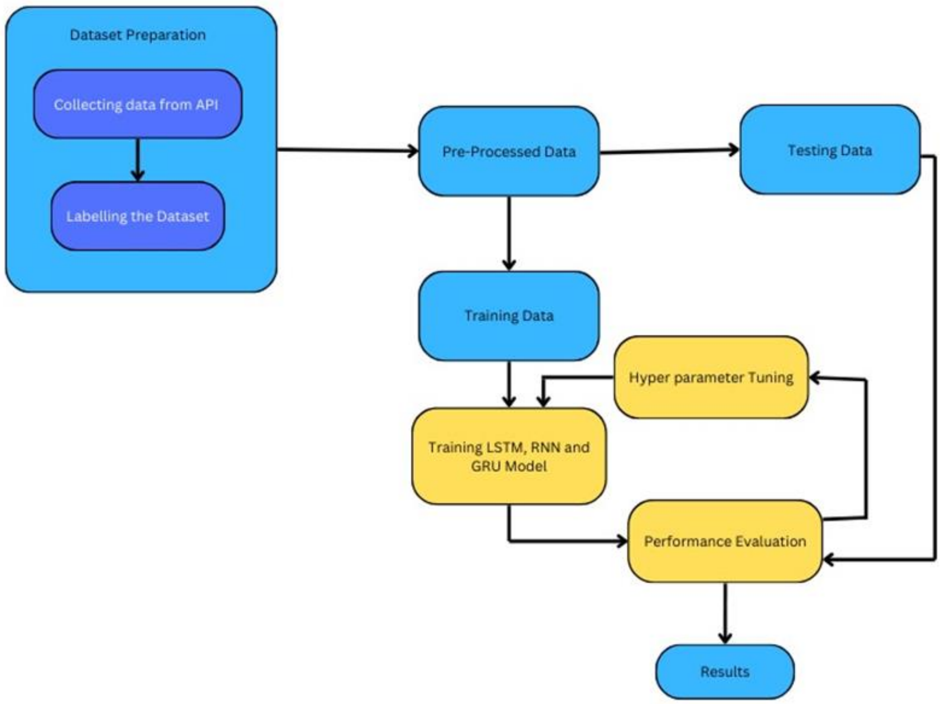

Figure 1: Flowchart Diagram

4. Data Collection

Dataset Description: RSMC Best Track Dataset (1982-2024)

The dataset contains 7,961 records detailing tropical cyclone activity in the North Indian Ocean, including the Bay of Bengal (BOB) and the Arabian Sea (ARB), from 1982 to 2024.

Key Features:

- **Basin of Origin:**

  This tells you where the cyclone first formed:

  i) Bay of Bengal (BOB): The majority of cyclones in this dataset originated here, which is typical for the North Indian Ocean. This basin is known for favourable conditions like warm waters and low wind shear.

  ii) Arabian Sea (ARB): While less frequent than in the Bay of Bengal, cyclones in the Arabian Sea can be powerful and have become more common in recent years, potentially due to climate change.
- **Date & Time (UTC)**: Provides timestamped cyclone observations.
- **CI No (T. No)**: CI No. (T No.) represents cyclone intensity using the Dvorak Technique, ranging from T1.0 (weak depression) to T8.0 (super cyclone). It estimates wind speed and strength based on satellite imagery. Higher values indicate stronger storms, aiding in cyclone classification, forecasting, and disaster preparedness
- **Estimated Central Pressure (ECP)**: Estimated Central Pressure (ECP) is the approximate atmospheric pressure at the center of a cyclone, measured in hPa (hectopascals). Lower ECP values indicate stronger cyclones, as intense storms create a deep low-pressure area. It helps estimate cyclone intensity, wind speeds, and potential impact.
- **Maximum Sustained Wind Speed:** This is another crucial measure of cyclone intensity. Higher wind speeds indicate a more powerful storm. The dataset captures a wide range of wind speeds, from weak systems (2 kt) to very intense cyclones (140 kt).
- **Grade (text):** This provides a categorical classification of cyclones based on their intensity. The exact categories used might vary depending on the specific agency or scale (e.g., the India Meteorological Department's scale).

Common categories include:

- Depression (D)
- Deep Depression (DD)
- Cyclonic Storm (CS)
- Severe Cyclonic Storm (SCS)
- Very Severe Cyclonic Storm (VSCS)
- Extremely Severe Cyclonic Storm (ESCS)
- Super Cyclonic Storm

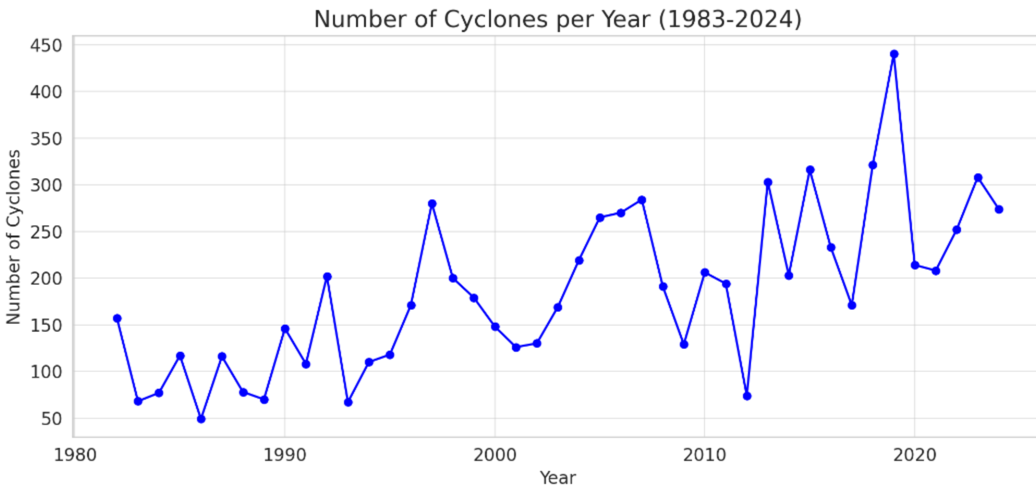

Figure 2: Yearly Cyclone Frequency

Yearly Cyclone Frequency (1983-2024):

- Significant variability in cyclone counts across years
- Lowest annual record count: Around 50-60 records
- Highest annual record count: Peaks at around 440-450 records (likely in recent years)
- Apparent increasing trend in record count from 1980s to 2020s
- Notable fluctuations with multi-year cycles of higher and lower record counts

