FEATURE EXTRACTION AND TRACKING SYSTEM FOR TROPICAL CYCLONES

THU ZAR HSAN

UNIVERSITY OF COMPUTER STUDIES, YANGON

JUNE, 2024

Feature Extraction and Tracking System for Tropical Cyclones

Thu Zar Hsan

University of Computer Studies, Yangon

A thesis submitted to the University of Computer Studies, Yangon in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

June, 2024

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research and has not been submitted for a higher degree to any other University or Institution.

28.6.2024

…..…… ……..……

Date Thu Zar Hsan

ACKNOWLEDGEMENTS

First of all, I would like to thank His Excellency, the Minister, the Ministry of Science and Technology for full facilities support during the Ph. D Course at the University of Computer Studies, Yangon.

I would like to express very special thanks to Dr. Mie Mie Khin, Rector, the University of Computer Studies, Yangon, for allowing me to develop this thesis and giving me general guidance during period of my study.

I would like to say a very big thanks to my supervisor, Dr. Thin Lai Lai Thein, University of Computer Studies, Yangon, for her excellent guidance, caring, patience, and providing me with excellent ideas for doing research. I have been extremely lucky to have a supervisor who cared so much about my work. Without her guidance and constant feedback, this dissertation would not have been achievable. I will always remember her for being a mentor to me.

I would also like to extend my special appreciation and thanks to Dr. Ah Nge Htwe, Professor, the University of Computer Studies, Yangon and Dean of the Ph.D Courses for the useful comments, advices and insight which are invaluable to me.

I am also very grateful to express my deepest gratitude to my teacher, Dr. Thandar Win, Pro-rector of the University of Computer Studies, Myeik, for her supporting and encouraging to do my research.

I would like to express my respectful gratitude to Daw Mya Thandar, Associate Professor, and English Department for her valuable supports from the language point of view and pointed out the correct usage in my dissertation.

I am very much indebted to my mother for always believing in me, for her endless love and support. She is always supporting and encouraging me during the years of my Ph.D study. And finally, I want to thank my husband and my sons who gave me their love, kindness, patience, physically and mentally supports, and constant encouragement along the way of my life.

ABSTRACT

There are numerous sub-continents in the world where cyclones yearly hit a certain region. Since cyclones directly affect people's lives and homes, their prediction is crucial to preventing the loss of life and property. There are many ways of techniques that is able to forecast tropical storms systems such as Dvorak technique, different kinds of time series analysis, Artificial Neural Network (ANN), numerical weather prediction system (NWP) model, machine learning, etc. Machine learning theory such as regression analysis is still challenging for forecasting tropical cyclone's track. It is very useful and suitable for predicting and great impact on independence and random data for time series.

Tropical cyclones that occurred in the Northern Indian Ocean affected Myanmar Land. Historical datasets are obtained from Joint Typhoon Warning Center (JTWC) and provided from 1945 to 2022 years. Feature extraction has a critical role in machine learning theory and also strong features impact the outcome of the cyclone trajectory. The movement of the cyclone trajectory points out the value of Latitude and Longitude. In this research, these values are changing in the direction and magnitude of the movement. The main contribution is stand on the correlation coefficient value of the direction and magnitude of the historical trajectory data and test data. Not only the latitude and longitude of the cyclone but also metrological data such as wind speed and sea level pressure are also used the input data to extract the features. Features of Direction and Movement are extracted to build the model based on similar cyclones and tested one. Logistic regression method is used to forecast the latitude and longitude of a cyclone's location 24 hours ahead of time by using the last twelve hours of observations (two positions, at six hourly intervals, and the current position).

The threshold value is also an essential decision-maker or forecaster of the system. According to the value, the accuracy of the system can also change. Three threshold values of the sigmoid function are tested which is based on two similar and three similar cyclones are tested. For the evaluation of the system, three matrixes are selected such as mean absolute percentage error (MAPE), mean absolute error (MAE), and root mean squared error (RMSE). By adding a maximum wind speed and minimum sea level pressure from the historical dataset, performance evaluation is gradually improved for these regression methods.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

CHAPTER 1

INTRODUCTION

Myanmar suffers the risk of natural disasters and the cost of damage is enormous and very affected to coastline regions such as Rakhine State and Ayeyarwady division. Heavily rainfall, extreme temperature, and floating and cyclone surges can hurt the suffering nations and result in significant financial losses. The Bay of Bengal, situated in the northeast of the Indian Ocean, is subjected to the formation of some of the most powerful and damaging Tropical Cyclones (TC) in the world [3]. The Bay of Bengal's coast is shared among Bangladesh, Myanmar, Sri Lanka, and the western part of Thailand. Our country was affected annually by these cyclones. The direction of tropical storms that occur in the Bay of Bengal sometimes goes to Bangladesh, India, or Myanmar. So, predicting the trajectory of tropical cyclones plays a major task in mitigation and prevention of disasters for our country.

Many processing methods have been used to predict weather forecasting in the past year but it remains the challenge to forecast the storm track. Numerical image processing techniques are fundamental tools for the initial preparation of input satellite pictures where image processing is applied to these images. Dvorak technique was a famous technique for intensity estimation science in 1972. Thousands of lives in regions that were affected by cyclones were saved by the application that used Dvorak technique.

Today world, artificial intelligence has increased to present itself rapidly and generally. Machine learning is also one of the sectors of artificial intelligence and it is very important in many areas [1]. In the field of meteorology and climatology, it can be used to classify the region of tropical cyclones achieved from multi-dated satellite images. The rapid development of satellite observations and numerical modeling capabilities has led to improvements in the strength and track of the cyclone forecast methodologies. The location of the cyclone is an essential point of the prediction model. Many researchers presented artificial intelligence and morphological image processing algorithms to extract the cyclone center automatically [4]. These data can be obtained in two forms, the statical data (latitude and longitude) and the formation of the cyclone extracted from satellite images. This technology employs a machine learning algorithm to recognize cloud patterns from satellite imagery. The exact location of the cyclone

can be obtained easily from past statistical data. The past statistical data of tropical cyclones is enormous and many types of learning techniques are still solving this system. Specifically, the development of ensemble approaches and data assimilation employing a range of observational data has led to continual improvements in cyclone track prediction. There is still a lack of accuracy in cyclone track forecasting, particularly when it considers the metrological data such as wind speed, sea level pressure, wind intensity, and eye diameter as well as the area covered [8]. This is despite the rapid advancement of observational technologies and numerical modeling. Machine learning systems that can take nonlinear and complex data have only been slightly studied for tropical cyclone trajectories.

However, several parameters, including the thermodynamics and kinetics of the tropical storm structure and the surrounding meteorological conditions, influence the development of a tropical cyclone. The greatest sustained wind speed, the lowest sea level pressure, the wind intensity, and the eye diameter are only a few of the different factors that can affect a cyclone's path. Forecasting tropical cyclone trajectories is a huge difficulty due to the interplay of these elements. Researching new tropical cyclone track forecasting techniques is essential because of tropical storms have enormous effects on people and the challenges involved in predicting them.

At the moment, there are two main kinds of forecasting techniques for TCs: numerical weather prediction (NWP) is the dominant method. When the initial situation and atmospheric boundary situation are known, NWP computes the equivalent result of partial differential equations, including atmospheric state variables. A statistical model, which often makes use of multiple regression, is the alternative forecasting technique. The relationship between the TC's mobility and its unique historical features forms the basis of the statistical model [11]. Nowadays, numerous deep learning models are used to forecast cyclone tracks; however, they have certain limitations, such as high computational cost, overfitting, lack of interpretability, and dependency on data accuracy. When utilizing deep learning to solve a problem, several limits must be considered.

The major goal of this system is to create a new tracking model for the TC track using the multidimensional logistic regression method. The proposed model uses only a simple multiple logistic regression method, which uses the location of the past 18 hours in terms of latitude and longitude to Magnitude and Direction by using a

mathematical equation, maximum sustained wind speed, and minimum sea level pressure for new cyclone track and two most correlated cyclones that happened in the same basin (Northern Indian Ocean) to extract features for the model. This process can reduce time complexity and improve the prediction accuracy of the next timestamp for 6-hour short-term prediction. Features of direction and magnitude are passing through the multiple logistic regression method with sigmoid function, predicted results are not labeled as binary form as “0” or “1”, “yes” or “no”, and “true” or “false” which is assumed as the probability value and changing as the direction and magnitude of the next point. This research shows less powerful and time-consuming prediction model can be developed by using mathematical equations with a logistic regression model changing the nature of the input variable. Finally, this proposed system is developed in a real-world dataset of cyclones from the Joint Typhoon Warning Center (JTWC), which supports the storms’ geographical factor at 6-hour intervals and this proposed system carries out better than some existing traditional methods, statical method, and deep learning technique for short term prediction.

1.1 Problem Statement

The trajectory of tropical cyclones plays an essential role in understanding the areas it can affect Myanmar Land. The damaging power of TCs is growing in reaction to global warming [2]. The two primary statistical forecasting models and numerical prediction models used in traditional methods of tropical cyclone track prediction are [3]. For numerical models to handle complex thermodynamic formulas and mimic a tropical cyclone's interior structure, they need a strong computational capacity [4]. Although the numerical model has gained widespread use with the advancement of computer technology and the installation of ground stations, it still suffers from issues related to low forecast accuracy and relatively high computational complexity. In the case of cyclone track forecasting during 24, 48, and 72 hours, for instance, errors of about 97.4, 188.2, and 302.7 km have been recorded when using a numerical model; in contrast, these mistakes were only about 84.2, 145.6, and 205.4 km when utilizing a subjective empirical approach, according to data released by the Shanghai Typhoon Institute [5]. More improvement in prediction accuracy is possible using statistical models that identify tropical cyclones based on features from historical data [6]. Tropical storm path prediction is becoming a big data challenge due to the increase in

data volume that comes with the installation of ocean observation stations, terrestrial stations, and meteorological satellites.

Recently, machine learning algorithms have been used in object detection, natural language processing, and image processing, and they have shown a strong ability to analyze vast amounts of complex data in two different ways. In this system, processing time for a significant amount of historical data can be decreased by the correlation coefficient on one side. This technique uses the direction and magnitude of historical data instead of the latitude and longitude of the cyclone center. In contrast, short-term forecasting data can be processed simply and effectively with multiple logistic regression techniques. The precision of the method is further increased by using additional meteorological data. According to the above problem statement, forecasting tropical cyclone tracking systems is still necessary to prevent natural disasters.

1.2 Motivation of the Research

One of the motivations for doing this research is to develop a computer program that supports tropical cyclone mitigation and incorporates the plan before the TC strikes to lessen destruction and injury from the storm. Moreover, the area of tropical cyclone strikes from Bay of Bangel effects Bangladesh, Myanmar, Sri Lanka , and the western part of Thailand every year. Although many channels in other countries announce weather forecasting, Myanmar has only the Department of Meteorology and Hydrology. For this reason, developing a prediction system for tropical cyclone tracks is needed to construct.

Furthermore, there are many other techniques used in the prediction system, regression model was not proposed in the system. Regression model is also the prediction model that constructs the relationship between the variables. This model can construct not only linear interconnection but also nonlinear variables.

The last reason is to propose an automatic tracking system by using the latitude, longitude, wind speed, and sea level pressure as statical input data from the Joint Typhoon Warning Center (JTWC). Therefore, this research on feature extraction and forecasting the track of tropical cyclone motion system for Myanmar Land.

1.3 The Objectives of the Research

The main objective of this system is to propose feature extraction and prediction of tropical cyclone tracing systems using historical data. Nowadays, artificial intelligence approach and machine learning approaches are usually used in the prediction TC system. However, much historical data is needed to train these approaches. Features are extracted from the historical data to create a database that can be used directly without making the training step. Therefore, a newly created database has been tested by using multiple logistic regression techniques with three threshold values to compare the accuracy of the system. The objectives of this proposed system area are as follows:

1. To reduce the injuries and damage of natural disasters to Myanmar Land.
2. To develop an automatic tracking system for short-term forecasting.
3. To decrease the time consumption, a correlation coefficient value is used in this system.
4. To create new features that are extracted from tested cyclones and similar cyclones.
5. To forecast the direction of cyclones accurately by using a simple multiple logistic regression technique.

1.4 Contributions of the Research

Many tropical cyclone tracking systems were used by various machine learning algorithms. The reality of the machine learning system, and the training and testing data are important to learning the model. Many kinds of features are used in the machine learning system. An automated tracking system is comfortable for the user who doesn't know the weather prediction system. Some systems are semi-automated and some systems need experts to use the application. This system is an automated tracking system and accuracy is the most important factor for the research.

This system is an automatic feature extraction and tracking system for tropical cyclones. The first contribution is to decrease time consumption for using huge amounts of data, the correlation coefficient technique is used between tested data and historical data. There is no time to compare all historical data with testing data, only the two most similar cyclones need to be emphasized to extract the feature.

Moreover, getting a high accuracy hit-and-miss ratio is important for this forecasting system. The second contribution is to create the data that seven features of direction and eight features of magnitude with sea level pressure and wind speed are extracted from the tested cyclone and similar cyclones' statical data.

There are also many tracking systems using machine learning techniques, but a regression model has not been used for these systems before. Therefore, the multiple logistic regression model is used for the forecasting of the tropical cyclone trajectory and getting high accuracy is the third contribution of this research.

1.5 Organization of the Research

The structure of this study is with six chapters, including a general introduction, literature review, theory background of machine learning systems, detection and tracking, theory of correlation coefficient, multiple regression model, and experimental results.

Chapter 1 includes the study areas, the motivations, the research issues, and the goals and aims of the study. The contributions of the research work are also presented.

Chapter 2 focuses the literature review on cyclone formation, cloud tracking, and related works concerning existing tropical cyclone tracking systems.

Chapter 3 describes the background theory of Machine learning theory and mentions possible approaches to the solution of cyclone tracking.

Chapter 4 presents the implementation and the basic concepts for the correlation coefficient and multiple regression model.

Chapter 5 discusses the experimental results of the proposed system.

Chapter 6 concludes the thesis and outlines avenues for future research.

CHAPTER 2

LITERATURE REVIEW

This chapter displays an introduction to tropical cyclone formation, machine learning systems and tropical cyclone tracking systems. In addition, the current different approaches to the prediction of tropical cyclone tracking systems are reviewed.

2.1 Tropical Cyclone

Low atmospheric pressure, high winds, and heavy rain are characterized of tropical cyclones and it develops over warm tropical oceans. A tropical cyclone makes extreme winds because it gets energy from the sea surface. Tropical cyclones generally generate between 6 to 30 degrees of Latitude. The surface water must be about 80° F. The regions of Bangladesh, North Indian Ocean, Gulf Coast of North America and the Southern Pacific were struck by cyclones in every late summer month. Tropical cyclone is sometimes called typhoon or hurricane depends on the regions they are formed. Hurricanes is called over North-Eastern, North Atlantic and Caribbean Ocean. Tropical Cyclone or Cyclones are called South Pacific and Indian Ocean. In Northwest Pacific is known as Typhoons. It is a small size but can generate high-speed wind [13].

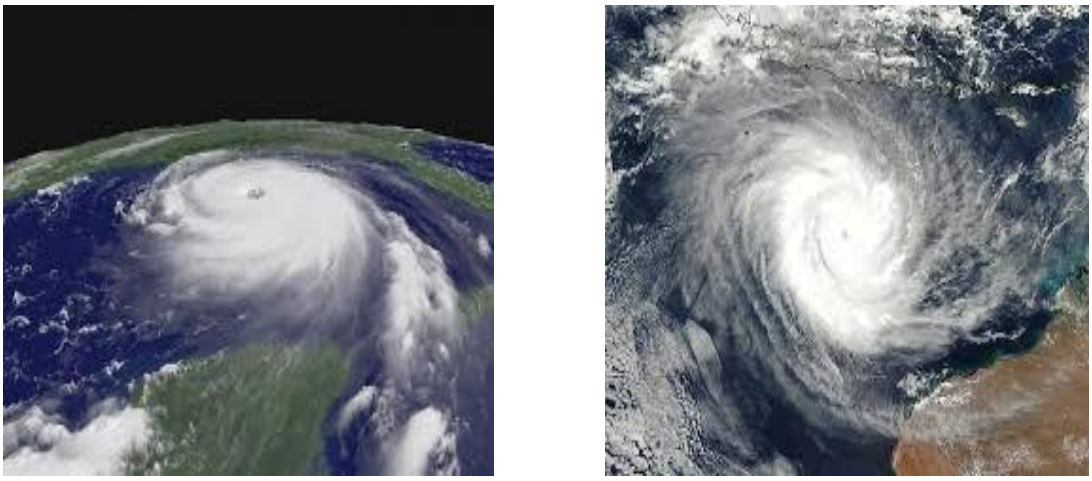


Figure 2.1 Sample Images of Cyclone

2.1.1 Categories of Cyclone

The harshness of a tropical cyclone is expressed in category 1 to category 5 and connected to the maximum mean wind speed demonstrates in the below table.

Table 2.1 Categories of Cyclone

Name	Category	Strength of Gust (km/h)	Effects
Tropical Low TC	Tropical Depression	>63	Gales
Moderate TC	1	90-125	Damaging Wind
Serve TC	2	125-164	Destructive Wind
TC	3	165-224	Very Destructive Wind
Intense TC	4	25-279	Very Destructive Wind
Very Intense TC	5	>279	Extremely Destructive Wind

2.1.2 Impacts of Tropical Cyclones

Tropical cyclones can bring out destructive winds, heavy rainfall that causes flooding and storm surges that can cause overflow to low-lying coastal areas [26].

1. **Wind**

A consequence resulting from tropical cyclones is wind with gusts in above 70 km/h. Sometimes, very intense tropical cyclones can overtake 280 km/h. These destructive winds can destroy buildings and roofing loss. There will be an interruption in the wind when the cyclone's eye crosses over a location, but the destructive wind will come from another direction.

2. **Rain**

Heavy rainfall is connected with a tropical cyclone that can cause immense flooding. This can affect intensive loss and damage. The heavy rain can continue as long as the cyclone goes to the interior and degrades into a low atmosphere air pressure system. Therefore, intense tropical cyclones cause heavy rain and flooding.