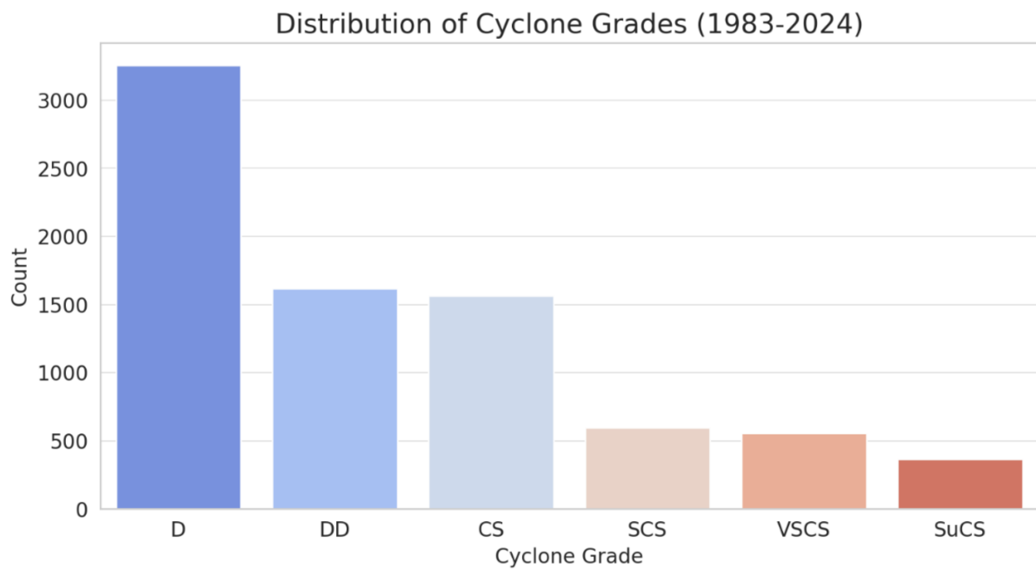

Figure 3: Cyclone Grade Distribution

Cyclone Grade Distribution:

- D (Depression) grade is most frequent, with approximately 3,300-3,400 records
- DD (Deep Depression) follows with around 1,600-1,700 records
- CS (Cyclonic Storm) has similar count to DD, around 1,500-1,600 records
- Severe categories (SCS, VSCS, SuCS) have progressively fewer records
- Demonstrates a typical pyramidal distribution of cyclone intensities

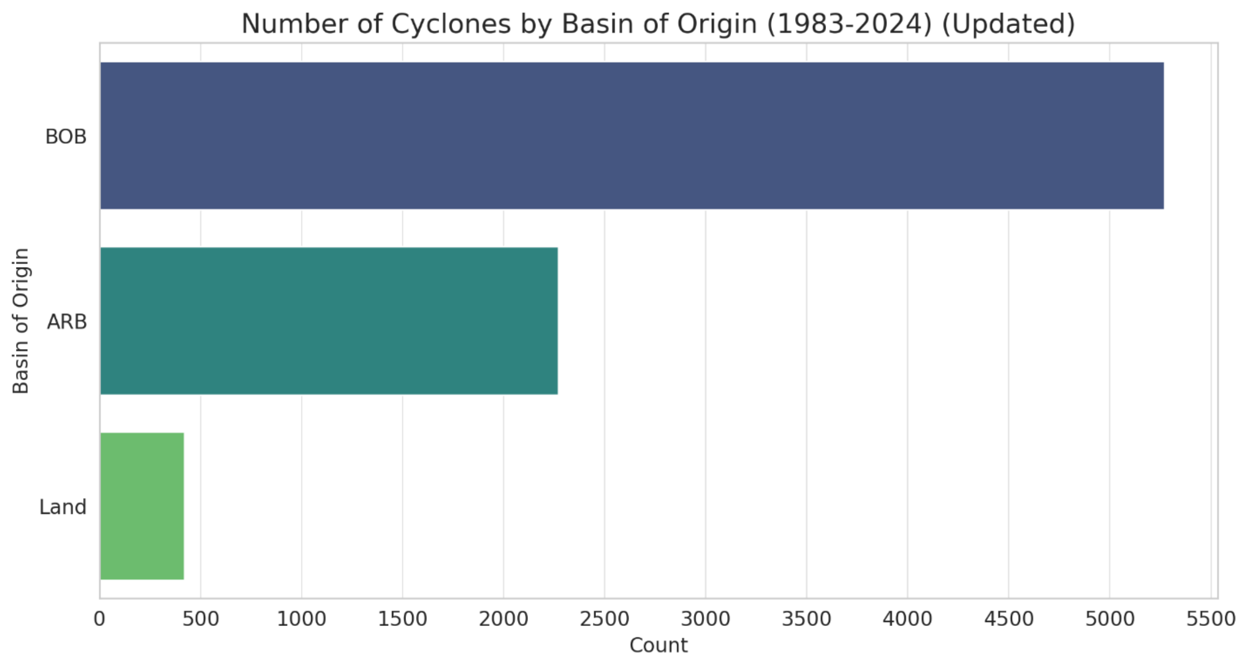

Figure 4: Basin of Origin Analysis

Basin of Origin Analysis:

- BOB (Bay of Bengal) dominates with approximately 4,500-5,000 cyclone records
- ARB (Arabian Sea) shows around 2,500-3,000 cyclone records
- Land-based cyclones are minimal, suggesting maritime origin predominance

Some of the significant cyclones covered in the dataset are:

- **Tauktae (May 14, 2021 – May 19, 2021):** An extremely severe cyclonic storm that impacted India's western coast, primarily Gujarat.
- **Yaas (May 23, 2021 – May 28, 2021):** A very severe cyclonic storm that made landfall in Odisha, India, shortly after Cyclone Tauktae. Dorian (August 24, 2019 – September 10, 2019), severely impacting the Bahamas and Southeastern United States.
- **Phailin (October 4, 2013 – October 14, 2013)**, hitting Eastern India, especially Odisha and Andhra Pradesh.
- **Amphan (May 13, 2020 – May 21, 2020):** An extremely severe cyclonic storm that made landfall in West Bengal, India, and Bangladesh
- **Hudhud (October 6, 2014 – October 14, 2014)**, impacting Eastern India, particularly Andhra Pradesh.
- **Fani (April 26, 2019 – May 4, 2019)**, striking Odisha, India, and parts of Bangladesh.

5. Data Preprocessing

Preparing the dataset for model training is a crucial phase that significantly impacts the accuracy and efficiency of the predictive model. A well-processed dataset ensures that the model can learn effectively and generalize well to new data. The preprocessing phase involves multiple key steps, each designed to enhance data quality and usability. The steps include consistency checks, outlier removal, handling missing values, and normalization.

1. **Consistency Checks**

   Consistency checks involve examining the dataset for uniformity across different sources and formats. This step is fundamental to ensure that all data points are comparable and reliable, preventing inconsistencies that may affect model training. The following aspects are considered during consistency checks:

   - **Format Standardization**: Ensuring that all data is in a uniform format, such as date and time representation, categorical data encoding, and consistent units of measurement.
   - **Data Type Verification**: Checking that numerical data is correctly formatted as integers or floats, and categorical data is properly labeled.
   - **Duplication Removal**: Identifying and eliminating duplicate records that may skew the model's learning process.
   - **Cross-Source Validation**: If data is collected from multiple sources, cross-referencing the values helps detect discrepancies and ensure integrity.

2. **Outlier Removal**

   Outliers are data points that significantly deviate from the expected range and can distort the model's predictions. These anomalies may arise due to sensor errors, recording mistakes, or genuine but rare events. The process of outlier removal involves:

   - **Statistical Analysis**: Methods like Z-score analysis, interquartile range (IQR), and boxplots help detect outliers.
   - **Domain Knowledge Application**: Understanding the dataset's domain allows differentiation between valid extreme values and erroneous data points.

- **Removal or Adjustment**: Based on the nature of the outlier, it may be removed or adjusted using statistical techniques such as mean or median substitution.

3. **Handling Missing Values**

Missing values in a dataset can lead to biased results and reduced model accuracy. Depending on the extent and nature of the missing data, different strategies are applied:

- **Deletion Methods**:
- **Listwise Deletion**: Removing records with missing values if the proportion is small and does not significantly reduce the dataset size.
- **Pairwise Deletion**: Retaining records where missing values do not affect critical variables.
- **Imputation Techniques**:
- **Mean/Median/Mode Imputation**: Filling missing values with the mean, median, or mode of the respective column.
- **K-Nearest Neighbors (KNN) Imputation**: Using the values of similar records to estimate missing values.
- **Regression Imputation**: Predicting missing values based on relationships between variables.
- **Interpolation**: Estimating missing data points in a time-series dataset using linear or polynomial interpolation.

4. **Normalization**

Normalization is essential to standardize the range of input variables, ensuring that no single feature dominates the model's learning process. This step improves model convergence, stability, and accuracy. Common normalization techniques include*:*

- **Min-Max Scaling**: Rescales values between 0 and 1 using the formula: This technique is suitable when the data distribution is not normal and has fixed boundaries.
- **Z-score Standardization**: Transforms data into a standard normal distribution with mean 0 and standard deviation 1 using the formula: This method is preferred when the dataset follows a Gaussian distribution.

- **Decimal Scaling**: Scales values by moving the decimal point based on the largest absolute value.

Each of these normalization methods is chosen based on the nature of the dataset and the requirements of the machine learning model.

6. Model Selection

6.1 RNN:

A recurrent neural network is a neural network that is specialized for processing a sequence of data x(t)= x(1), . . . , x(τ) with the time step index *t* ranging from 1 to τ. For tasks that involve sequential inputs, such as speech and language, it is often better to use RNNs. In an NLP problem, if you want to predict the next word in a sentence it is important to know the words before it. RNNs are called *recurrent* because they perform the same task for every element of a sequence, with the output being depended on the previous computations. Another way to think about RNNs is that they have a "memory" which captures information about what has been calculated so far.

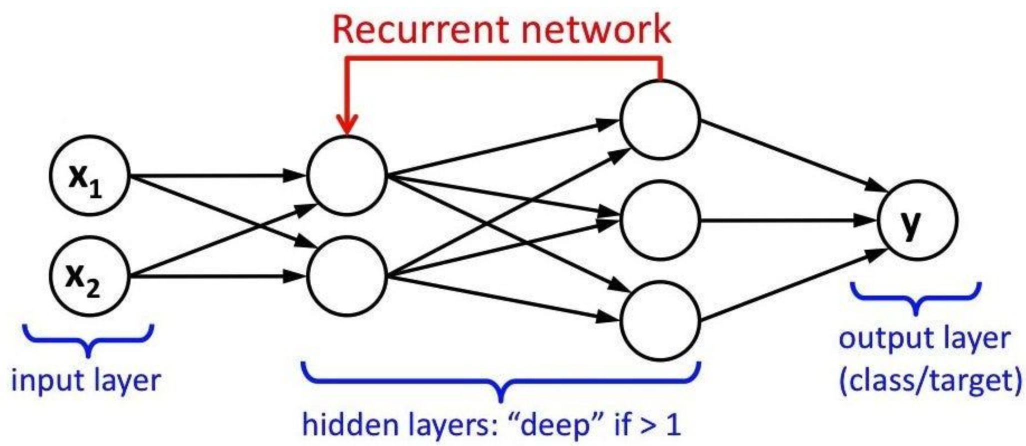

Figure 6.1: RNN Structure

- **Input:** *x(t)* is taken as the input to the network at time step *t*. For example, *x1*, could be a one-hot vector corresponding to a word of a sentence.
- **Hidden state*:*** *h(t)* represents a hidden state at time t and acts as "memory" of the network. *h(t)* is calculated based on the current input and the previous time step's hidden state:

  $$h(t) = f\,(U\,x(t) + W\,h(t-1)).$$

  The function *f* is taken to be a non-linear transformation such as *tanh*, *ReLu.*
- **Weights:** The RNN has input to hidden connections parameterized by a weight matrix U, hidden-to-hidden recurrent connections parameterized by a weight matrix W, and hidden-

to-output connections parameterized by a weight matrix V and all these weights (*U*, *V*, *W)* are shared across time.

- **Output:** *o(t)* illustrates the output of the network. In the figure I just put an arrow after *o(t)* which is also often subjected to non-linearity, especially when the network contains further layers downstream.