3. **Storm Surge**

As well as extreme winds, a tropical cyclone can affect not only the extreme winds but also can raise the sea to become the highest tide. Strong and onshore winds

can produce storm surges and decreased atmospheric pressure. Mainly, the storm surge is a harmful and dangerous natural disaster related to a tropical cyclone [17].

2.2 Tropical Cyclone Pattern

The cloud features of the tropical cyclone are divided into two sets; one set is used to estimate its current intensity and the other its future intensity. A cloud system center may be cloud-free and located within the curve of a cloud line on one day and be obscured under a central dense overcast (CDO) on the next day [46]. The characteristics of CDO are illustrated in Figure 2.3 and examples of each pattern are shown in Figure 2.4. The more spiral the cloud pattern is the more intense is the cyclone.

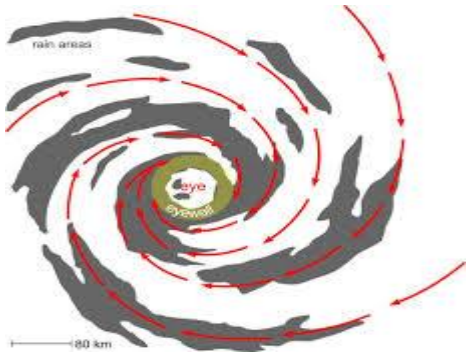


Figure 2.2 Spiral Pattern of Cyclone

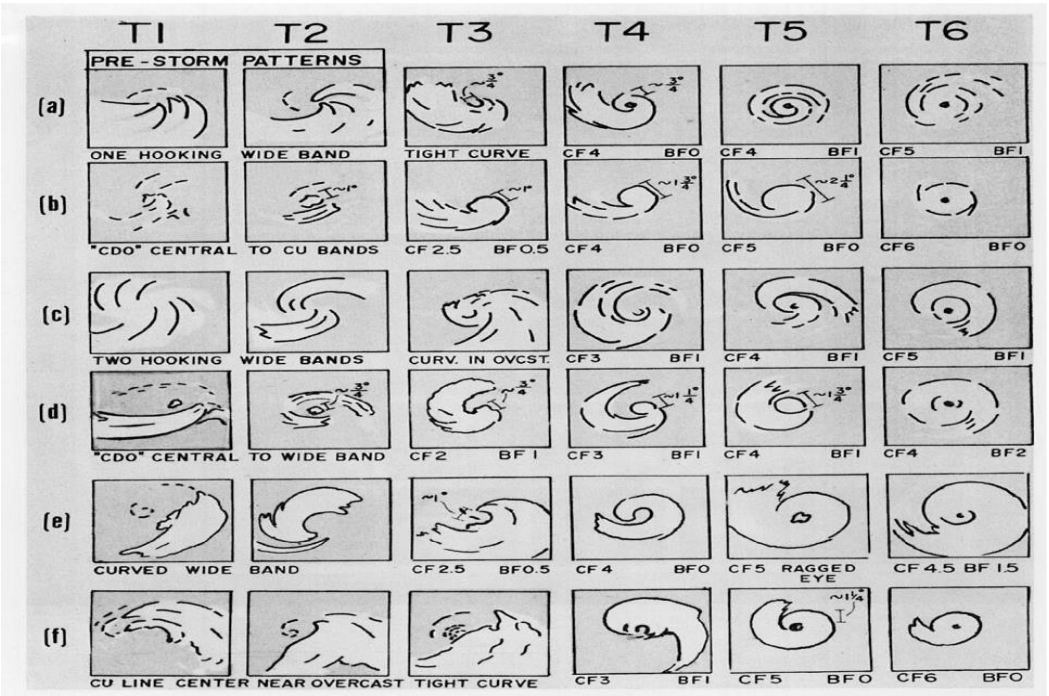


Figure 2.3 Common Tropical Cyclone Patterns and their Corresponding T Numbers

2.3 Cyclone Cloud Classification Techniques

Location of the tropical cyclone cloud from satellite images is a basic and important fact of the forecasting system. Especially used satellite images-based system, there are many different techniques to locate the tropical cyclone from the input satellite image. In early 1975, the most essential and weather forecasting tool was developed by Dvorak technique. It is to determine the intensity of the cyclone according to the pattern of the cloud. Nowadays, the high beneficent Tropical Cyclone (TC) archives gathered by this technique. The best track of historical TC datasets is the foundation for the prediction of risks from TCs effect the region with only computer program. Satellite based TC intensity estimation method is the major interest to meteorologist. This technique based on the pattern of the cloud such as curved band, shear, cloud's eye the central dense, embedded center and central cold cover. There are two steps in this method, step one is to find where the cloud's center is. The second step is to analyze the cyclone cloud pattern [38]. Figure 2.4 illustrate the references for typical cloud pattern.

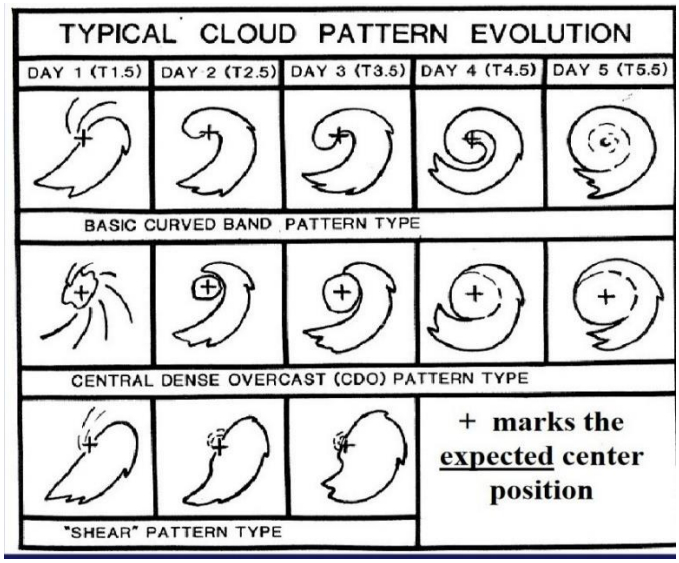


Figure 2.4 Typical Cloud Pattern Evolution

Based on the Dvorak technique, the researchers were developed many advanced Dvorak techniques [22] developed ADT. The Advanced Dvorak Technique (ADT) was developed by the Timothy L. Olander and Christopher S. Velden, Space and Science Engineers at the University of Wisconsin-Madison. This technique can be used to identify the cyclone intensity from the beginning of the cloud formation form development to disappearance. In Dvorak method, pattern evolution and rules for associated patterns are determined to the intensity of the cloud [9]. The approach for

determining intensity and incorporating several rules and analytic techniques has been designed in the ADT to closely resemble the SDT methodology. The ADT amended portions of the original SDT rules, and the rules and performance analysis of the ADT are predicated on the modified rules. Moreover, based on regression methods, ADT modified some intensity relationships. The original objective satellite estimating methods used in the development of the ADT were created at Colorado State University's Cooperative Institute for Research Applications and the University of Wisconsin's Space Science and Engineering Center. The ADT has undergone significant additions and revisions, with the consequence that the algorithm's technique, functionality, and content are very different from those of its predecessors. [34]. The original modifications from earlier digital Dvorak methods cover the adding of historical data, implementation of a time-weighted averaging arrangement, new meanings, and resolution of different environmental temperature values, parameters and intensity estimation. These modifications have brought more reliable and accurate intensity estimation.

Relationships between tropical cyclone cloud and intensity are also important considerable data in cyclone classification. Charles P. Arnold describes the study of cyclones based on infrared Satellite images. Four stages of cyclones are the formation of storms, tropical depressions, tropical cyclones and typhoons. Characteristics of cloud, wind and thermodynamics are connected to these stages. Storm size, types of banding storm, type of circulation center, intensity and cyclone's eye are the physical characteristics [48]. The ratio of cyclones is surrounded by cirrus, different cyclone type and distribution are the cloud characteristics. vertical motion, velocity, tangential wind, moisture content, and temperature were among the wind and thermodynamic parameters examined. The results of this research expose a consequential variance made in the cloudiness of tropical cyclones. This study has also displayed that great alternation in the tropical cyclone typically is discovered in multi-cellular complexes.

In this paper [24], an automated tropical cyclone detection system uses a novel deep learning model. This model applies a mask R-CNN model to segment the cyclone and wind speed filter. To discover a TC for the maximum amount of number of its life process, ML pipeline method was proposed by satellite images hold at every 6 hours' gap. This result also produced confused dataset with segmentation masks for each satellite image and can easily be at GitHub repository2 publicly. The satellite data were

gathered from Meteosat 5 and & in Indian Ocean Data Coverage. From year 2002 to 2016, every 6 hours frequency satellite images can be gathered. ML pipeline was developed in four stages. The first stage is to pass an input satellite image to the detector. If one or more enclosing boxes are found, the wind speed needs to be examined. In the third stage, the classifier supports to extract ROI more than one forecasting for a satellite image using the box coordinates [7]. The last stage is to determine the cyclone from the cropped images with the highest assurance sore from the classifier. 86.55% accuracy rate is obtained by using this method.

Another technique is to detect the center of TC automatically from multi-band cloud images by using SCBeM technique. This research studies the digital IR1 and visible and water vapor satellite images gathered by FY-2C from year 2012 to 2016, 4-year data. From an input image, the fixed-size domain was extracted by the following steps. First of all, the expected region was gained from the NMC of China or the latest predicted location [13]. Then, the expected region as the center of the input image, 10° was extracted from an input image as a square area. To determine the fixed data, a fusion of multi-band images and TCSS extraction is used. Different number of TC intensity, the area of the Latitude and the location where TC occurred are the essential facts to form the TC spiral cloud belts. There are two conditions in which the TC makes landfall or the formation of TC, the cloud bet is chaotic without ambiguous clouded and it is not a spiral shape [50]. TC can get easily in the condition of a strong cloud. And necessary addition needs to get the ground-based radar that TC will make landfall. The accuracy is good but many conditions need to be considered to complete this technique, such as terrain, wind shear value and flow need to be considered, there will be more accuracy rate will be gathered in the construction of the CBT.

Typhoon is a type of cyclone and its intensity is greater than another cyclone. Typhoon also brings flooding, severe winds and thunderstorms. Many researchers were developed different type of method to detect typhoon centers and tracking the moving of the typhoon from the satellite images. This paper [16] was developed a novel method that partition the input satellite image into sub pieces to bring out feature of the typhoon. Morphological image processing was used to image enhancing to get the location of typhoon and statistical image classification method to discover the center of the typhoon [48]. By using different image resolution and ROI, vector quantization method is used to encode satellite images to reconstruct image. From year 1995 to 2006, 71

typhoons' historical infrared satellite images are used in this research. By using this technique, the location of the typhoon cloud can be extracted from the satellite image with a fusion background. This expected method could recognize the detection of typhoon clouds simply and fast without complex calculation by using the features such as eyed section, spiral cloud belt and cloud wall compared with other methods. Getting the accurate location of the typhoon cloud can be helpful to predict the track of the typhoon movement.

Deep learning Convolutional Neural Network (CNN) is one type of deep learning algorithm and also a feed-forward neural network, that can easily feedback to nearby disclosed networks through artificial neurons, so getting a deep learning algorithm for fast feedback to data. To form a network topology, a convolutional layer and a sampling layer are needed [29]. CNN applies the method of back propagating neurons to understand the renewed network of each neuron's data. Many type of research have been developed to predict the typhoon tracking mode, but some problems remain such as the prediction accuracy rate being poor, cannot to recognize the location, and confused feature extractions discovered in the previous prediction model. This paper [35] contributed to overcoming this problem with a modified GCLSTM model to locate typhoon clouds and prediction models effectively. The performance of using this algorithm sounds good. Compared with other neural network models, 13.3% was improved and the accuracy of the prediction of the typhoon was 95.12%. But pros and cons are also together with this research. Some limitations of this research are only one GCN neural network being performed and the second limitation is the actual application will involve a huge amount of data but this model does not involve it. For future work, this research is analyzed by using the satellite-based cloud typhoon prediction model to prevent loss and damage to people’s lives [42].

Prevention disaster mitigation is the major role for environmental prevention system. Determining TC intensity is the basic step for the tracking system. Machine learning also used to create this application [19]. Supervised and unsupervised learning are the type of machine learning. Supervised learning must have the historical statistic data needed. Support Vector Machine (SVM) is a supervised learning system and it can use in the classification process. In this research, the input infrared image is extracted into 14 GLCM features such as Grayscale, Ycbcar and Red Green Blue (RGB) to recognize image. Combination of these features to become the new features that to be

bring out the classification stage. OAO and OAA coding design was used to test the classification [25]. From the result, 88% of accuracy rate of tropical cyclone intensity are gathered by using with the saffir-simpson hurricane wind scale is shown in Figure 2.5.

Category	Wind speeds (for 1-minute maximum sustained winds)			
	m/s	knots (kn)	mph	km/h
Five	≥ 70 m/s	≥ 137 kn	≥ 157 mph	≥ 252 km/h
Four	58–70 m/s	113–136 kn	130–156 mph	209–251 km/h
Three	50–58 m/s	96–112 kn	111–129 mph	178–208 km/h
Two	43–49 m/s	83–95 kn	96–110 mph	154–177 km/h
One	33–42 m/s	64–82 kn	74–95 mph	119–153 km/h

Figure 2.5 Saffir-Simpson Hurricane Wind Scale

Satellite based tropical cyclone detection system can be performed by using only Morphological image processing method. It can detect the location of the cyclone cloud automatically. In this research, several preprocessing methods were used to get the clear structure of the cloud from the input satellite image [37]. After doing this step, extract the feature from the skeleton cloud. Least square error method was used to compare the inner, outer and skeleton curves of the cloud structure. Some logarithmic helix to suit the tropical cyclone feature and the core point of the helix is the core point of the cyclone; confirming to its moving structure, a rotation matching technique, the rotation center is the cyclone core point. Some experimental results show a high accuracy rate and low error [49].

To detect the eye of typhoon center was proposed by using a new AI algorithm. This paper was also used morphological image processing as the pre-processing stage. Opening and closing are the basic structures of image processing. Using the entropy-based threshold techniques, the features can be extracted and the gravity calculation technique can locate the core of the typhoon. Compared with other typhoon center location methods, the deviation value is small [10].

An effective method for typhoon core position is presented using fractal feature and gradient of satellite image. Before the typhoon makes landfall, the centers of typhoon clouds are normally pointed in this region for a typhoon [31]. The features of

the cloud region are smooth and have a higher level of gray values than the thickness of borderlines. So, the window evaluates technique is proposed to locate the right cloud position. Finding the window with the largest difference value between the total of the fractal dimension and the gray-gradient co-occurrence matrix yields the closed cloud region. The temperature value near the eye of typhoon clouds is small. These factors must be considered to construct the model. This system uses a window to traverse the dense cloud region. If there is a closed curve, the region of the curve is considered as the typhoon center region [42]. If there is a curved cyclone region, it can be assumed as the typhoon center. Another considerable factor is the largest density and most texture structure. The information that is bring out from the strong and dense eye typhoon is an important factor. If the location of the eye can't be examined, the small region of the geometric center can be considered. This research is tested by using Chinese FY-2C stationary satellite image. Many other detection typhoon centers in "Tropical Cyclone Yearbook" which was produced by Shanghai typhoon institute of China Meteorological Administration is compared with this result. The outcomes of the experiments show that high precision is achievable. [23].

2.4 Type of Tropical Cyclone Tracking Systems

When the pressure is low with warm core, a tropical cyclone forms over tropical and subtropical waters. A huge amount of energy from the sun are given to the tropical and subtropical oceans and become water vapor. The charge of heat drives an upward direction of air, causing a low-pressure zone that is set into spiral by the circulation of the earth [28]. A cyclone occurs when the energy involved in the circling airflow is prominent. Cyclones are formed, made motion and finally disappeared over sea; they cannot be correctly detected from land. TIROS-1 was the first weather satellite and it was launched in 1960s. Satellite images have become input data sources for prediction weather forecasting system. Cyclone's location, eye, intensity, storm surges and rainfall are also correlated features for forecasting system [27]. All forecasting methods use the recent past behavior of the current cyclone as well as the historical behavior of similar cyclones that have already occurred. Current cyclone is similar to the previous cyclones so their behavior and manner will be same. In the prediction stage, based on the previous cyclones the manner of current cyclone is called predictors. There are many other tropical cyclone forecasting techniques are existed and all of them are used historical data to forecast [14].

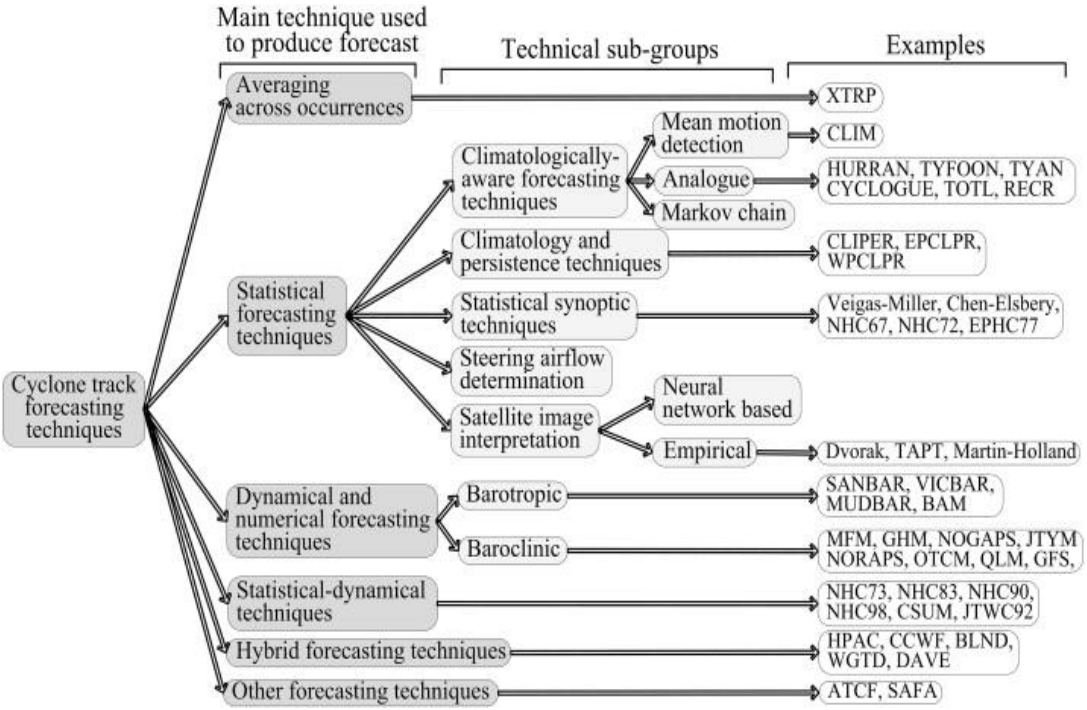


Figure 2.6 Main Techniques, Technical Sub-groups and Examples of Cyclone Track Forecasting Techniques

2.4.1 Tropical Cyclone Formation Basins and their Characteristics

There are seven tropical and sub-tropical basins around the world as shown in figure 2.7. They are

1. Northwest Pacific basin
2. Northeast Pacific basin
3. North Atlantic basin
4. North Indian basin (Bay of Bengal and Arabian sea)
5. Southwest Pacific basin
6. Southeast Indian Ocean basin
7. Southwest Indian Ocean basin

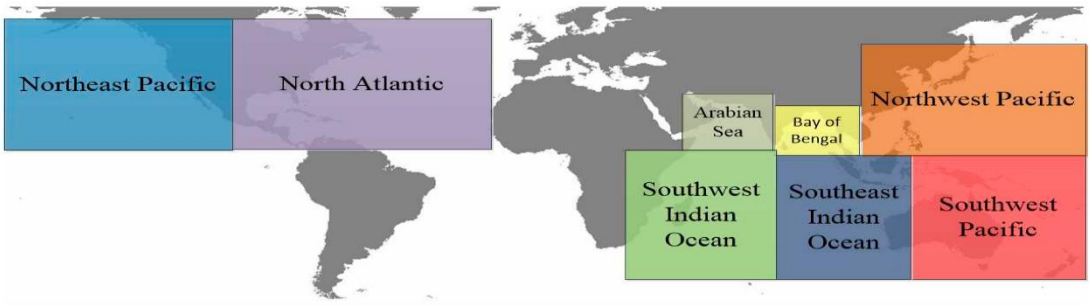


Figure 2.7 Basins Where Tropical Cyclones Form on a Regular Basin

2.4.2 Statistical Forecasting Techniques

Regression model are mainly contributed in statistical forecasting method. There are five major types of statistical forecasting techniques [15]. They are:

1. Climatologically-aware forecasting techniques
2. Climatology and persistence forecasting techniques
3. Statistical synoptic techniques
4. Steering airflow determination
5. Statistical-dynamical techniques

Short-time means (24 hours) and long-time means (72 hours) forecasts can be created using statistical data. The predicator data set is made up of information from the current cyclone, the prior cyclone, the synoptic research, and numerical experiments. Any fusion of criterion or variables in the observed data set can be taken into consideration is the major advantages of statistical method. Generally, statistical methods have also some disadvantages [36]. Statistical regression technique makes predictions that display the average manner of cyclones in the historical data set. Therefore, statistical method operates the best when the current synoptic condition, as obvious in current cyclone movement, does not come too much from the typical climatology of the basin. Moreover, these statistical techniques relay on high quality of data in order to reliably forecast the statistical trends in the cyclone motion.

This research proposed combination of short-range tropical cyclone system with upgraded traditional statistical method by using satellite images that are extracted as 9 features as the input data. Input satellite images are gathered from two resources that obtained from Japan Metrological Agency Satellite and Tokyo Typhoon Center [18]. Traditional method uses very large amount of input variables and it needs to run high end computer, this proposed method can reduce the input variables and it is portable to run the model. The performance of the average error is smaller than the traditional model. But it is still challenging for automated typhoon center location and forecasting of the typhoon movement have high error [43].

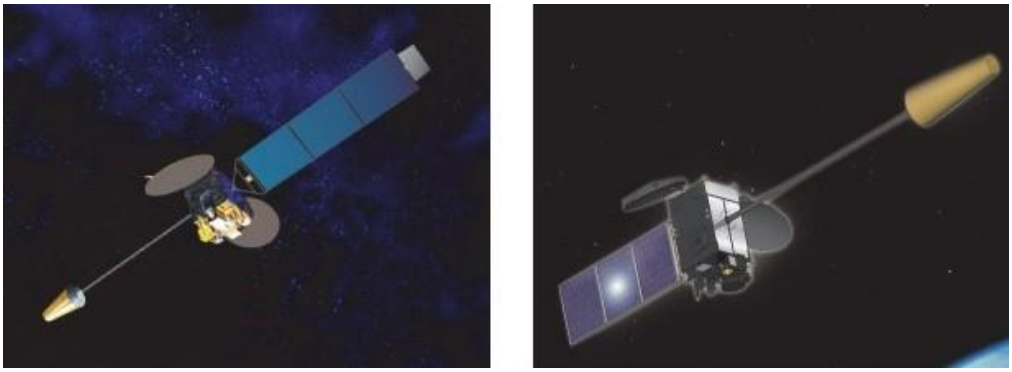


Figure 2.8 Satellite from Japan Metrological Agency and Tokyo Typhoon Center

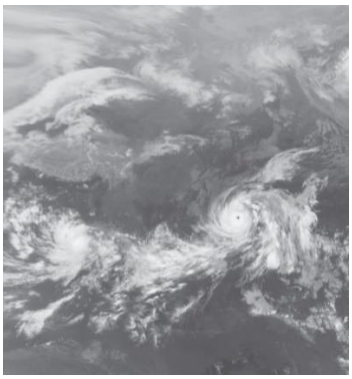


Figure 2.9 High Resolution of Typhoon Image by MTSAT-2 on 8 October 2014

2.4.3 Climatology and Persistence Techniques

The motion of the cyclone goes in the same direction when the atmospheric conditions are constant. An averaging across occurrences is based on this obedience or can cause satisfactory predictions for short-term duration [34]. When the atmospheric conditions are not consistent within a cyclonic occurrence, for long-term forecasts it is essential to know which weather climates are managing the cyclone motion. This data can be gained by including actual storm tracks in the same basin as the present storm, and under similar weather climates to the currently determining ones. The climatologically-aware forecasting model is to consider previous data under similar weather climates when forecasting the present cyclone motion. These methods perform satisfactorily in long-term projections. [35].

2.4.4 Statistical Synoptic Techniques

Air pressure is inversely proportion to the height of the sea level. This incremental decreasing of air pressure with increasing sea level may also make other changes in weather climate conditions. Statistical synoptic methods are contributed by

the differences of air pressure at different pressure levels [20]. Another use of pressure level data as one type of forecaster has made these methods bring out from the climatology and persistence method.

2.4.5 Forecasting Techniques Based on Satellite Image

Cyclones generally move long lengths over the basin before they make landfall. Sometimes ground-based observations cannot be affected because of the long lifetime of cyclones over the oceans [19]. At that time, forecasting the cyclone track must be observed with detailed knowledge from the formation of the cyclone to make landfall. The pattern of the cyclone movement is also important for forecasting. In this condition, satellite images can support valuable instruction for the entire life of a cyclone. Nowadays, many researchers have developed various methods to predict cyclone tracks by using satellite images as historical data [33]. To translate satellite images, statistical equations are used and make predict cyclones, they are computed under statistical methods. Moreover, most other techniques use satellite images to get more accuracy rate.

In predicting of the weather forecasting system, clouds play an essential role. Satellite images can give distinct view of cloud motion and behavior of cloud manner in every hour. Forecasting the cyclone trajectory is the major role for disaster mitigation and decreases loss of lives and households. Previous satellite images are mainly gathered from the Doppler Radars in India. But it has a limitation and semi-automatic system, the researcher proposed to create an automatic tropical cyclone prediction. To extract the location of the cyclone and features such as shape, texture, and color, Fuzzy C-means clustering method is implemented. Observation of the different features between multi-dated satellite image compares this difference and predicts the tropical cyclone movement [40].

According to the difference satellites have launched, satellite images have different resolutions and background noisy. This paper is proposed the automated tropical cyclone motion system that uses multiple satellite images from five satellite sources and wind field images [12]. There are three main parts that are data processing to detect the location of cyclone from the satellite image, eye detection algorithm and combining the eye detection results from different sources by using filter-based predictor. Data processing using morphological image processing method. To detect

the cyclone eye, the first step is to segment the cyclone eye by using wind speed value and ensemble classification using histogram of wind speed and wind direction, histogram of speed-to-direction, and outstanding wind direction. Graph based tropical storm eye detection algorithm is used [37].

To establish a tropical storm forecasting system using a single orbiting satellite to capture a extended behavior is inappropriate because of its limitation of spatial and temporal scope. One finding way is to use multiple images from multiple satellite sources for cyclone forecasting. Sometimes, data from some satellites do not support features as useful as other orbiting satellites in the detection of cyclones. Moreover, weak cyclone cannot provide distinct features while strong cyclone has obvious features. In this paper, a tropical cyclone prediction system is used weak cyclone data from multiple sources. Images of 3 hours using TRMM and Quiescent satellites. Linear Karman filtering method is used in this research. Using only the satellite data with strong cyclone selection features reduces the detrimental effects of blocking and coarse temporal resolution associated with this framework.

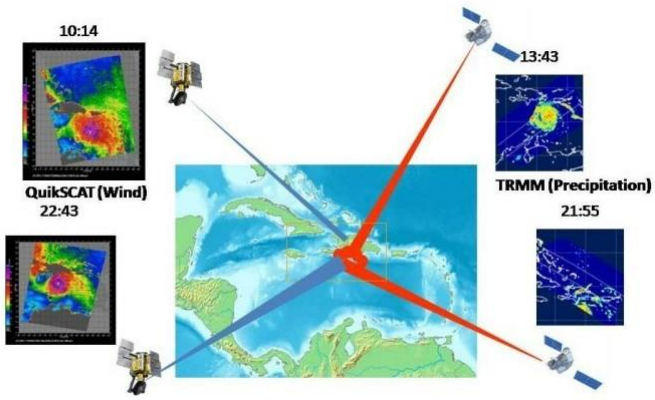


Figure 2.10 Satellite Images from Multiple Sources

2.4.6 Empirical Forecasting Techniques

This method needs the expert who decides on the motion of the cyclone. The forecaster's experience level is mainly impacted by this techniqueMajor errors related to storm acceleration into mid-latitudes and missed re-curving can be eliminated by a forecaster with pattern recognition expertise. Therefore, major errors will occur by these experts recognize the pattern of cyclones. The main disadvantage of these techniques is that the forecaster needs to take the time to become an expert. The Dvorak technique is also empirical forecasting technique and determine the intensity of the

cyclone based on the manner of the cloud by using Satellite Images. Different stages of cyclone development, cloud patterns will be changed and the forecasting is based on these patterns. Not only the cyclone's intensity but also the cyclone's near-future are also being predicted through translation of current cloud patterns in the satellite [40].

2.4.7 Techniques using Artificial Neural Networks

In typhoon prediction and tracking systems, satellites images are also the most important for gathering information about the tropical cyclone and its near environment. Traditional prediction techniques face challenges that combination of information gathered through the sensors, and recognition of spatial pattern. Numerical, and statistical techniques can perform these barriers but it is complex task and needs high-end computers to operate [45]. Moreover, satellite images are very huge in size and noisy. According to these barriers, Artificial Neural Networks (ANNs) is proposed for data processing.

This paper [28] is the typhoon track forecasting system using a generative adversarial network (GAN) with satellite images as inputs. As a training date, the historical multi-dated satellite images of typhoons which occurred in the Korea Peninsula are used the trained GAN is operated to create a 6 hours motion of a typhoon for which the GAN was not trained. The forecasted motion of a typhoon image generously detects the next position of the typhoon eye as well as the shapeless cloud manners. Distance errors between forecasted centers and actual typhoon centers are calculated numerically in kilometers. 10 typhoons are tested and the average error is 95.6km. There is a limitation about the changes of cyclone track direction into westward northward directions is pointed out when the prediction is remarkably improved when satellite photos are combined with velocity variables.

Since 1975, Dvorak technique has mainly used as to detect the intensity of cyclones. Detecting the intensity of tropical storm is the most important step in TC motion forecasting system and the prevention of natural disasters. But Dvorak technique is sometimes misleading to consistent intensity estimation. So, deep-learning model using satellite images was developed to evaluate storm intensity. Compared with the Dvorak technique, CNN model is flexible with the variable among various basins [33]. A revised approach is also suggested in this study, which aims to further increase accuracy by attempting a post-smoothing step and utilizing the basin, day of the year,

local time, longitude, and latitude. CNN-TC is the first CNN model that can forecast the intensity of tropical cyclones using regression analysis. The training dataset consists of 1407 worldwide TCs from 2003 to 2014 and further data from 188 TCs from 2015 to 2016. 94 TCs during 2017 were used for testing data to determine the performance of the CNN model [16].

Another paper proposed Long Short-Term Memory LST based Recurrent Neural network RNN model to forecast cyclone track using GridIDs. That uses data of the initial time of occurring cyclones and forecasts cyclone track for the coming several hours. Central pressure, maximum sustained surface wind speed, latitude, and longitude are used as input data and which are accessible at a constant time interval. This network can represent the fusion nonlinear manner of cyclones and plays reasonably in terms of MAEs outperforming past grid forecasting results. Forecasting cyclone motion well in advance supports in performing preventive measures much more quickly to decrease damage and property loss [30].

This research presented to establish the multi-layer neural network by using satellite images from NOAA-AVHRR. This proposed network consisted of five layers, an input layer, three hidden network layers, and an output layer. This is a bidirectional network layer except for an input layer. The features of small clouds in input satellite images are encoded by the first layer V1. 66 × 66 pixels satellite images are preprocessed to transform grayscale image and chaining the value into the interval [0,1]. The trained network can create the correct cyclone forecasting for 98% with little variations in the cyclone position [39].

2.4.8 Hybrid Radial-basis-function Network

A hybrid radial-basis function (HRBF) network can identify cyclones automatically from input satellite pictures and can accurately forecast the cyclone's track and strength. There are two components to this network. The initial step involves identifying the characteristics of the retrieved cloud and extracting the cloud patterns from the input satellite photos. The second component aims to categorize the data from module one and forecast the cyclone's velocity and strength. Module 1 is further distinguished into three different parts: Sub 1: Bring out of the cloud manners of tropical storms using Gabor filters that are used to carry out the cloud manner and these filters are used as locators of small, isolated cloud extends or balls in the satellite image [44].

As these predictors are isolated, a huge number of Gabor filters are operated to conceal the entire image. The products from these predictors are then linked and operated to identify the cloud structure at massive scales in the satellite image. Sub 2: Composite Neural Oscillatory Model is used as segmentation. Sub 3: NOEGM Model is used for recognition and classification method. HRBF network is used to define the intensity of the cyclone and to track the tropical cyclone motion.

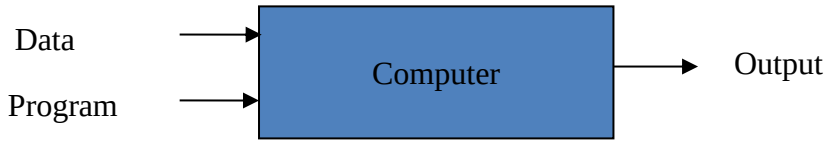
2.5 Chapter Summary

This section explains the nature of tropical cyclones, their formation, common tropical cyclone patterns, impacts, and types of cyclones. This chapter intends to highlight the types of cyclone cloud classification techniques and types of tropical cyclone tracking systems within the research area by making a carefully analyzed literature review. And also explains the forecasting techniques such as statistical forecasting, climatology and persistence, statistical synoptic technique, forecasting based on satellite images, empirical forecasting, artificial neural networks, hybrid radial-basic function networks, etc.

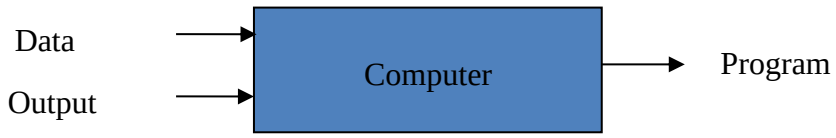
CHAPTER 3

THEORETICAL BACKGROUND

There are many definitions of Machine Learning (ML) from the diversity of area. Form the side of technical view, machine learning is different from traditional programming language. In traditional computer programing, data and program are passed the computer and computer produce the output. In machine learning, data and output form experiences are passed to the computer and computer produces the program according to the nature of data and experience from the output [21]. Computers are trained by machine learning, which uses the experiences of humans and animals to teach them tasks. Machine learning algorithms use computer approaches to "learn" information directly from data, as opposed to using a pre-existing equation as a model. The more examples that are available for learning, the more adaptively the algorithms work. [32]. Figure 3.1 describes the nature of traditional computer programming and machine learning system.



Traditional Computer Programming



Machine Learning

Figure 3.1 Traditional Programming Versus Machine Learning

ML is a subgroup of computing theory in artificial intelligence and fraction of computer science that developed from the appliance of pattern recognition. Machine learning algorithms establish the model constructed on dataset that can train and test the new dataset to produce forecasting. These algorithms execute by developing a prototype from a training dataset to make data-driven forecasting or determinations,

rather than following exactly defined program orders. Machine learning is nearly related to and usually interconnects with geometrical information; a regulation that also practices in forecasting-making. It has powerful attaches to mathematical improvement, which produce models, assumption and application areas to the sectors. Machine learning is occupied in an area of computing functions where developing and programming explicit theory is unusable [50]. Different disciplines of knowledge and the discipline of machine learning systems are shown in Figure 3.2.