6.2 LSTM:

Long Short-Term Memory (LSTM) networks are a specialized type of Recurrent Neural Network (RNN) designed to effectively manage long-term dependencies in sequential data. They were introduced to address the limitations of traditional RNNs, particularly the vanishing gradient problem, which hampers the learning process when dealing with long sequences. LSTMs achieve this through a unique architecture that incorporates memory cells and gating mechanisms, allowing them to retain and manipulate information over extended periods.

The fundamental building block of an LSTM is the memory cell, which serves as a long-term storage unit. Each memory cell is equipped with three types of gates:

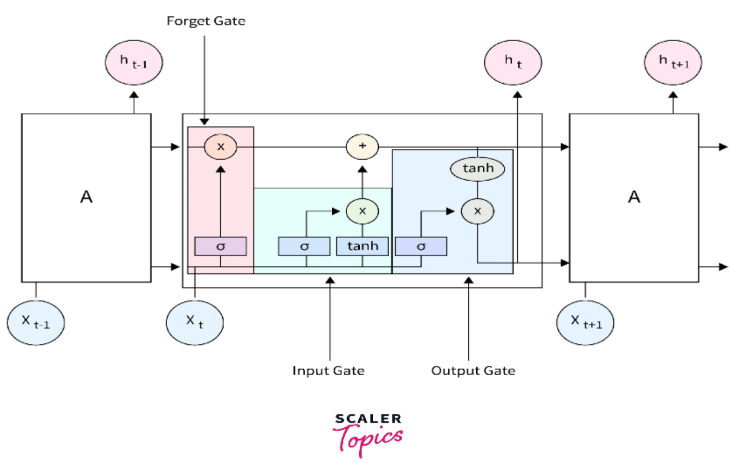

Figure 6.2: LSTM Structure

- **Forget Gate**: This gate determines what information from the previous time step should be discarded. It outputs a value between 0 and 1 for each element in the cell state, where 1 indicates “keep this information” and 0 means “forget this information”.
- **Input Gate**: This gate controls what new information is added to the cell state. It also outputs values between 0 and 1, allowing the model to selectively update its memory based on current inputs.

- **Output Gate**: This gate decides what information from the cell state will be outputted to the next layer or as part of the model's prediction. It filters the cell state based on both the current input and the hidden state from the previous time step.

The forward training process of the LSTM can be formulated with the following equations:

$$ft = \sigma(Wf \cdot [ht{-}1, xt]{+}bf)$$

$$it = \sigma(Wf \cdot [ht{-}1, xt]{+}bi)$$

$$Ct = ft{*}Ct{-}1{+}it{*}tanh(Wc \cdot [ht{-}1, xt]{+}bc)$$

$$Ot = \sigma(W0 \cdot [ht{-}1, xt]{+}b0)$$

$$ht = ot{*}tabh(Ct)$$

where it, ot, and ft denote the activation of the input gate, output gate, and forget gate, respectively; Ct and ht denote the activation vector for each cell and memory block, respectively; and W and b denote the weight matrix and bias vector, respectively. In addition, $\sigma(°)$ denotes the sigmoid function.

6.3 GRU:

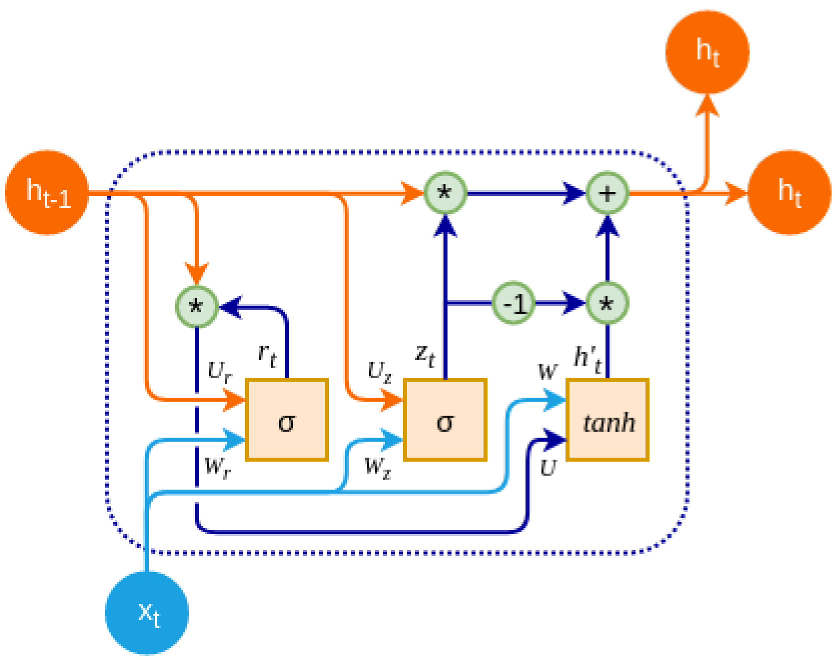

Figure 6.3: GRU Structure

Gated Recurrent Units (GRUs) are a specialized type of Recurrent Neural Network (RNN) designed to effectively model sequential data while addressing the vanishing gradient problem seen in traditional RNNs. GRUs offer a simpler alternative to Long Short-Term Memory (LSTM) networks, reducing computational complexity while maintaining strong performance in sequence modeling tasks. The key advantage of GRUs lies in their ability to selectively retain or discard information, enabling efficient handling of long-term dependencies in data.

The fundamental building block of a GRU is the hidden state, which acts as the memory mechanism. Instead of maintaining a separate memory cell like LSTMs, GRUs rely on two primary gating mechanisms to regulate information flow:

- **Reset Gate**: This gate determines how much of the previous hidden state should be forgotten. It outputs a value between 0 and 1 for each element in the hidden state, where 0 represents complete forgetting and 1 represents full retention.
- **Update Gate**: This gate controls how much of the new candidate activation vector should be incorporated into the hidden state. It helps balance between keeping past information and incorporating new information from the current input.

The forward training process of a GRU follows these equations:

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r)$$

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z)$$

$$\tilde{h}_t = \tanh(W_h \cdot [r_t * h_{t-1}, x_t] + b_h)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

Where:

- x_t is the input at time step t.
- h_t is the updated hidden state.
- W_r, W_z, W_h are weight matrices, and b_r, b_z, b_h are bias vectors.
- $\sigma(\cdot)$ is the sigmoid activation function

Overall, GRUs provide a computationally efficient alternative to LSTMs, making them particularly suitable for scenarios where limited resources or faster training times are required. Their ability to capture long-range dependencies while maintaining a simpler architecture makes them a popular choice for sequential data modeling tasks such as natural language processing and time series forecasting.