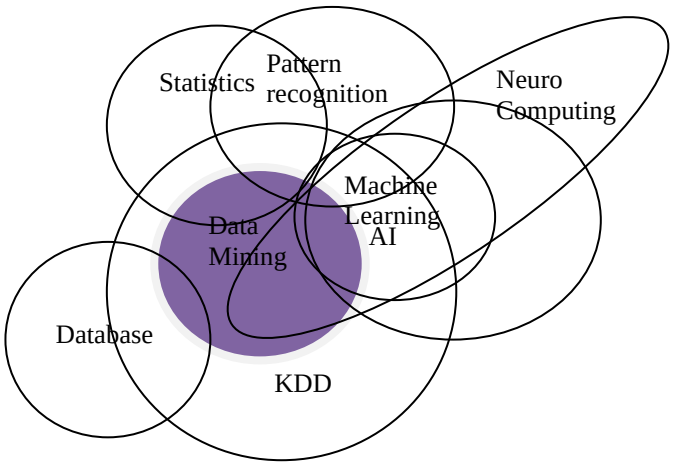


Figure 3.2 Different Disciplines of Knowledge and the Discipline of Machine Learning

3.1 Real World Applications Using Machine Learning

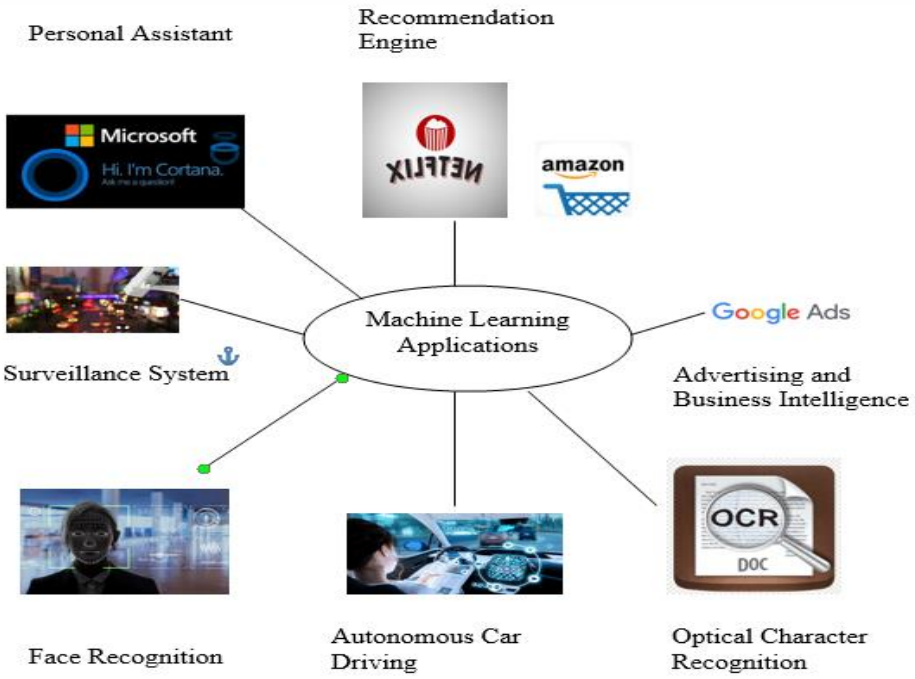


Figure 3.3 Applications Using Machine Learning for Different Purpose

Real world applications by using machine learning algorithm are shown in Figure 3.3. With the increasing in big data, machine learning has developed generally important for determine difficulty in application like these below [22]. Example areas include

1. Computational finance, for credit scoring and algorithmic trading
2. Image processing and computer vision, for eye recognition, weather forecasting, and GIS system
3. Computational biology, for tumor detection, drug discovery, and DNA sequencing
4. Energy production, for price and load forecasting
5. Automotive, aerospace, and manufacturing, for predictive maintenance
6. Natural language processing [41]

Moreover, machine learning is not an easy step. There are seven steps bring to machine learning and each step decides with these input datasets.

3.2 Seven Steps of Machine Learning

There are seven steps of machine learning algorithm as shown in Figure 3.4.

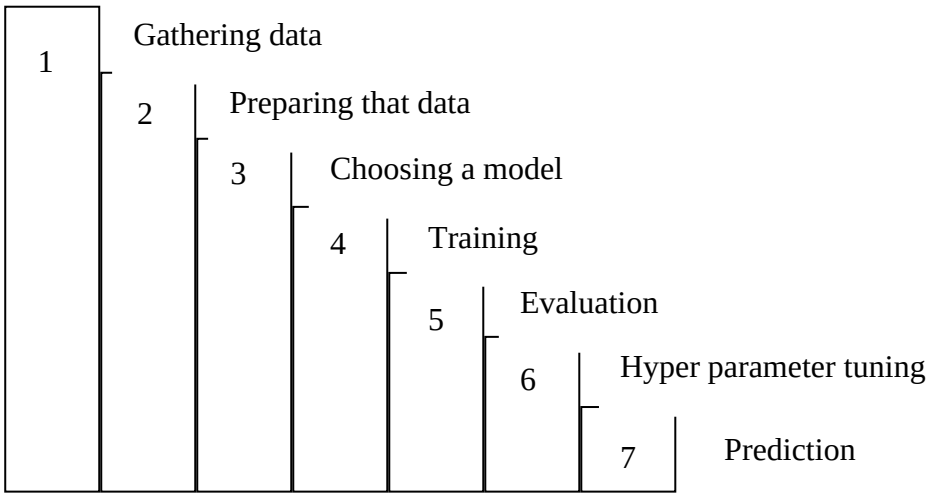


Figure 3.4 Seven Steps of Machine Learning

1. **Gathering data:** Machine learning needs training dataset. Texts, Graphs, Tables, Numbers, Clickstreams, Videos, Images and Transactions are also datasets and it come in different dimensions [36].
2. **Preparing data:** In machine learning system, raw data alone is not very convenient. The data needs to be developed, standardized, non-copied and

mistake and bias required to be extracted. Envision of the data can be applied to consider patterns and outliers to see if the correct data has been gathered or if the data is removing.

3. **Choosing a model:** The next step includes choosing the correct model. Many different models can be applied for many reasons. Based on choosing the model, need to confirm that the model catches the business plan. In addition, three facts need to be known, the first one is how much preparation the model needs, how exact it is and how scalable the model is. Sometimes the simple model is more suitable for the plan even though some models are more complex. Various techniques are available, including but not limited to logistic regression, decision trees, K-means, principal component analysis (PCA), support vector machines (SVM), random forest, neural networks, and naïve Bayes. [11].
4. **Training:** Training the model is the key portion of machine learning system. Training the data is more increases the accuracy of the forecasting decision for the model [44]. One training step is one epoch of weight and bias change. While unsupervised machine learning attempts to generate hypotheses from raw data, supervised machine learning builds the model using classified sample data.
5. **Evaluation:** Evaluation is also the main step after training model. This step is needed to know which data set is unused. Evaluation step can determine how the model operates in the business application. The enormous amount of training and testing data are needed for the real-world application.
6. **Parameter tuning:** Parameter must be tested to get more accuracy to develop the Artificial Intelligence. Increasing the epochs of the training data can get the more accuracy. Moreover, changing parameter can get more accuracy for the mode. This is an experimental process.
7. **Prediction:** After doing the six steps, the final step is prediction. Prediction is the answer of the whole problem. Using training and testing data, evaluation step to determine the strength of the model and making the changes the parameter value for the model, prediction is the key point for the machine learning system [27].

3.3 Types of Machine Learning

There are three types of machine learning. They are supervised learning, unsupervised learning and semi-supervised learning as shown in Figure 3.5.

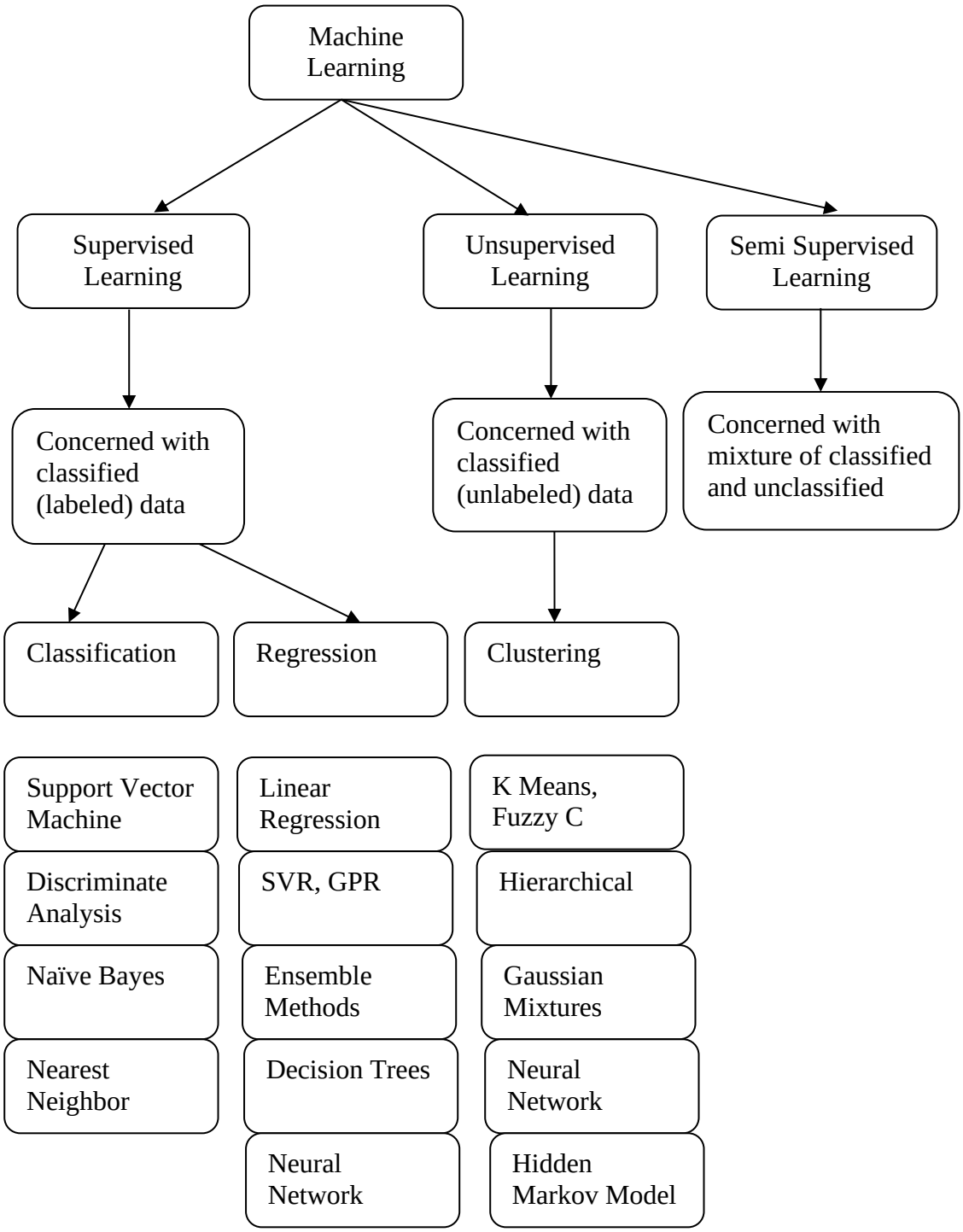


Figure 3.5 Types of Machine Learning

3.4 Supervised Learning

The goal of supervised machine learning is to construct a model that makes decisions depend on documentation in the presence of uncertainty. A supervised learning algorithm uses a known set of source data and known feedbacks to the output information and trains a model to produce acceptable forecasting for the answer to new data. Supervised learning applies classification and regression methods to improve predictive models [29]. Figure 3.6 describes the nature and feature of supervised learning system.

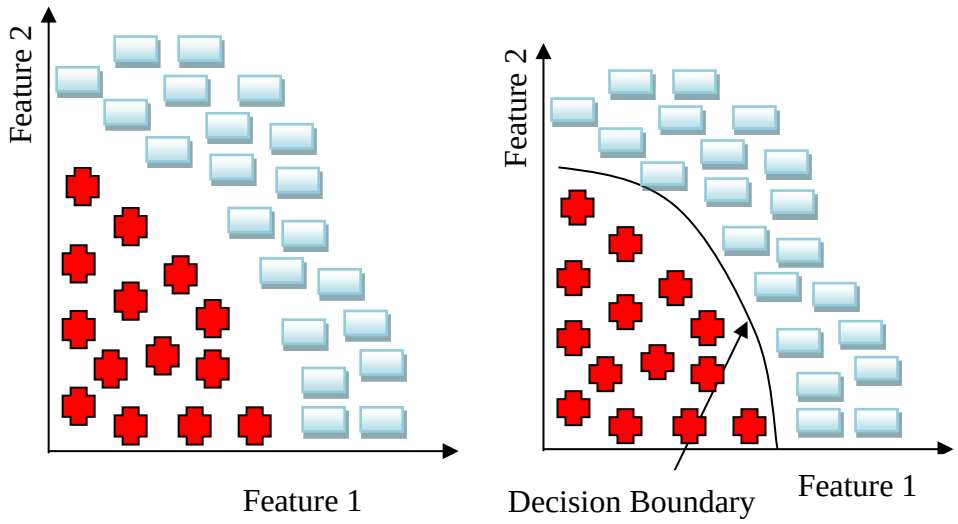


Figure 3.6 Supervised Learning

Classification techniques react discrete values; for instance, whether a tumor is malignant or benign, or if an email is spam or real. Input data is categorized using classification models. Credit scoring, speech recognition, and medical imaging are examples of common uses. [42].

3.4.1 Support Vector Machine (SVM)

Support vector machines (SVMs) are type of supervise learning system for classification and regression [1]. SVM can be clarified as a system which apply hypothesis area of a linear functions in a high-dimensional feature space, taught using optimization approaches' learning theory that performs a learning bias come from statistic. It can be used for real world applications, such as OCR, face recognition and so on, mainly for pattern classification and regression-based applications [15].

There are three types of SVM such as

1. Optimal Linear Classification
2. Nonlinear Classification
3. Nonlinear SVM and Kernels

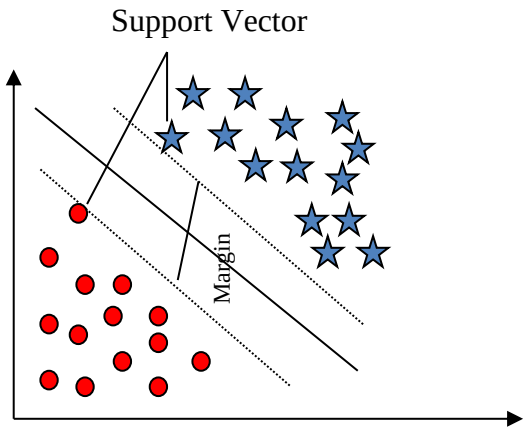


Figure 3.7 Optimal Classification Hyper-plane.

It depends on the hyperplanes that distinguish decision margins between the different class labels. In SVM method, the classification is applied by detecting the hyperplane that maximizes the boundaries between the two class labels. SVM can used when input data has exactly two classes. The best hyperplane for an SVM expresses the largest boundaries between the two labels. Boundary expresses the highest width of the slab parallel to the hyperplane that has no data exist in this boundary. The following figure demonstrates these meanings, ● denoting A type data, and ★ denoting B type data as shown in Figure 3.7 [26].

A single hyper-plane P can be separated linearly within the training data is the simple form in SVM. This hypothesis is comparable to state that the parameters (y,a) of the classifier must meet the set of constraints

$$(y^T x_i + a) d_i \geq 0, i = 1, \ldots, B. \quad (3.1)$$

3.4.2 Nonlinear Classification

Nonlinear classification means that some input datasets cannot be spate linearly by borderline margins and in this situation nonlinear classifier needs to produce features immediately from the input data set. In most applications, this outcome can be performed simple but adequate non-linear transformations of the input data. This strategy is simpler to the multilayer perceptron [47] In any MLP, the output layer

performs as a linear classifier but the pre-processing of the input by one or more non-linear hidden layers may perform a transformation of the input data such that the output layer executes on a linearly separable feature space [19].

Nonlinear SVM and kernels

Kernel function is used in nonlinear classification systems and there are many different types of kernel functions. Some kernel functions are described as

1. Polynomial
2. Radial basis function
3. Sigmoidal type

3.4.3 Point in Using SVMs

Every classification method has benefits and drawbacks, which vary in importance depending on the data being studied and have a relative relevance. When there is non-regularity in the data, such as when the data are not regularly distributed or have an uncertain distribution, SVMs can be a helpful tool for insolvency analysis. It can assist in evaluating data, such as financial ratios, which need to be converted before being entered into the scoring system for traditional categorization methods. The benefits of using the SVM technique can be summed up like this:

1. Due to the kernel's non-parametric and locally operating function, SVMs are now more flexible in selecting the threshold form that separates solvent from insolvent enterprises. This threshold need not be linear or even have the same functional form for all data. Because of this, they are able to work with financial ratios that have a non-monotone relationship to both the score and the default probability, or that are non-linearly dependant, all without having to perform specialized work on each non-monotone variable. Since the kernel implicitly contains a non-linear transformation, no assumptions about the functional form of the transformation, which makes data linearly separable, is necessary. The transformation occurs implicitly on a robust theoretical basis and human expertise judgement beforehand is not needed.
2. If the parameters C and r (in the case of a Gaussian kernel) are selected properly, SVMs offer a good out-of-sample generalization. This indicates that SVMs can be robust even in cases when the training sample has some bias, provided that

the right generalization grade is selected. Outliers can be rescaled by selecting different r values for various input values.

3. The optimal problem is convex, which allows SVMs to provide a unique solution. This is advantageous over Neural Networks, which might not be as robust across various samples because to the many solutions linked to local minima.
4. By selecting a suitable kernel, like the Gaussian kernel, one may emphasize the degree of similarity between two organizations. This is because the greater the kernel value, the more similar the financial structures of the two companies are.

3.4.4 Discriminant Analysis

The process of determining which weightings of quantitative values or predictors may distinguish between two or more groups of cases more effectively than by chance is known as discriminant function analysis, or discriminant discriminant analysis. Depending on a linear combination of the weightings and counts on these values, the analysis causes a discriminant algorithm. The vast quantity of functions can be expressed as the smaller of the two following values: the number of groups minus one or the number of predictors. [34]. There are two possible objectives in discriminant analysis: seeking a predictive function for classifying new groups or interpreting the predictive function to better recognize the interconnection that may consist among the values [8].

3.4.5 Naïve Bayes

In a supervised learning system, many famous classification methods including Bayesian are also statistical methods for supervised learning systems. Naïve Bayes is an underlying probabilistic model and it permits catching uncertainty about the model in a principled way by deciding probabilities of the results. It can solve diagnostic and predictive applications. Bayesian classifier supports effective learning methods and prior knowledge and discovered datasets that can be cooperated. Bayesian Classifier supports a helpful perspective for realizing and estimating many learning theories. It computes explicit probabilities for the hypothesis and it is robust to noise in the input data set [20].

This algorithm can be coded up easily and the predictions made quick is an advantage of Naïve Bayes Classifier. Therefore, it is simply scalable and is

traditionally the algorithm of choice for real-world applications that are needed to reply to user's requests instantaneously.

3.4.6 Nearest Neighbor

The Nearest Neighbor classification is determined by the closest neighbors' class. The simplest prediction algorithm for a test data set is this one. It's easy to complete the training phase: just retain each indicated training. The basic idea behind this technique is to calculate the distance between each training data set and the input. Then, store the k closest training data, where $k \geq 1$ is a fixed value. Find the class that appears most frequently in these records. The prediction for this testing data is this class. Using the same set notation as above, the nearest-neighbor method is a function of type $(A \times B)^n \times A \rightarrow B$. A distance function has type $A \times A \rightarrow R$. This basic method is called the k Nearest Neighbor algorithm. To select a design, you must know the value of k and the distance function to utilize. The most popular option for k when there are two different classes is a tiny odd integer, like k = 3, to prevent ties. Even when k is odd, ties are still conceivable if there are more than two classes. When two distance values are the same, ties can also form. There is disagreement over the most effective method for breaking ties in a kNN implementation.

Euclidean distance is the most common function to find the distance between the example data:

$$d(a,b) = \|a,b\| = \sqrt{(a,b).(a,b)} = \left(\sum_{i=1}^{m}(a_i - b_i)^2\right)^{1/2} \quad (3.2)$$

where a and b are points in X = R m. But it is necessary complex time to compute the distances for each examples and it is needed to recomputed until the fixed value is the disadvantage of kNN algorithm. Assume that Rm has n training data sets. A linear classifier, like a perceptron, can be applied in O(m) time, however the method takes O(nm) time when applied to a single test set of data. Should a complex data structure be used to hold the training data, for example a kd-tree, then finding nearest neighbors can be done much faster if the dimensionality m is small [11]. However, for dimensionality $m \geq 20$ about, no data structure is known that is useful in practice.

3.5 Regression Techniques

These techniques forecast continuous values for example, changes in temperature or fluctuations in power demand. Typical applications include electricity load forecasting and algorithmic trading [23]. There are many different regression techniques including assemble method, decision tree and neural network are famous.

3.5.1 Ensemble Methods

Ensemble methods are mathematical and estimation learning process expressive of the human social learning manner of finding various point of view before doing any major conclusion. Combining the idea of different "experts" to achieve general "ensemble" decision is basic. In the large borderline classification framework, the first step is needed to examine the ensembles that develop the borderlines, enhancing the conclusion skill of Output result and boosting-based ensemble algorithms. The second step is depended on the labeled bias values decomposition of the error, and it displays the ensembles can decrease values or both bias and values. All different classifiers will create the same error is the major advantage of ensemble methods. However, a minority of the classifiers make every mistake but optimal classification will be achieved. All different learning algorithms are highly connected so all algorithms are ready to similar types of mistakes. In particular, ensembles serve decrease the value of classifiers [16].

3.5.2 Decision Tree Algorithm

The decision tree is a famous supervised regression algorithm in SVM. The input data is derived as a hierarchical model and it is developed when the subspace is connected with a labeled class. This algorithm is realized for being easy to translate and easy to use. Because they facilitate an intuitive method of resolving challenging decision-making issues, they are widely employed in applications related to business, science, and healthcare. In the business domain, decision trees are utilized for various purposes such as standardizing staff behavior about client wants and facilitating high-value investment decisions. In the medical field, decision trees are employed to diagnose ailments and make decisions about individual or community treatment. [33].

There are nodes in the decision tree that make up a rooted tree. The root node has no incoming edges, and the tree is directed. Every other node has precisely one edge coming in. A node that has outward edges is referred to be an internal or test node.

"Leaf" refers to any other node (also known as terminal or decision nodes). In a decision tree, each internal node divides the branches into two or more sub branches based on a definite the input values. In deciding tree algorithm, choose one single input attribute and according to the attribute's value, the space is divided. In the case of numeric value, the situation refers to a class. Each leaf is allocated to one class means that the maximum suitable target value. Alternatively, the leaf retains a possibility vector preferable the possibility of the target attribute meeting a definite value. Samples are divided by operating them from the root of the tree down to a leaf, based on the result of the example value across the route. Figure 3.8 expresses a decision tree that define whether or not a potential customer will replay to a direct mailing. Internal nodes are described as circles, whereas leaves are represented as triangles. These decision tree integrated both nominal and numeric attributes.

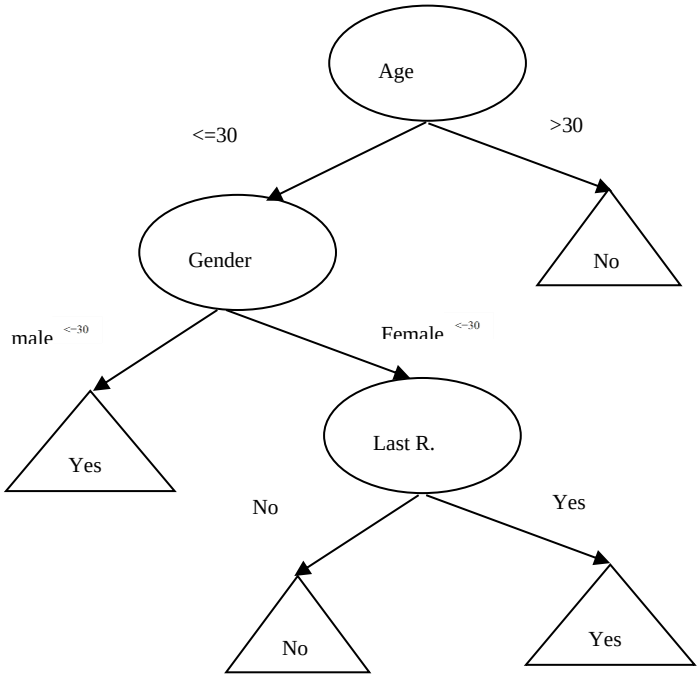


Figure 3.8 Decision Tree Presenting Response to Direct Mailing

Each directed edge of the tree can be translated to a Boolean expression (e.g., x1 > 5); So, a decision tree can simply be translated to a set of rules. Each path from root to leaf creates one rule as follows: form the conjunction (logical AND) of all the decisions from parent to child. Decision trees can be used with both ordered and unordered attributes [28]. Given this classifier, the interpreter can forecast the feedback of a potential customer, and estimated the detectable features of the entire possible consumer population concerning direct mailing. Each node is categorized with the

attribute it evaluates, and its branches are categorized with its corresponding values [61].

3.5.3 Neural Network

Artificial Neural Networks (ANNs) are motivated to biological neural networks forming the brain. Integrating with mathematical theories, ANNs are well applicable for linear programming, arithmetic and logic calculation such as image recognition and matching, classification and clustering. The first ANN was very simple and consist of some neural connections. A supervised learning system of artificial neural network decreases the errors between actual and expected outcome. A training data set predicts the expected outcome. This difference is back-propagated across the whole network and permits the updated weights for each nodes. Figure 3.9 describes symbolically how the perceptron accepts the inputs x_1, x_2, …, x_m and the upper left box with symbol "1" describes a bias for the input data. These are combined with the weights w_i in order. To calculate the Net input function, the input data are passed on to the activation function that creates a binary output -1 or +1 that correlate to the predicted class label of the example data. To compute the error, the output is used in the learning phase. To improve the weights for decreasing the error between the actual and the desired outcome is called backpropagation.

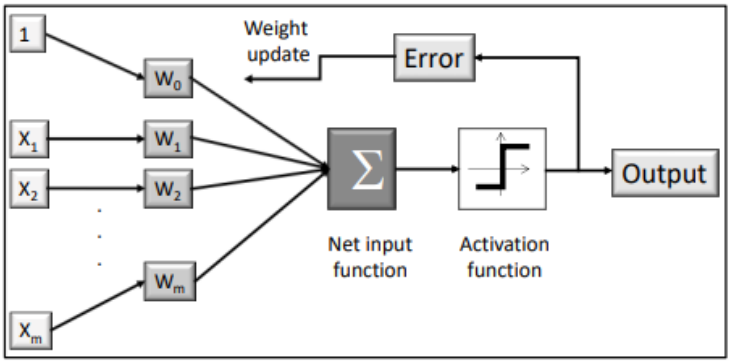


Figure 3.9 The Perceptron Work Flow

The Perceptron's algorithm can be abstracted in the following steps:

1. Load the weights to 0 or to small random numbers.
2. Arrange a training data set and use it as input for the Perceptron.
3. Calculate the output value, y_{out} using the unit step function.
4. Compare the output y_{out} with the desired output y_{true} (the true class label for the selected training sample).

5. Use the error (the difference between y_{true} and y_{out}) for updating the weights.
6. Repeat until the error is below a certain desired threshold.

These steps can be constructed as following. The unit step function produces the outcome result that is the class label. The updated weight w_j in the weight vector w can be generally described as:

$$w_j = w_j + \Delta w_j$$

The update value

$$\Delta w_j \text{ is } \Delta w_j = \eta(y_{true}(i) - y_{out}(i))xj(i) \text{ A5-4} \quad (3.3)$$

The equation (3.3) is generally described as the Perceptron learning rule. It permits the updated weights to reduce the error. However, the merging is only approved if the two classes are linearly separable and if the "learning rate" η is acceptably small. This is easy to create in two dimensions. These two sets are linearly separable if there exists at least one line in the plane with all of the circles on one side of the line and all the triangles points on the other side. In greater dimensional Euclidean spaces, the same issue is approximately if the line is substituted by hyperplane. Its permanent restriction to linearly separable classes described the main reason why the Perceptron reach serious analysis [45].

The activation function, denoted by φ(v) means the activation function and it is the outcome of a neuron in terms of the induced local field v. Two basic types of activation function are Threshold Function and Sigmoid Function. In Threshold Function, the outcome of neuron n using such a threshold function is defined as

$$y_{n=}\begin{cases}1 \; if \; v_n \geq 0 \\ 0 \; if \; v_n < 0\end{cases}$$

where v_n is the induced local field of the neuron

$$v_n = \sum_{j=1}^{m} w_{nj}x_j + b_n \quad (3.4)$$

The another basic activation function used in the neural network is Sigmoid function and whose graph is "S" shaped. It is described as exactly improving function that expresses an elegant stability between linear and nonlinear nature. The logistic function is an example of the sigmoid function,

$$\varphi(v) = \frac{1}{1+\exp(-av)} \tag{3.5}$$

where a is the slope value of the sigmoid function. This function comes simply a threshold function. Whereas a threshold function assumes the value of 0 or 1, a sigmoid function assumes a continuous range of values from 0 to 1 [11].

3.6 Unsupervised Learning

Unsupervised learning does not need to have label class and it defines based on their own input structure [6]. This learning seeks concealed structure or permanent pattern in data. It is applied to sketch conclusion from datasets consisting of unlabeled input data [19]. It is hard to imagine data set in more than two dimensions, and most data is in hundreds of dimensions. Dimensionality decreasing is the issue of having great dimensional data and embedding it in a lower dimension area. Another thing that might want to do is automatically derive a partitioning of the data into clusters. The k-means algorithm is unsupervised clustering algorithm and it is automatically partition into the class. Gene sequence analysis, market research, and object recognition is an example of unsupervised clustering application. Figure 3.10 describe the feature of unsupervised learning structure.

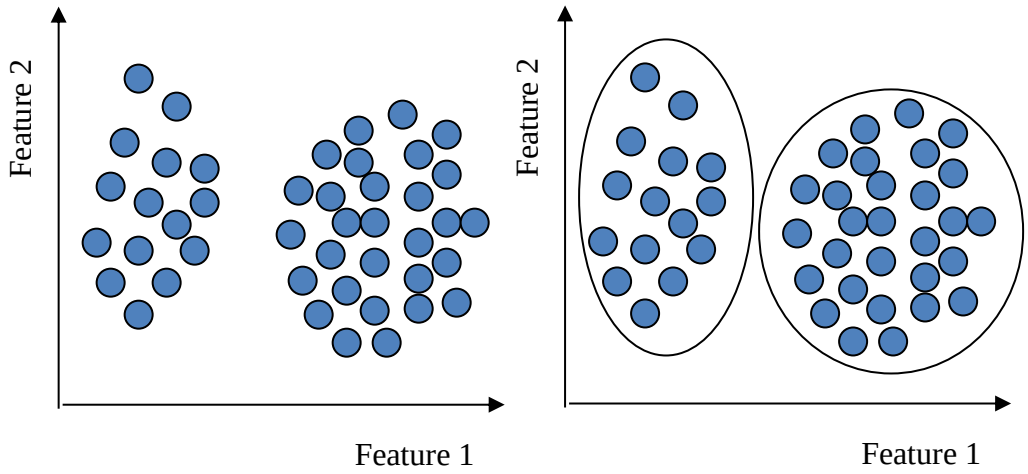


Figure 3.10 Unsupervised Learning

3.6.1 K means

Data clustering is frequently used in many applications, such as data mining, image recognition, decision support system, machine learning and data segmentation. K-means clustering is unsupervised clustering system and no labeled input data doesn't need. It repeatedly detects the k centroids and allocates each data to the nearest centroid, where the coordinate of each centroid is the mean of the coordinates of the objects in

the cluster. Sadly, K-means clustering algorithm is recognized to be subject to the initial cluster centers and simple to get stuck to the internal optimal conclusion. Additionally, when the data set is enormous, it takes long time to detect the conclusion. Simple K-means algorithm works as follows:

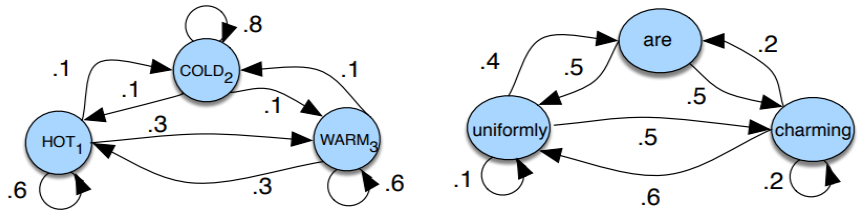
1. Initialize k (random) data set (seeds) to be the initial centroids, cluster centers
2. Locate each data set to the closest k centroids
3. Re-calculate the k centroids by using the current cluster data
4. If a convergence criterion is not met, repeat steps 2 and 3

Different methods have been developed to improve the performance of the K-means algorithm. There are many advanced K-means clustering algorithm. Some researchers developed to improve K means in various ways such as randomly divide input data set into 10 subsets. They choose randomly same initial seeds to all 10 subsets. The outcome of the 10 computes is 10K center points. These 10K points are then themselves input to the K-means algorithm and the algorithm run 10 times, each of the 10 runs initialized using the K final centroid locations from one of the 10 subset runs. The resulting K center locations from this run are used to initialize the K-means algorithm for the entire dataset. Huang [7] and Sun et al [8] developed the K-means prototype to cluster into categorized data. Strehl and Ghosh [9] proposed to collect many clusters data set into a single combined clustering without achieving the features or algorithms that solved these partitions. Likas et al [10] developed the global K-means algorithm (The GKM algorithm), which is an incremental way to clustering that dynamically adds one cluster center at a time through a deterministic global search procedure consisting of N executions of the K-means algorithm from suitable initial positions [8].

3.6.2 Hidden Markov Model

The Hidden Markov Model is based on increasing the Markov chain. A Markov chain states the probabilities of sequences of random variables, states, each of which can draw on values from some set. These sets can be words, or tags, or symbols describing anything, like the weather. A Markov chain causes a very powerful acceptance to forecast the next step in the sequence, based on the current state. The states before the current state have no impact on the next step except via the current state. It is likely to predict tomorrow's weather is based on today but don't need to consider yesterday's weather [12]. A Markov chain is helpful to calculate a probability

for a sequence of obvious occurrence. In many examples, the events are hidden and can't simply discover part-of-speech tags in a text. The nature of Markov chain is shown in Figure 3.11.



(a) (b)

Figure 3.11 A Markov chain for weather (a) and one for words (b) showing states and transitions

Figure 3.12 describes a generalized automate architecture of an operating HMM λi with the two integrated stochastic procedures.

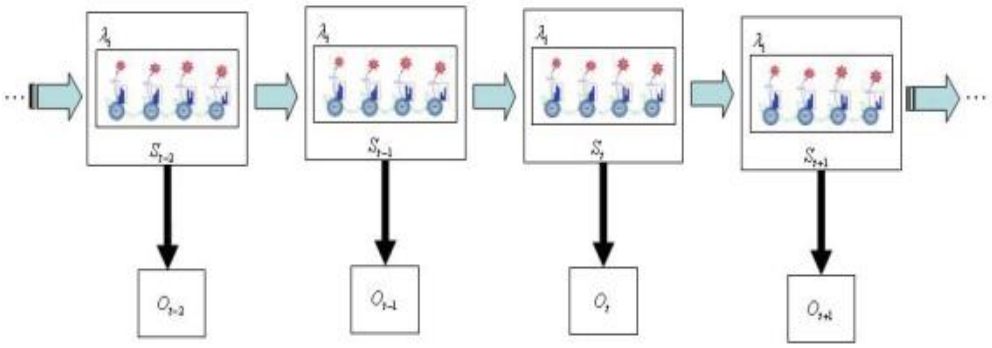


Figure 3.12 Generalized Architecture of an operating Hidden Markov Model

Each curve describes a random value that can accept any values. The random variable s(t) is the hidden state at time t. The random variable o(t) is the examination at the time t. The rule of provisional probability of the HMM value s(t) at the time t, expressing the values of the hidden variables at all times based on the value of the hidden variable s(t-1) at the time t-1. By the second stochastic model, the value of the observed variable o(t) based on the value of the hidden variable s(t) also at the time t [28]. A hidden Markov model (HMM) allows us to talk about both observed events Hidden Markov model (like words that we see in the input) and hidden events (like part-of-speech tags) that we think of as causal factors in our probabilistic model. The sequence of HMM is described as the following steps.