7. Analysis & Results

7.1 LSTM (LONG-SHORT TERM MEMORY)

The **Sequential LSTM model** is designed to classify cyclone severity by leveraging **Long Short-Term Memory (LSTM) networks**, a type of recurrent neural network (RNN) well-suited for handling time-series data. The model architecture is structured to progressively **reduce dimensionality**, extract meaningful features, and improve classification accuracy while preventing overfitting. Below is a detailed breakdown of the model's architecture and its effectiveness in cyclone severity classification.

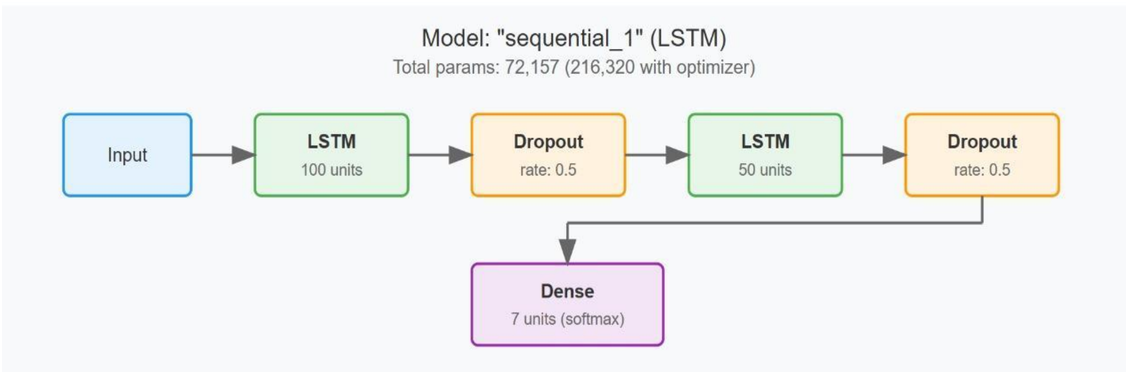

Figure 7.1 LSTM Model

1. **Layer-by-Layer Architecture and Dimensionality Reduction**

The model processes cyclone data through **three LSTM layers**, each refining the feature representation:

i. **First LSTM Layer (100 features)**
   - Captures initial temporal dependencies from raw weather data.
   - Extracts high-dimensional patterns related to cyclone formation and progression

ii. **Second LSTM Layer (50 features)**
   - Further refines feature representations, ensuring the model focuses on critical aspects of cyclone intensity.
   - Provides a compact yet informative feature set for the classification layers.

7.2 SIMPLE RNN

Analysis of the Simple RNN Model for Cyclone Severity Classification

The Simple RNN model is designed to classify cyclone severity by analyzing sequential weather data. It consists of **two RNN layers** for capturing temporal dependencies, followed by **two Dense layers** for refining features and making final predictions. The input layer processes sequences of length **20**, with **8 features per time step**, while the output layer classifies the data into **five severity categories**.

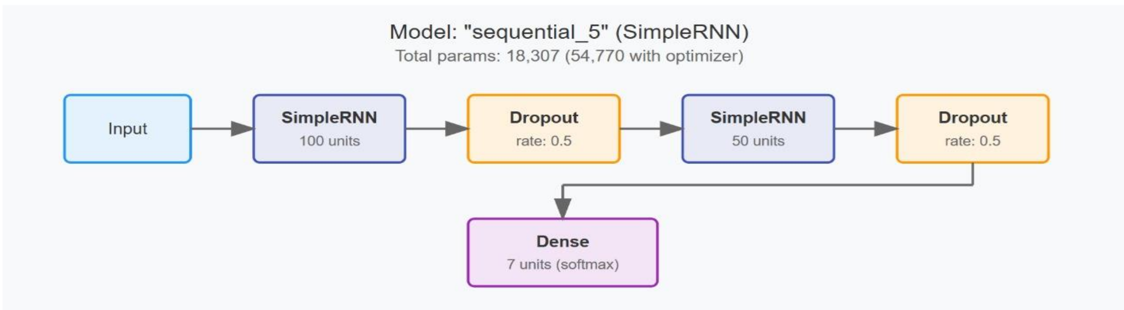

Figure 7.2 RNN Model

Model Architecture & Layer-Wise Breakdown

i. **First Simple RNN Layer:**
   - Captures short-term dependencies in cyclone data.
   - Preserves *sequential relationships between time steps.*

ii. **Second Simple RNN Layer:**
   - Further refines the temporal patterns extracted from the first layer.
   - Enhances learning by processing longer-range dependencies.

iii. **Dense Layer:**
   - Extracts meaningful features from the RNN output.
   - Helps in dimensionality reduction while retaining essential information.

iv. **Final Output Layer:**
   - Uses an activation function (e.g., SoftMax) to classify the cyclone severity into five categories.
   - Outputs probability scores for each class, enabling precise classification.

7.3 Gated Recurrent Unit

Gated Recurrent Units (GRUs) are a variant of recurrent neural networks (RNNs) designed to handle sequential data efficiently. Their simplified architecture, using only reset and update gates, enables faster computations and reduced memory requirements compared to Long Short-Term Memory (LSTM) networks. In the context of cyclone classification and forecasting, the study highlights both the strengths and limitations of GRUs in handling different time window sizes.