$R = r_1 r_2 \ldots r_N$ a set of N states

$M = m_{11} \ldots m_{ij} \ldots m_{NN}$	a transition possibility matrix M, each a_{ij} describing the possibility of changing from state i to state j, s.t. $\sum_{j=1}^{N} m_{ij} = 1 \; \forall i$
$E = e_1 e_2 \ldots e_T$	a sequence of T examination, each one gathered from a vocabulary $C = c_1, c_2, \ldots, c_V$
$L = L_i(e_t)$	a sequence of examination likelihoods, also called emission possibility, each describing the possibility of an examination e_t being produced from a state i
$\pi = \pi_1, \pi_2, \ldots, \pi_N$	an initial possibility distribution over states. π_i is the possibility that the Markov chain will begin in state i. Some states j may have $\pi_j = 0$, describing that they cannot be initial states.

3.7 Semi-Supervised Learning

Semi-supervised learning is a classification technique for the application that input data consist of a small amount of labeled data and large amount of unlabeled data. Semi-supervised learning drops between unsupervised learning and supervised learning. Many machine-learning researchers have found that unlabeled data, when used in conjunction with a small amount of labeled data, can produce considerable improvement in learning accuracy. The addition of labeled data for a learning issue often needs a skilled human or a physical experiment. The charge connected with the labeling procedure thus may deliver a fully labeled training set unusable, whereas collection of unlabeled data is nearly cheap. In such condition, semi supervised learning can be of great practical value. Semi supervised learning is also of theoretical interest in machine learning and as a model for human learning [49].

3.8 Regression Analysis

Regression analysis is a type of forecasting model which examination the connection between a **dependent** and **independent variable**. This method is used for prediction, time series modelling and discovering the connection between the variables. The aim of regression analysis is to describe the reply value as a consequence of the predictor variables. The including of fit and the accuracy of result based on the input data. Therefore, for powerful use of regression analysis one must contain:

1. Examine the input data collection process,
2. Find any limitations in collected data
3. Limit consequences accordingly.

Once a regression analysis connection is achieved, it can be applied to forecast values of the response variable, recognize variables that most impact the reaction, or confirm hypothesized causal prototype of the reaction. The value of each predictor variable can be evaluated through statistical trials on the predicted coefficients of the predictor variables [38]. There are many advantages of applying regression analysis. They are as follows:

1. It illustrates the prominent connection between target and predictor.
2. It illustrates the powerful effect of multiple predictor based on a dependent variable.

Regression analysis also permits to contrast the impacts of variables calculated on different variables, such as the impact of price changes and the number of promotional activities. These advantages assist data scientists, market researcher and data analysts to decrease and calculate the efficient variables to build the predictive models.

3.8.1 Type of Regression Model

There are three types of regression model as shown in Figure 3.13. These are linear regression, multilinear regression and nonlinear regression models. Linear and multiple linear are basic type of model and these are easy and simple. Nonlinear regression analysis is used in sophisticated application and its connection exists between the target and predictors.

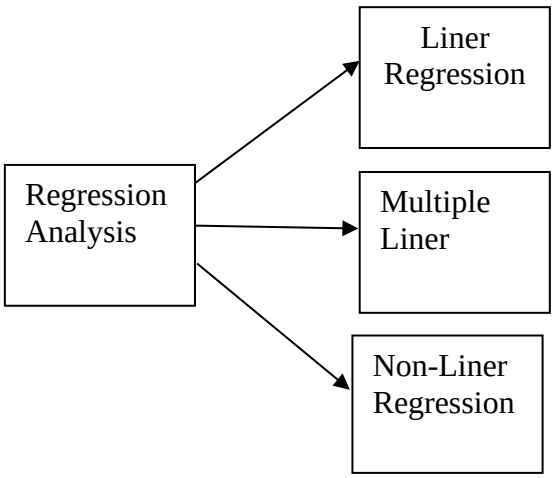


Figure 3.13 Type of Regression Model

3.8.1.1 Linear Regression Model

In practical applications, linear regression is utilized to create and measure the relationship between variables using statistical input data. The line of greatest fit for the relationship between the z target value and the i predictor value is described by the mathematical equation z = ni + b. The goal value is Z, and the informative value is I. The intercept is b, and the slope of the line is n. The regression coefficient, or r2, shows how much z is influenced by i. A line with advanced features is referred to as a hyper-plane when it has several inputs (i). The complexity of a regression model such as linear regression is a common topic of discussion. This explains how many coefficients are used in the model. When a coefficient approaches zero, the input value's influence on the model and the prediction derived from the model is essentially reduced (0 * x = 0). The values of the coefficients utilized in a linear regression model are estimated in four different ways.

1. **Simple Linear Regression**

When estimating a coefficient with a single input, statistical variables like means, standard deviations, correlations, and covariance can be computed from the data. All of these data must be accessible to cross and compute statistics.

2. **Ordinary Least Squares**

When there have more than one input value, Ordinary Least Squares used to construct the values of the coefficients. These model finds to decrease the sum of the squared residuals. This means that given a regression line through the data we The input data point is used to compute the distance from each point to the regression line, square it, and sum all of the squared errors together. This model handles the value as a matrix and applies linear algebra equation to calculate the optimal values for the coefficients. This approach need huge memory location to construct the matrix and all of the input data must be accessible. It is simply like to call a procedure in a linear algebra library. This approach is very fast to compute.

3. **Gradient Descent**

When there are one or more inputs you can use a process of optimizing the values of the coefficients by iteratively minimizing the error of the model on your training data. To minimize the error, iterating one or more input training data are used to get the value of coefficient. This method is called Gradient Descent and select

randomize value used for each coefficient. The means square error is calculated for each input and output values. A learning rate is applied as a scale factor and to minimize the error, the coefficients are renewed. Before the changes are not occurred, this process is repeated to get a minimum sum squared error. In this method, the researcher must choose a learning rate value to determine the optimization of the process to use each iteration. Gradient descent is sometime trained using a linear regression model because it is easily to realize. In reality, this method is useful in the large dataset.

4. Regularization

The regularization approach is an updated linear model training procedure. This technique uses ordinary least squares to minimize both the sum of the squared errors of the model on the training data and the sum of the squared errors on the training dataset. There are two well-known instances of linear regression regularization processes:

1. Lasso Regression: this method updates Ordinary Least Squares to lower the absolute sum of the coefficients as well (referred to as L1 regularization).

2. Ridge Regression: In this method, also known as L2 regularization, Ordinary Least Squares are updated to lower the squared absolute sum of the coefficients. These techniques work well when the input data set exhibits collinearity, and simple least squares would overfit the training set.

3.8.1.2 Multi-Linear Regression Model

Multiple linear regression (MLR) is a statistical model that applies many explanatory variables to forecast the result of a dependent variable. The aim of MLR is to model the linear connection between the independent values and dependent values. Basically, multiple linear regression is the addition of original least-squares regression that consist more than one expository value defined as P and the connection between the dependent variable and the expository value is described by the following equation:

$$y_i=\beta_0+\beta_1 x_i^1+\beta_2 x_i^2+\ldots+\beta_p x_i^p+\epsilon \quad (3.6)$$

i=n surveying

y_i=dependent variable

$x_{i=}$ expository value

β_0=y-interception point (constant value)

β_p=slope coefficients for each expository value

ϵ= error value (the residuals)

Some assumption need to consider to construct multiple regression model.

The first one is: There must be linear connection between the dependent values and independent values. The second one is: The independent values are not too highly interconnected with each other. The third one is; y_i surveying is chosen independently and erratically from the population. The last one is: Error value must be the mean value is 0 and variance represented as σ. The coefficient R-squared vale is a mathematical metric that is applied to compute how much of the changes in result can be determined by the changes in the independent values. R^2 always develops as more forecaster are combined to the MLR model after all the forecasters may not be connected to the result value. R^2 by itself cannot be applied to point out which forecasters should be involved in a method and which must be reject. R^2 value must be between 0 and 1, where 0 defines that the result cannot be forecasted by one of the independent values and 1 defines that the result can be forecast without error from the independent values. Using a multiple linear regression model has three benefits. Initially, it could be applied to identify the extent of the influence an independent value has on a dependent variable. It can also be used to predict the consequences or repercussions of changes. In other words, multiple linear regression analysis aids in our comprehension of the degree to which changes in the independent variables affect the dependent variable. For example, you can use a multiple linear regression to find the expected gain or decrease in GPA for each one-point change in IQ. Third, future values and trends are predicted by multiple linear regression analysis. One method for obtaining point estimates is to employ multiple linear regression analysis.

Another major consideration is the model fit. Combining independent values to a MLR model will regularly develop the number of confirmed changes in the dependent values. So, combining extremely independent values without any theoretical confirmation may outcome in an over-fit model.

3.8.1.3 Nonlinear Regression Prediction Model

Using a basic linear relationship to represent the shifting and trend of a time series is very awkward and inaccurate because of the complexity and variety of real-world data. When one or more independent variables and model parameters are modeled as a non-linear function, the dependent or criterion variables are said to be in a nonlinear regression. Nonlinear regression is typically employed in place of hypothesis testing for estimating the parameters of a nonlinear model. It is not necessary to make the standard assumption on the residuals' normality in this instance. Rather, the essential premise that must be made is that the data might be accurately represented by the model. [10].

In real world application, simple linear prediction model can't resolve all application because of the trend of a time series and some nature of application is complex. In this situation, nonlinear regression prediction model is used for the dependent values are constructed as nonlinear function. Nonlinear regression model is normally used to construct the parameters without executing hypothesis tests.

To fit a nonlinear regression model is using quadratic or higher order trend is the simplest way and it can be specified as x1,t=t,x2,t=t2,….x1,t=t,x2,t=t2. But quadratic or higher order trends is not suitable for the application of the forecasting and it is not suggested. The results in forecasting application are sometime unrealistic.

A better attitude is to apply the piecewise definition introduced above and suit a piecewise linear movement which curves at some point in time. So, it can be constructed for a nonlinear curve using linear piecewise. If the trend curves at time τ, then it can be defined by simple replacing x=tx=t and c=τc=τ for above definition such that the predictors are defined as,

$$x1,t=tx2,t=(t-\tau)+=\{0t<\tau(t-\tau)t\geq\tau x1,t=tx2,t=(t-\tau)+=\{0t<\tau(t-\tau)t\geq\tau \quad (3.7)$$

in the model. If the connected coefficients of x1,tx1,t and x2,tx2,t are β1β1 and β2β2, then β1β1 permits the slope of the curve before time ττ, while the slope of the line after time ττ is offered by β1+β2β1+β2. Another curves can be combined in the connection by adding another values of the form (t−τ) + (t−τ) + where ττ is the "knot" or point in time at which the line should curved.

Nonlinear regression is any connection between an independent value X and a conditional value Y which outcomes in a non-linear function modelled data. Basically,

any connections that is not linear, can be defined as non-linear and is usually described by the polynomial of k degrees (maximum power of X). In fact, many distinct NLRs model survives that may be applied to fit anywhere the data set favors and these can continue and on to endless degrees. There are many types of non-linear regression model and the most popular functions are defined the below:

1. Cubic
2. Quadratic
3. Exponential
4. Logarithmic
5. Sigmoidal / Logistic

3.9 Uses of Regression Analysis

There are many application areas that regression analysis model is applied. Commonly, it can be distinguished as six groups. These are prediction model, specification area, parameter estimation, the strength of predictors, to forecast the result and to predict tendency forecasting. First of all, the equations of the regression analysis are planed only to create predictions. Second, to develop good predictions model, it should be not only the corrected specified model but also the accuracy of the parameter is defined. Moreover, these two-application areas such as accurate prediction and model specification want that all applicable variables be studied for all data and the equation be described in the accurate effective construction for all forecaster values. Third, the highest challenging is to build parameter estimation because not only the model is required to be accurately detected, the forecasting must also be accurate and the data should allow for better estimation. For example, multi collinearity expands a difficulty and wants that some estimators may not be applied. So, data limit and incapacity to calculate all predictor value correlate in a study limit the use of prediction equations [48]. Fourth, the cause of the independent values has on a dependent value, the regression can realize the power of the effect. Fifth, it can be applied to forecast results or effect of changes. That is, the regression analysis provides to realize how much the dependent value changes with a change in one or more independent values. Last, regression analysis predicts trend and prospective values. This analysis can be developed to gain point estimates.

3.10 Chapter Summary

This chapter introduced real-world applications using machine learning techniques. And also explains about seven steps of the machine learning system. This chapter aims to highlight the type of machine learning system in detail. The chapter also presents supervised learning, unsupervised learning, semi-supervised learning, regression techniques, and regression analysis in detail. Logistic regression is a type of nonlinear regression prediction system.

\

CHAPTER 4

THE ARCHITECTURE OF THE PROPOSED SYSTEM

Humans can manage the avoiding plans when they notice the information of cyclone's track and direction before it falls on land. Therefore, tropical cyclone position detection and cyclone movement estimation are the essential preventing procedures for people's lives and decreasing the small impact of the damage for the attacked regions. Many input data are used to predict the tropical cyclone track but mainly two types of input data are used such as statistical and satellite images of historical and in time. Many researchers have proposed the detection of cyclone position and track by using statistical data. The historical data are developed in many research areas such as biodiversity conservation, meteorology, geology, landscape, agriculture, forestry, regional planning, education, intelligence, weather prediction, warfare, and so on. In this chapter, a research paper "Feature Extraction and Tracking System for Tropical Cyclone" is proposed to prevent natural disasters and mitigate the loss of human lives and properties. Briefly, this chapter will describe with multiple sections to present the proposed detecting tropical cyclone and tracking system. The workflow of the proposed system is outlined. The suggested system operates in four basic steps as shown in Figure 4.1.

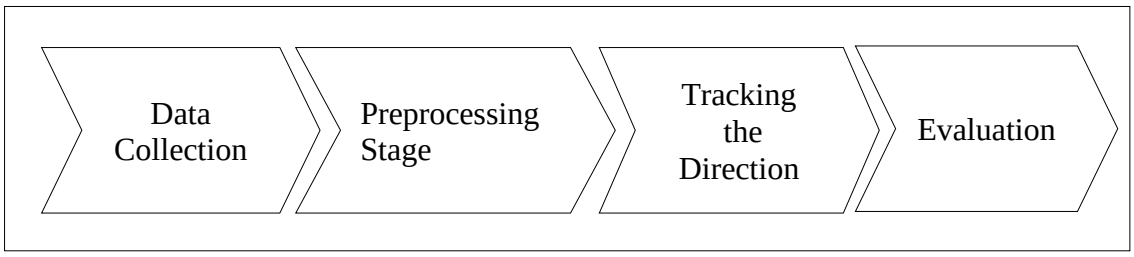


Figure 4.1 Proposed System Workflow

Data Collection Stage: The first stage is the collection of historical data on the North Indian Ocean basin year 1995 to 2022 from Joint Typhoon Warning Center (JTWC).

Preprocessing Stage: The second stage is the preprocessing stage which in order contains three phases: changing the latitude and longitude of the historical data as magnitude and direction by using Pythagoras theorem.

Tracking the Direction Stage: The third stage is to forecast the track of the cyclone direction with various threshold values of multiple logistic regression to detect the

cyclone tracking. Before detecting the position, calculate the highest correlation coefficient value between the tested cyclone and past cyclones. After that, seven features for direction and eight features for magnitude are extracted based on the deviation in time series structure.

Evaluation Stage: The assessment of the suggested paper recommendation system is the final phase based on the calculation of root mean square error, mean absolute error, and mean absolute percentage error to find the distance in kilometers between two points of Latitude and Longitude of ground truth data and predicted data. The detailed evaluations are discussed in Chapter 5.