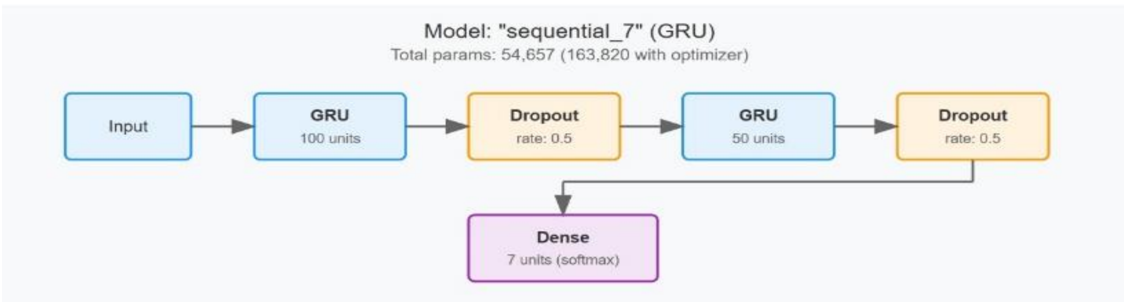

Figure 7.3 GRU Model

Performance of GRU in Cyclone Forecasting:

Efficiency in Shorter Sequences

i. GRUs are well-suited for processing sequential data where short- to mid-range dependencies are critical. In cyclone prediction, this means they can effectively capture immediate weather fluctuations and evolving patterns within a limited time window.

ii. Their ability to retain relevant information without excessive computational overhead makes them a practical choice for near-real-time forecasting applications.

7.4 Analysis:

1. Accuracy Curves:

This graph presents a comparative analysis of the accuracy progression of three deep learning models—RNN, GRU, and LSTM—over multiple training epochs. The X-axis represents the number of training epochs, while the Y-axis represents the accuracy achieved by each model at different stages of training.

Recurrent neural networks (RNNs) and their advanced variants, GRU (Gated Recurrent Unit) and LSTM (Long Short-Term Memory), are widely used for sequence-based tasks such as time-series forecasting and natural language processing. The purpose of this graph is to illustrate how each model learns and improves accuracy as training progresses.

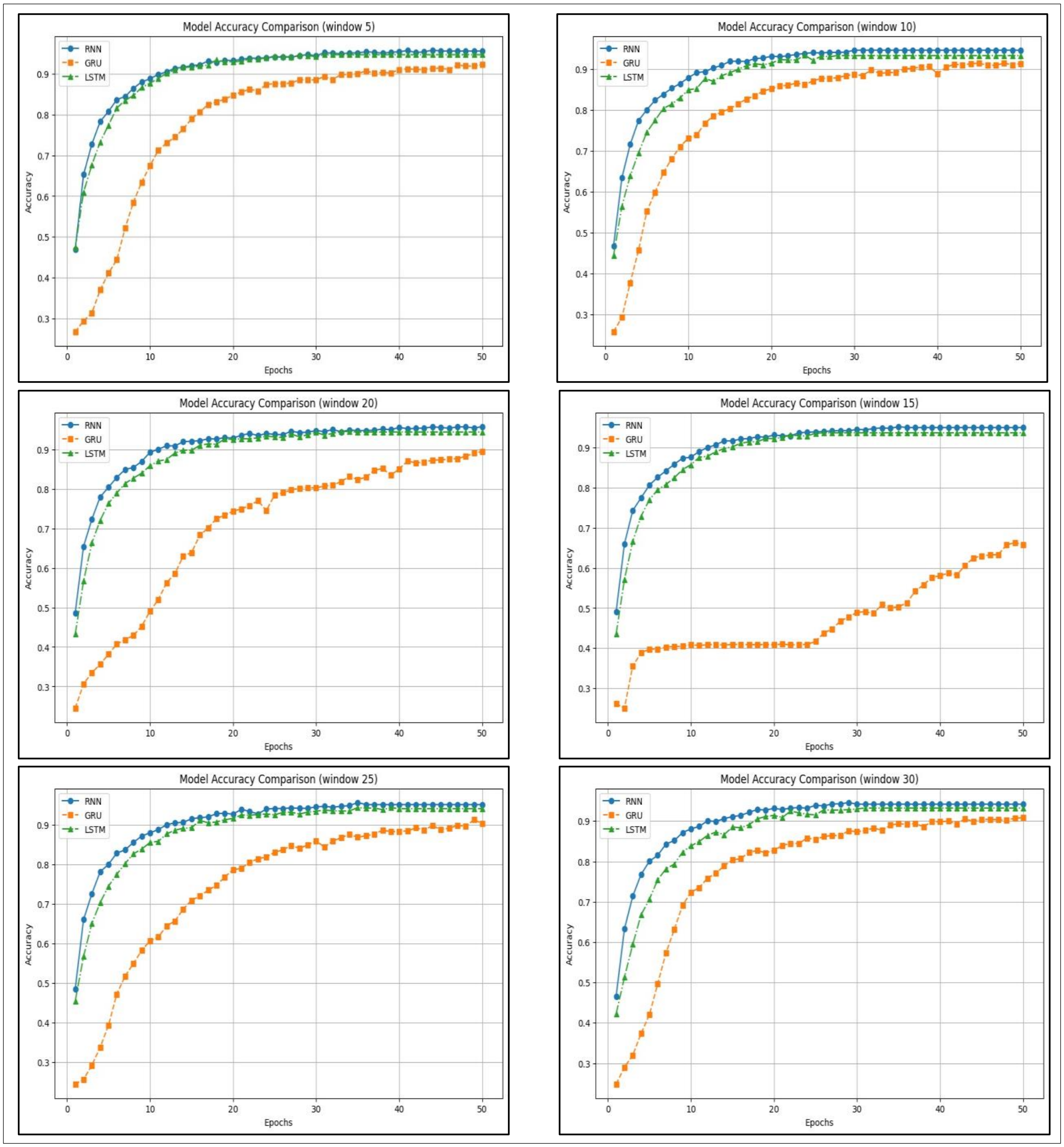

Figure 7.4 Comparison of Accuracy between LSTM, Simple RNN and GRU

LSTM, RNN, and GRU were compared across different window sizes for forecasting accuracy. RNN performs best for short-term forecasting (window size ≤ 15), excelling at capturing immediate dependencies. LSTM outperforms for long-term forecasting (window size ≥ 20), effectively handling long-range temporal patterns. GRU shows inconsistent accuracy across all window sizes, with a significant drop at window size 15. Therefore, it is better to use RNN for short-term predictions, LSTM for long-term forecasting, and avoid GRU due to instability.

LSTM's memory gating makes it ideal for long sequences, while RNN is more responsive to short-term changes.

2. ROC Curve

A **Receiver Operating Characteristic (ROC) Curve** is a graphical representation used to evaluate the performance of a **classification model**. It plots **True Positive Rate (TPR)** against **False Positive Rate (FPR)** at different threshold values.