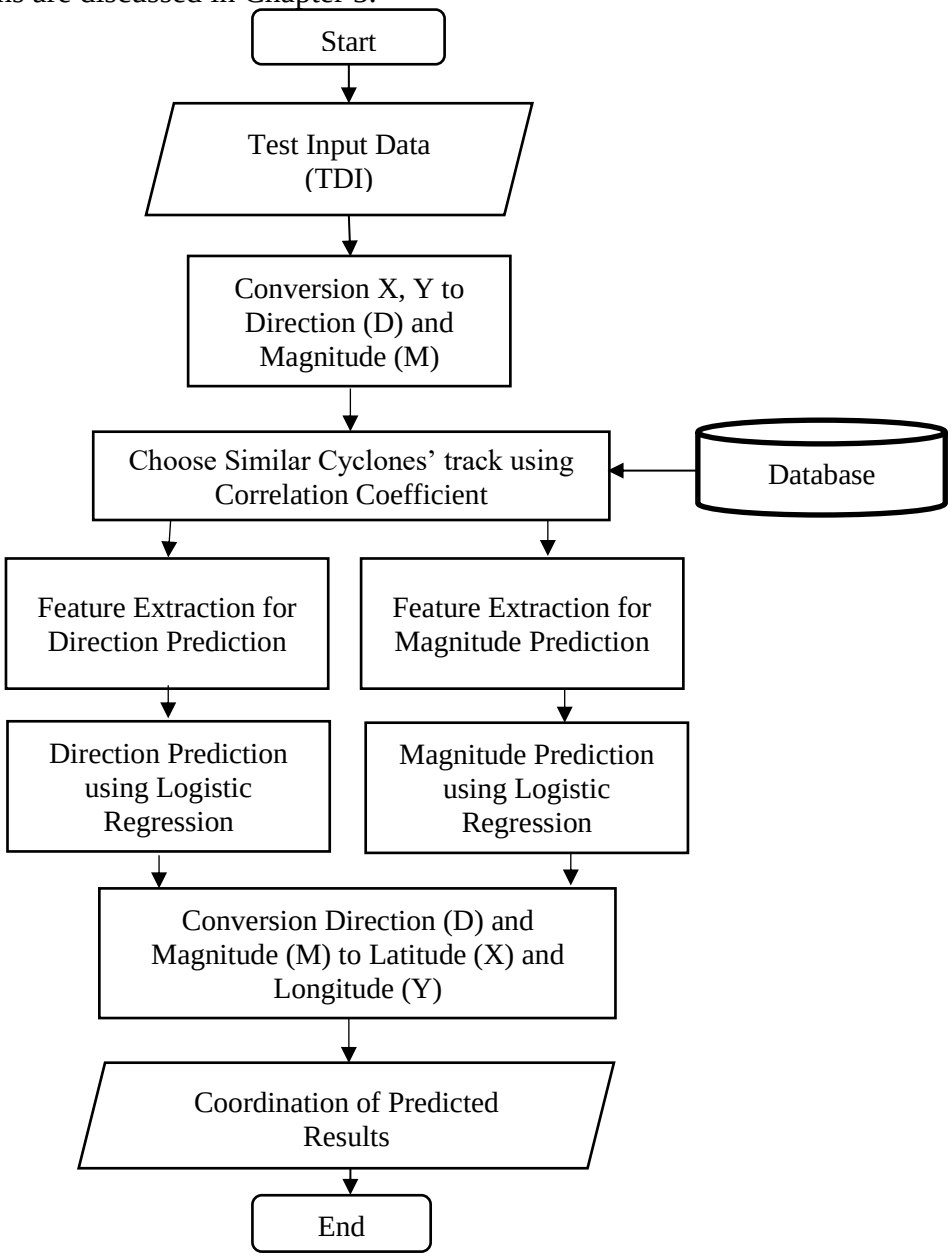


Figure 4.2 Overview of the Proposed System

Figure 4.2 depicts an overview of the proposed system, and the detailed steps are given in the following subsections. The latitude and longitude of all historical data and input test data are changed into direction and magnitude. The suggested approach mainly utilizes multiple logistic regression with pre-processed input data to generate useful features using mathematical equations. Cyclones occurred in the North Indian Ocean located from E40° to 100°, N 0° *t*o 30°.

4.1 Data Collection Stage

The data used in this research are trajectory and metrological factors. The 146 TC track data originate from the North Indian Ocean Best Track Data provided by the Joint Typhoon Warning Center [JTWC] between 1945 and 2022. In each historical data, there are six hourly tropical cyclone landing locations of 0.1 by 0.1 degrees and several meteorological factors, including the radius of the maximum winds (MRD), the minimum sea level pressure (MSLP), the level of tropical cyclone development (TY), the pressure in millibars of the last closed isobar (RADP), the maximum sustained wind speed (VMAX), the wind intensity (kts) for the radii (RAD), storm speed (SPEED), the eye diameter (EYE), storm direction (DIR), the maximum sea level (MAXSEAS), storm name (STORMNAME), system depth (Depth), wave height for the radii given in SEAS1-SEAS4, and radius code (SEASCODE). Meteorological factors were represented as one-dimensional points on the track at each timestamp.

Table 4.1 Data representation of cyclone Nargis

YYMMDDHH	Lat N/S	Lon E/W	VMAX	MSLP	STORMNAME
2008042512	105N	903E	20	1007	INVEST
2008042518	108N	895E	20	1007	INVEST
2008042600	107N	887E	20	1004	INVEST
2008042606	112N	885E	25	1002	NARGIS
2008042612	115N	879E	30	1000	NARGIS

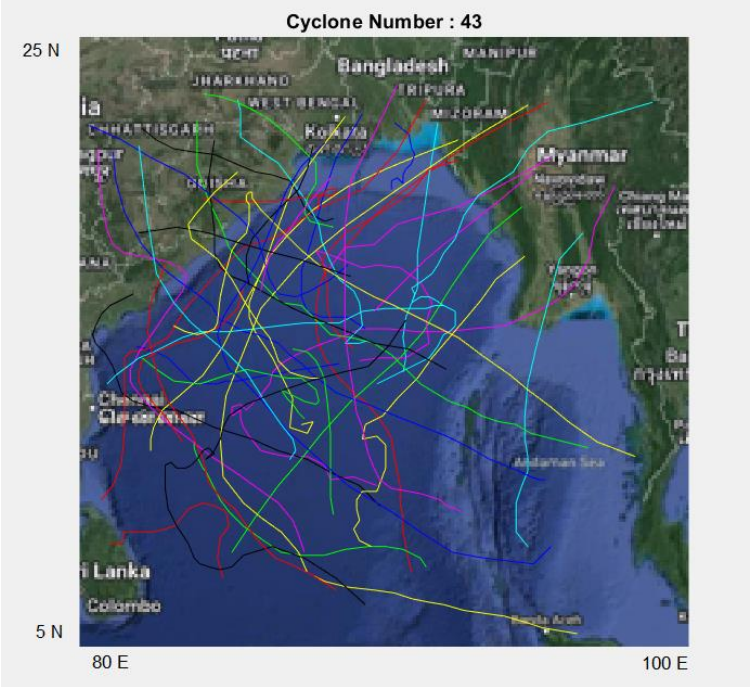


Figure 4.3 Cyclones' Tracks over the North Indian Ocean

Table 4.2 Tropical Cyclones' Climatology Elements and Persistence Elements in the Proposed System

Predictors	Description
X_0	Current longitude
Y_0	Current latitude
P_0	Current minimum sea level pressure
WS_0	Current maximum sustained wind speed
D	Direction
M	Magnitude
TDI	Test Data Input
D_1	Direction of correlated cyclone 1
D_2	Direction of correlated cyclone 2
M_1	Magnitude of correlated cyclone 1
M_2	Magnitude of correlated cyclone 2

4.2 Preprocessing Stage

Conversion of Latitude(Y) and Longitude(X) to direction and magnitude.

Direction (D) = θ

$$\tan\theta = \frac{Y}{X} \tag{4.1}$$

$$\theta = \tan^{-}\left(\frac{Y}{X}\right) \tag{4.2}$$

$$Magnitude(M)\sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2} \tag{4.3}$$

By using equation (4.2) latitude of the cyclone position is considered as Direction(D) and the longitude is transformed as Magnitude(M) by using equation (4.3).

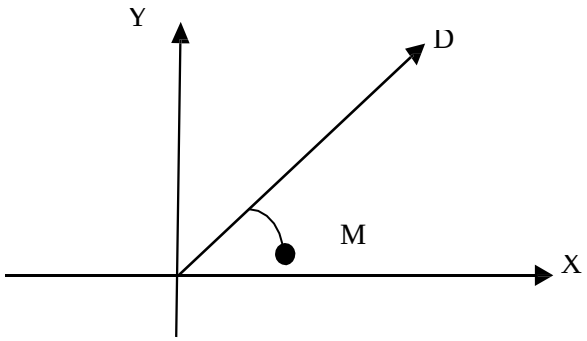


Figure 4.4 Direction and Magnitude

4.2.1 Selection of Cyclones with Similar Tracks

Similar cyclone tracks are selected by using the value of the correlation coefficient. This method is very simple and effected to reduce the most similar cyclones with the tested cyclone's track. One way to define correlation is the strength of the relationship between two variables [19]. For nominal and ordinal variables (as well as for time-series research), there are several measures of connection, and there are several correlation coefficients to address the unique properties of variables like dichotomies. To understand the notation, consider that the coefficient of determination is equal to the square of the correlation coefficient between x and y. That is the only true for a straight line [22].

$$r_{xy} = \pm\sqrt{R^2} = \sqrt{1 - \frac{S_{YY} - a_1{}^2 S_{XX}}{S_{YY}}} = a_1\sqrt{\frac{S_{XX}}{S_{YY}}} = \frac{S_{XY}}{\sqrt{S_{XX}S_{YY}}} \tag{4.4}$$

A popular statistical tool for describing straightforward relationships without demonstrating cause or effect, correlation is a measure of how much two variables are linearly related; correlations are also examined for statistical significance. The sample correlation coefficient, or r, indicates the stability of the connection. In this step, compute the correlation coefficient between the direction of test data input (TDI) and all historical data (HD). The two most correlated arrays are selected from all historical data. Pair the Direction of the test data input DTDI with directions D1, and D2, and the Magnitude of test data input MTDI with M1 and M2. The degree of link between two variables is known as correlation.

Figure 4.5 Tested Cyclone of Nargis with Two Correlated Cyclones

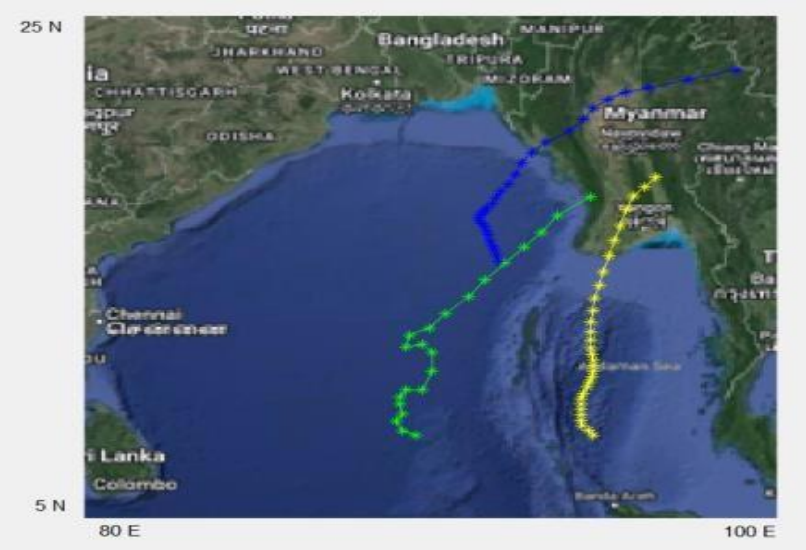


Figure 4.1 Tested Cyclone of Mala with Two Correlated Cyclones

4.2.2 Feature Extraction

Feature extraction is an important stage for this research. Depending on the input extracted features, there will be different accuracy results. In this paper, direction is based on the sea level pressure and magnitude may be changed upon maximum sea level pressure and minimum wind speed of tested data input (TDI).

Table 4.3 Seven Features for Direction

FE1	FE2	FE3	FE4	FE5	FE6	FE7(Target)
$D_{t_1-t}^{TDI}$	$D_{t_2-t_1}^{TDI}$	$D_{t_3-t_2}^{TDI}$	$P_{t_3}^{TDI}$	$D1_{t_3-t_2}$	$D2_{t_3-t_2}$	$D_{t_4}^{TDI}$

Table 4.4 Eight Features for Magnitude

FE1	FE2	FE3	FE4	FE5	FE6	FE7	FE8(Target)
$M_{t_1-t}^{TDI}$	$M_{t_2-t_1}^{TDI}$	$M_{t_3-t_2}^{TDI}$	$WS_{t_3}^{TDI}$	$P_{t_3}^{TDI}$	$M1_{t_3-t_2}$	$M2_{t_3-t_2}$	$M_{t_4}^{TDI}$

Direction Features

	Feature 1	Feature 2	Feature 3	Feature 4	Feature 5	Feature 6	Angle Dif
1	0	-41.6854	32.3824	0	6.0090	4.6896e-12	32.3824
2	-41.6854	32.3824	26.4082	-2	-3.4474	0	26.4082
3	32.3824	26.4082	-5.3509	-3	-4.6827	-17.7004	-5.3509
4	26.4082	-5.3509	-4.5696	-2	0.3649	-11.3542	-4.5696
5	-5.3509	-4.5696	-31.8746	-3	7.7652	2.9274e-12	-31.8746
6	-4.5696	-31.8746	-9.0589	-6	-5.9686e-13	-13.2824	-9.0589
7	-31.8746	-9.0589	-12.1004	-7	-21.0375	-2.6630	-12.1004
8	-9.0589	-12.1004	-41.4350	-11	-5.5275	1.2873	-41.4350
9	-12.1004	-41.4350	-23.9959	-4	-8.2594	11.2415	-23.9959
10	-41.4350	-23.9959	0.7418	1	-10.1755	4.5759e-12	0.7418
11	-23.9959	0.7418	13.2992	0	-2.7927	-22.8337	13.2992
12	0.7418	13.2992	-2.4966	4	-11.2435	-31.3287	-2.4966
13	13.2992	-2.4966	-22.6000	2	4.2917e-12	1.3785e-12	-22.6000
14	-2.4966	-22.6000	-10.1426	-2	-40.2364	-11.5932	-10.1426
15	-22.6000	-10.1426	-4.9349	0	-4.7636	-7.7187	-4.9349
16	-10.1426	-4.9349	-1.4747	-1	-4.6895	0	-1.4747

Figure 4.2 Seven Features for Direction

Magnitude Features

	Feature 1	Feature 2	Feature 3	Feature 4	Feature 5	Feature 6	Feature 7	Target Mag
1	-0.0481	-0.2768	0.1217	0	0	-0.0863	0.0377	0.6227
2	-0.2768	0.1217	0.0419	4	-2	0.0121	0	0.6645
3	0.1217	0.0419	-0.1417	1	-3	0.0207	0.0252	0.5228
4	0.0419	-0.1417	-0.1048	4	-2	-0.0073	0.0344	0.4180
5	-0.1417	-0.1048	0.1312	5	-3	-0.0868	3.8858e-15	0.5492
6	-0.1048	0.1312	-0.1264	8	-6	1.2629e-15	0.0200	0.4228
7	0.1312	-0.1264	-0.1063	10	-7	0.0088	0.0064	0.3165
8	-0.1264	-0.1063	-0.0118	16	-11	0.0057	0.0034	0.3046
9	-0.1063	-0.0118	0.1651	4	-4	-0.0050	0.0413	0.4698
10	-0.0118	0.1651	0.1185	-2	1	-0.0021	9.6589e-15	0.5883
11	0.1651	0.1185	-0.1022	0	0	0.0068	-0.0026	0.4861
12	0.1185	-0.1022	-0.0962	-5	4	0.0418	0.1061	0.3899
13	-0.1022	-0.0962	0.0908	-3	2	-5.1625e-15	-9.4369e-15	0.4807
14	-0.0962	0.0908	0.1589	3	-2	0.0554	-0.0462	0.6395
15	0.0908	0.1589	0.1336	0	0	0.0186	-0.0178	0.7731

Figure 4.3 Eight Features for Magnitude

4.3 Multiple Logistic Regression

A supervised machine learning technique called logistic regression is utilized for classification problems in which estimating the likelihood that an instance will belong to a particular class or not is the objective. A statistical method for analyzing the correlation between two variable components is called logistic regression. Logistic regression predicts the outcome of a categorical dependent variable. As a result, the outcome needs to be discrete or categorical. That can indicate Yes or No, 0 or 1, true or false, etc., but rather than providing a precise value between 0 and 1, it provides probabilistic values that fall in that In logistic regression, the sigmoid function fits a logistic function with an "S" shape, which predicts two maximum values instead of a regression line (0 or 1). When there are two result categories and numerous independent feature variables, multiple logistic regression is employed. For multiple logistic regression, the following equations are applied.

$$P(y^{(i)}=1 \mid x^{(i)};\theta) = \sigma(\theta_0 x_0^{(i)} + \theta_1 x_1^{(i)} + \ldots + \theta_n x_n^{(i)})$$

$$= \sigma(\theta^T x^{(i)}) = \frac{1}{1+e^{-\theta^T x^{(i)}}} \tag{4.5}$$

$$\theta^T = [\theta_0, \ldots, \theta_n]$$

$$x^{(i)} = \begin{bmatrix} x_0^{(i)} \\ .... \\ x_n^{(i)} \end{bmatrix}$$

i = ith training example

n = number of feature

$$J(\theta) = \frac{-1}{m}\sum_{i=1}^{m} y^{(i)} \log\left(\sigma\left(\theta^T x^{(i)}\right)\right) + \left(1 - y^{(i)}\right)\log\left(1 - \sigma\left(\theta^T x^{(i)}\right)\right) \quad (4.6)$$

$$\frac{\partial\left(J(\theta)\right)}{\partial\left(\theta_j\right)} = \frac{-1}{m}\sum_{i=1}^{m}\left(y^{(i)} - \theta^T x^{(i)}\right)x_j^{(i)} \quad (4.7)$$

$$\theta_j := \theta_j - \alpha\frac{\partial\left(J(\theta)\right)}{\partial\theta_j} \quad (4.8)$$

m = number of examples

i = an example

j = feature j

α = learning rate

The logistic regression model's output is converted into a probability using the non-linear sigmoid function. In this research, three types of thresholds are used in the experiential result.

$$sigmoid(x) = 1/(1+e^{(-x)}) \quad (4.9)$$

$$sigmoid(x) = 1/(1+e^{(-x*0.5)}) \quad (4.10)$$

$$sigmoid(x) = 1/(1+e^{(-x*1.5)}) \quad (4.11)$$

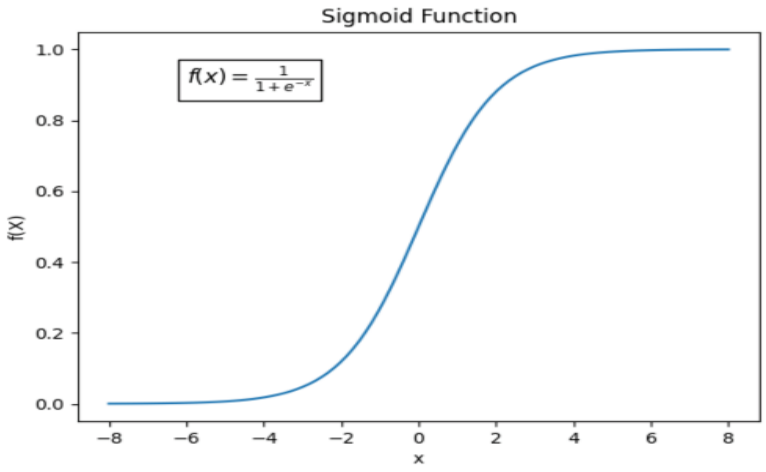


Figure 4.4 Sigmoid Function

4.4 Chapter Summary

This chapter explains how the suggested system is being implemented and explain the flow of the system step by step. Data collection stage, preprocessing stage and how to forecast the track of the cyclone by multiple logistic regression. Moreover, how to select the most correlated cyclones' track and extract the features are also explained. The following Chapter discusses the suggested system's experiments and evaluation findings.