Key Components of an ROC Curve:

- **X-axis (False Positive Rate, FPR):** Measures how many negative samples were incorrectly classified as positive.
- **Y-axis (True Positive Rate, TPR):** Measures how many actual positives were correctly classified.
- **Diagonal Line (Random Guess Line):** Represents a model with no predictive power (50-50 random guessing).
- **AUC (Area Under the Curve):** The higher the **AUC value (closer to 1.0)**, the better the model's performance.

The following curves illustrates the classification performance of the three models: LSTM, RNN and GRU respectively:

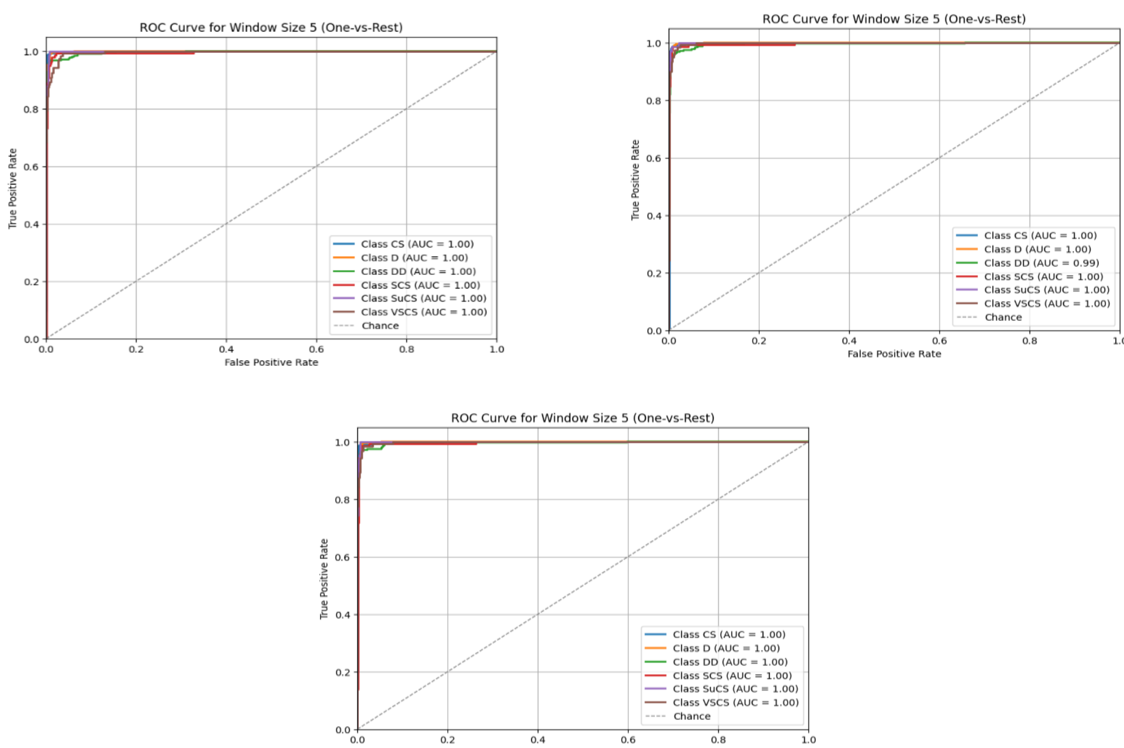

Figure 7.6 ROC Curves of LSTM, Simple RNN and GRU

The ROC curves reveal exceptional cyclone classification performance across LSTM, RNN, and GRU architectures, with near-perfect AUC scores (1.00) for most classes. The models demonstrate remarkable consistency and discriminative power, suggesting robust feature extraction and minimal misclassification. These results indicate that the underlying data representation is likely more crucial than the specific neural network architecture for this cyclone classification task.

3. Precision-Recall Curve

A precision-recall curve visualizes a classification model's performance by plotting precision against recall at various threshold settings. It illustrates the trade-off between correctly identifying positive instances (precision) and capturing all actual positive cases (recall). This curve is particularly valuable for imbalanced datasets, helping evaluate model effectiveness by showing how well a classifier maintains accuracy across different classification thresholds.

The following curves illustrates the precision-recall performance of the three models: LSTM, RNN and GRU respectively:

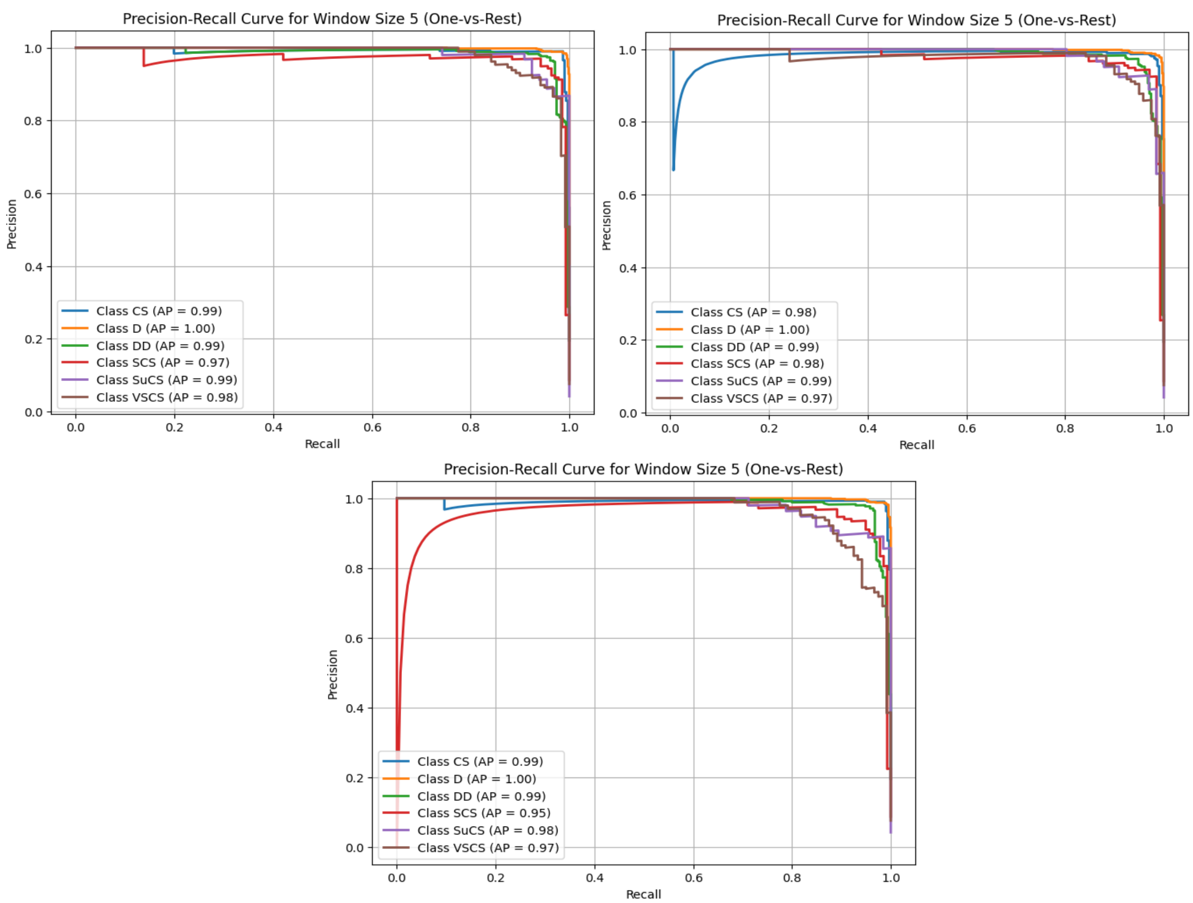

Figure 7.7 Precision-Recall Curves of LSTM, Simple RNN and GRU

The precision-recall curves reveal exceptional cyclone classification performance. GRU shows AP scores of CS (0.99), D (1.00), DD (0.99), SCS (0.97), SuCS (0.99), VSCS (0.98). RNN displays similar metrics: CS (0.98), D (1.00), DD (0.99), SCS (0.95), SuCS (0.98), VSCS (0.97). LSTM demonstrates comparable precision with CS (0.98), D (1.00), DD (0.99), SCS (0.98), SuCS (0.99), VSCS (0.97), indicating highly consistent and robust classification across neural architectures.

Table-7.1 State-of-the-Art

Reference	Method	Model Accuracy
Mohamad S. , Rodi A. and Mohammad A. et.al.[2023]	SVM	0.93
Mohamad S. , Rodi A. and Mohammad A. et.al.[2023]	ANN	0.97
J. S. Kumar, V. Venkataraman, S. Meganathan and K. Krithivasan et.al. [2017]	LSTM	0.95
Yao Qin, Dongjin Song, Haifeng Chen et.al. [2017]	DA-RNN	0.92
Proposed	LSTM	0.96
Proposed	Simple-RNN	0.94
Proposed	GRU	0.92

8. Conclusion

This project demonstrates the potential of machine learning, specifically Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) and Gated Recurrent Networks (GRU), to significantly enhance cyclone prediction accuracy. By analyzing historical weather data, these models can better identify patterns in cyclone formation, intensity, and trajectory, overcoming many limitations of traditional forecasting methods. Accurate and timely predictions allow authorities to issue more reliable warnings and implement targeted preparedness measures, helping to minimize the devastating impacts of cyclones on vulnerable communities.

8.1 Future Aspects:

- **Comparative Analysis:** Further research can compare LSTM, RNN, and GRU models across different datasets to evaluate accuracy, stability, and efficiency.
- **Automation Enhancements**: Implementing automated hyperparameter tuning and real-time data processing can improve forecasting efficiency.
- **Multimodal Data Integration:** Incorporating satellite imagery, oceanic sensors, and atmospheric data can enhance prediction accuracy.
- **Hybrid Model Development:** Combining LSTM, GRU, and RNN can leverage their strengths for improved short- and long-term forecasting.
- **Extreme Weather Adaptability:** Extending models to predict hurricanes, typhoons, and tornadoes can improve disaster preparedness.

In conclusion, this project has laid the groundwork for an advanced cyclone prediction system using neural networks. The architecture developed shows promise in handling complex temporal data, and the proposed future aspects indicate exciting directions for further research and development in this field.

9. Reference

[1] Manas Mondal, Anupam Biswas, Subrata Haldar, (2022) Spatio-temporal behaviours of tropical cyclones over the Bay of Bengal Basin in last five decades

[2] Peng Hao, Yaqi Zhao, Shuang Li (2024), Deep learning approaches in predicting tropical cyclone tracks: An analysis focused on the Northwest Pacific Region

[3] Hill, T., O'Connor, M., & Remus, W. (1996). Neural network models for time series

[4] Guo, T., Lin, T., & Antulov-Fantulin, N. (2019) Exploring Interpretable LSTM Neural Networks over Multi-Variable Data

[5] Etienne, B. (2019). Time Series in Python — Exponential Smoothing and ARIMA processes.

[6] Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation.

[7] Li, Y., Wu, H., & Liu, H. (2018). Multi-step wind speed forecasting using EWT decomposition. LSTM principal computing. RELM subordinate computing and IEWT reconstruction. Energy Conversion and Management.

[8] Qin, Y., Song, D., Chen, H., Cheng, W., Jiang, G., & Cottrel, G. W. (2017). A Dual Stage Attention-Based Recurrent Neural Network for Time Series Prediction.

[9] Mohamad S. , Rodi A. and Mohammad A. [2023]. Classification of Weather Conditions Based on Supervised Learning for Swedish Cities.

[10] J. S. Kumar, V. Venkataraman, S. Meganathan and K. Krithivasan [2017]. Tropical Cyclone Intensity and Track Prediction in the Bay of Bengal Using LSTM-CSO Method.

[11] Yao Qin, Dongjin Song, Haifeng Chen (2017) A Dual Stage Attention-Based Recurrent Neural Network for Time Series Prediction"

[12] Ribana Roscher, Bastian Bohn, Marco F. Duarte, and Jochen Garcke [2019]. Explainable Machine Learning for Scientific Insights and Discoveries.

[13] Subrata Paul, Vidhi Singrodia, and Anirban Mitra, [2024]. Understanding the requirements of the Indian IT industry using web scrapping.

[14] Buo-FuChen, Yu-TeKuo, Treng-ShiHuang [2022]. A deep learning ensemble approach for predicting tropical cyclone rapid intensification.

[15] Shuxian Liu, Yang Liu, Zhigang Chu, [2023] Evaluation of Tropical Cyclone Disaster Loss Using Machine Learning Algorithms with an eXplainable Artificial Intelligence Approach