CHAPTER 5

EXPERIMENTAL RESULTS AND EVALUATIONS

This chapter represented the experimental results of the proposed system. This system is a computer-based feature extraction and tracking system for tropical cyclones. An application is developed by MATLAB programming language with a Graphical User Interface (GUI). Users can choose two types of input data formats: Data Input Mode and Historical Data Mode for testing the track of cyclone. Moreover, the system evaluation is also described in this chapter. In Data Input Mode: Users must define the number of points of the cyclone location (latitude and longitude), the starting and ending value of wind speed and sea level pressure, maximum wind speed, and minimum sea level pressure. Changing magnitude and direction of all historical data are already preprocessed and saved in the database. By getting the input data, the system changes the latitude and longitude of the tested cyclone into direction and magnitude automatically. Users need to click the button "Find Correlation" to choose the highest similar cyclones to extract features by using the correlation coefficient value between the tested cyclone's data and historical data. Figure 5.1 illustrates the real-time Data Input Mode; the Input Point Number means the latitude and longitude of the cyclone position and the user needs to click on UI. The user must type the value of the starting point, ending point of wind speed, and maximum wind speed. Then also give the value of the starting point, ending point of sea level pressure, and minimum sea level pressure.

Figure 5.1 Input Data Mode

Figure 5.2 and Figure 5.3 illustrates the graph of wind speed and Sea Level Pressure.

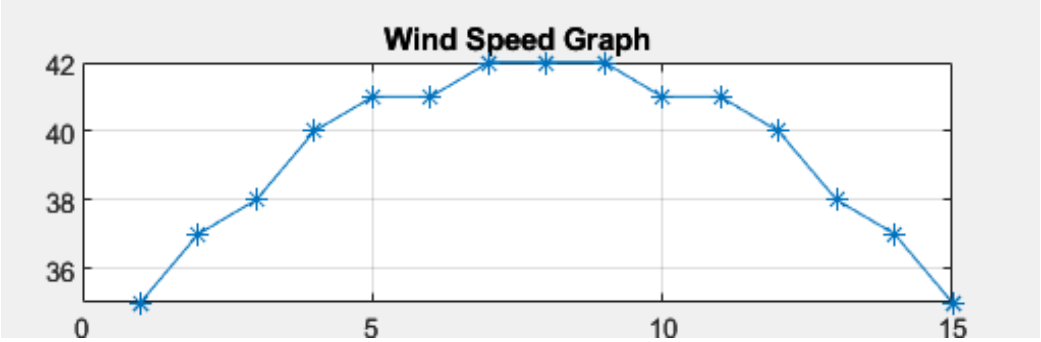


Figure 5.2 GUI of Wind Speed

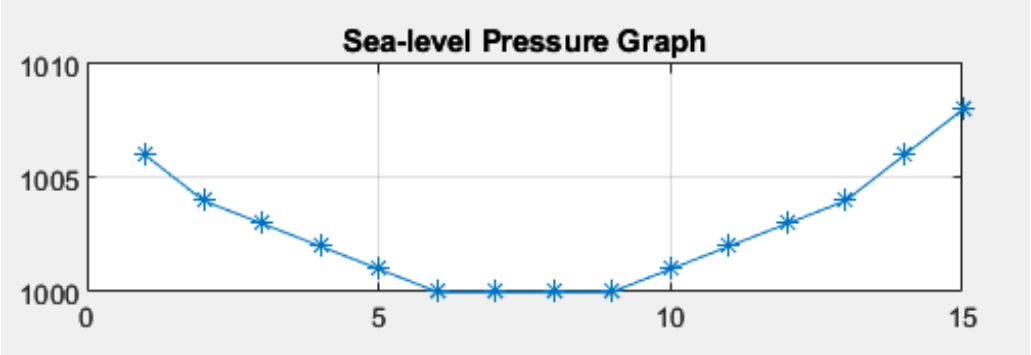


Figure 5.3 GUI of Sea Level Pressure

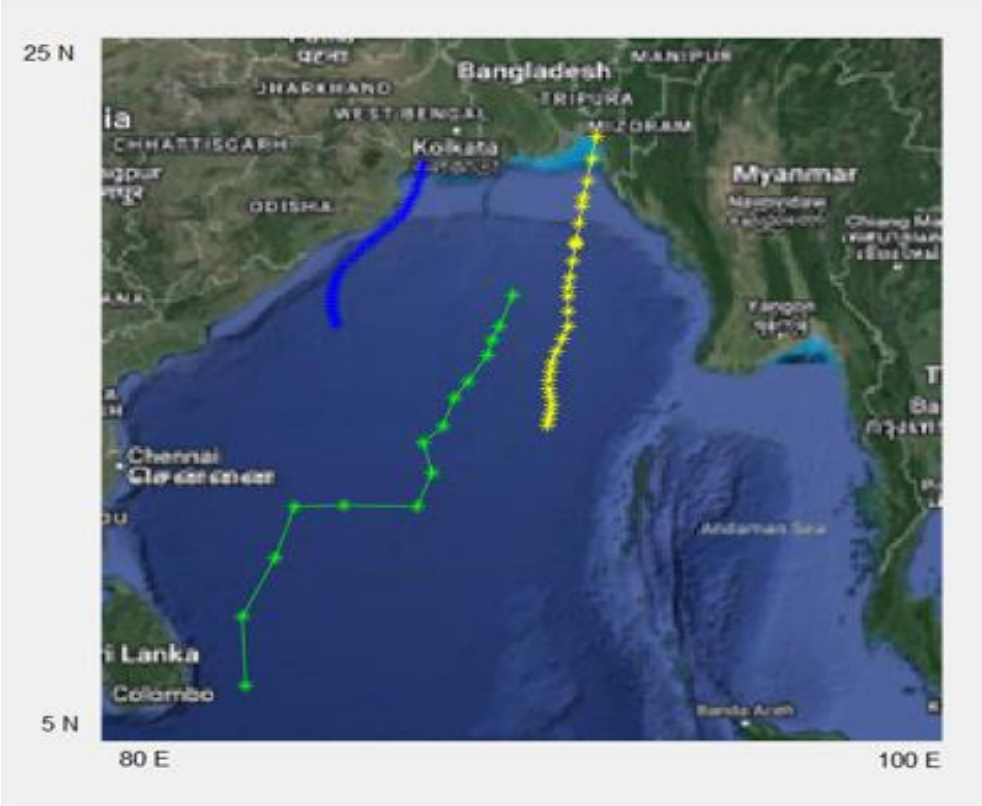


Figure 5.4 Tested Cyclone and Two Similar Cyclones

Figure 5.4 shows the tested cyclone in green color and two similar cyclones in blue and yellow color. And then extract the features for direction and magnitude as shown in Figure 5.5 and Figure 5.6.

Direction Features

	Feature 1	Feature 2	Feature 3	Feature 4	Feature 5	Feature 6	Angle Dif
1	8.5265e-13	-23.1770	-3.9891	-1	12.8160	-2.7792	-3.9891
2	-23.1770	-3.9891	5.0543	0	4.2633e-14	-8.5308	5.0543
3	-3.9891	5.0543	2.4368	-1	-8.8742	0	2.4368
4	5.0543	2.4368	-35.8067	0	-7.3860	0	-35.8067
5	2.4368	-35.8067	-34.8652	-1	4.2633e-14	-10.3048	-34.8652
6	-35.8067	-34.8652	-1.5303	0	-11.9414	-8.1301	-1.5303
7	-34.8652	-1.5303	0	-1	-1.7299	-4.9596e-12	0
8	-1.5303	0	10.2706	0	2.1476	4.5901e-12	10.2706
9	0	10.2706	59.2009	0	19.6538	-18.0939	59.2009
10	10.2706	59.2009	4.1339	0	0	-4.0724	4.1339
11	59.2009	4.1339	30.2254	0	0	-5.3291e-13	30.2254
12	4.1339	30.2254	-1.8900e-12	0	0	5.3291e-13	-1.8900e-12
13	30.2254	-1.8900e-12	-59.0362	0	0	-4.3987	-59.0362
14	-1.8900e-12	-59.0362	0	0	-8.7462	-3.0482e-12	0
15	-59.0362	0	23.0026	0	-5.2580e-13	7.6028e-13	23.0026
16	0	23.0026	2.8184	1	0	2.2879e-12	2.8184

Figure 5.5 Feature Extraction for Direction

Magnitude Features

	Feature 1	Feature 2	Feature 3	Feature 4	Feature 5	Feature 6	Feature 7	Target Mag
1	-1.2212e-15	-0.0599	0.0051	1	-1	0.0553	-0.0011	0.9049
2	-0.0599	0.0051	-0.1121	1	0	-1.7208e-15	-0.0015	0.7928
3	0.0051	-0.1121	-0.0428	1	-1	-0.0026	7.1054e-15	0.7500
4	-0.1121	-0.0428	-0.2194	0	0	0.0026	-3.5527e-15	0.5306
5	-0.0428	-0.2194	0.0451	1	-1	1.7208e-15	0.0157	0.5757
6	-0.2194	0.0451	0.1193	0	0	0.1098	0.0196	0.6950
7	0.0451	0.1193	0.1596	0	-1	0.0258	1.1380e-15	0.8546
8	0.1193	0.1596	-0.1833	0	0	0.0015	-4.4964e-15	0.6713
9	0.1596	-0.1833	-0.1693	1	0	0.0436	0.0553	0.5021
10	-0.1833	-0.1693	-0.0196	0	0	-3.5527e-15	0.0219	0.4825
11	-0.1693	-0.0196	-0.0511	0	0	-0.1600	-2.6923e-15	0.4314
12	-0.0196	-0.0511	-1.7208e-15	0	0	-0.2800	2.6923e-15	0.4314
13	-0.0511	-1.7208e-15	-0.1108	0	0	0	-0.0573	0.3206
14	-1.7208e-15	-0.1108	0	0	0	0.3069	1.0048e-14	0.3206
15	-0.1108	0	0.0851	0	0	0.1754	2.4980e-15	0.4057

Figure 5.6 Feature Extraction for Magnitude

The track of tropical cyclone is forecasted by using multiple logistic regression with three threshold Sigmoid functions as shown in Figure 5.7.

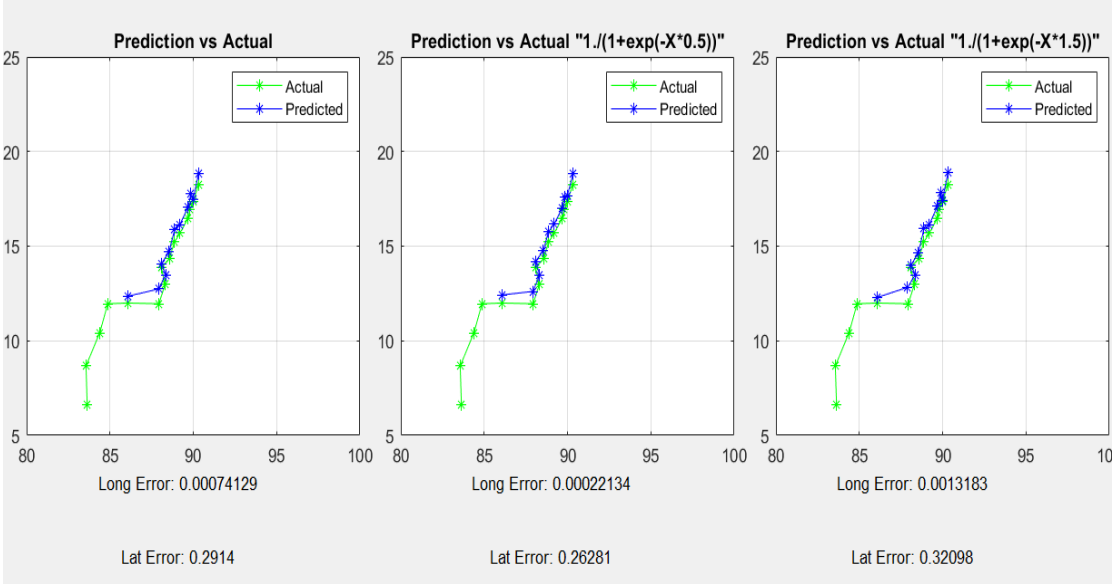


Figure 5.7 Compare the Cyclone Track by using Three Sigmoid Function

Error for latitude and longitude values are also described in the system. Another type of input data is very simple and easy for Users. Users can only choose the cyclone that occurred from 1945 to 2020 in the North Indian Ocean from the real-time datasets.

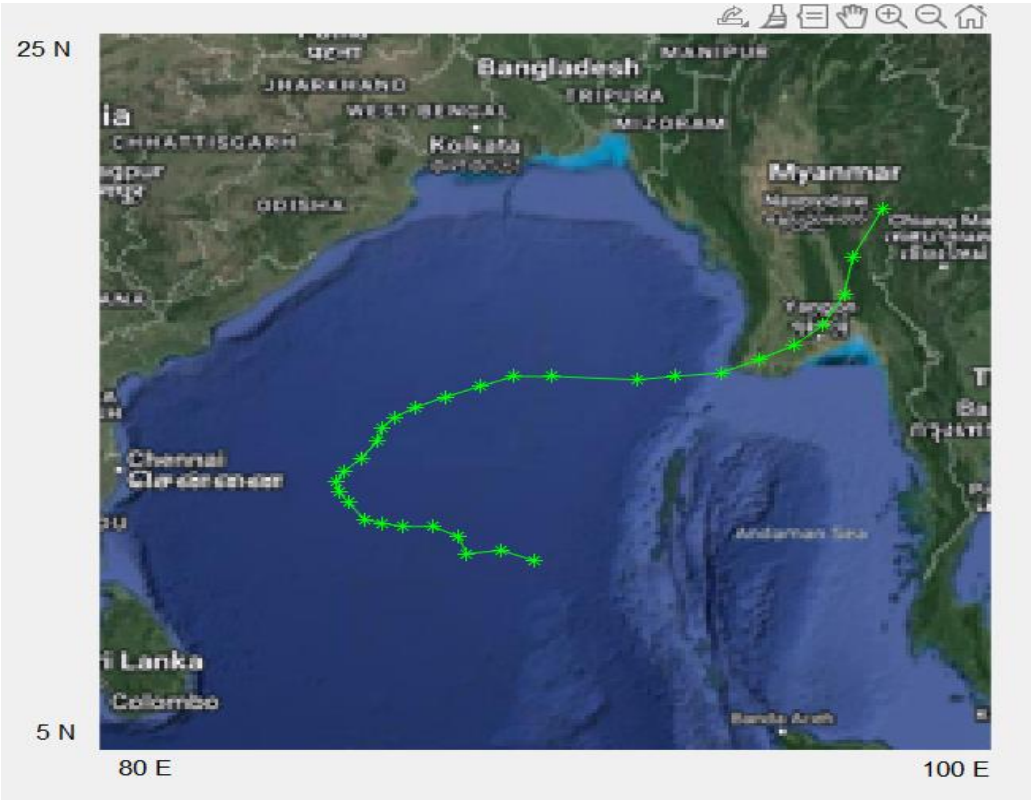


Figure 5.8 Nargis Cyclone from Choosing Historical Data

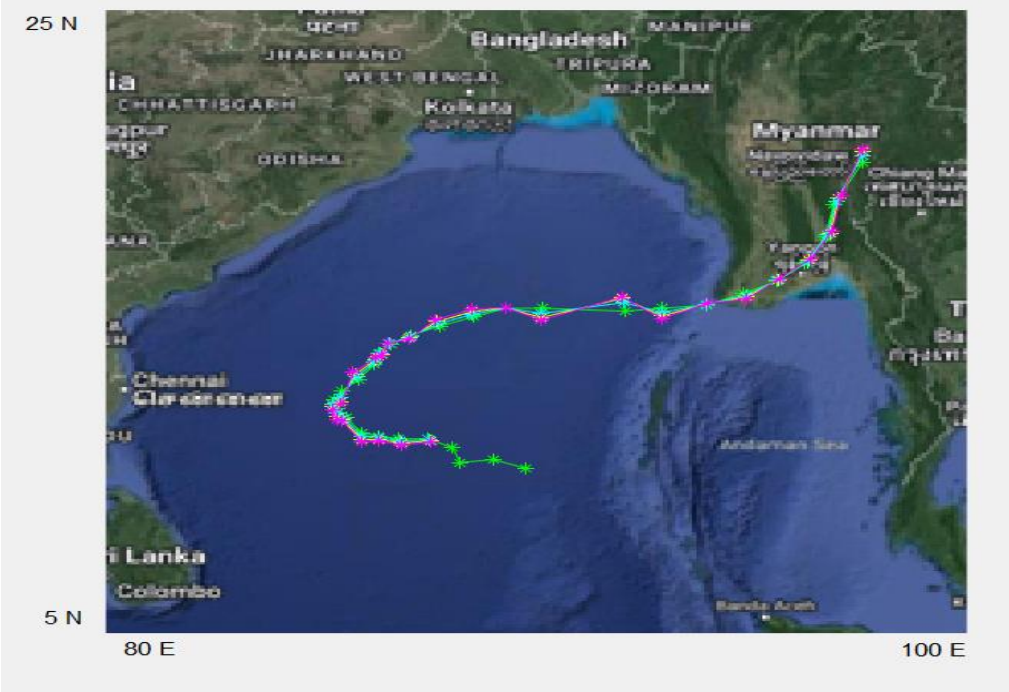


Figure 5.9 Forecast the track of Nargis Cyclone

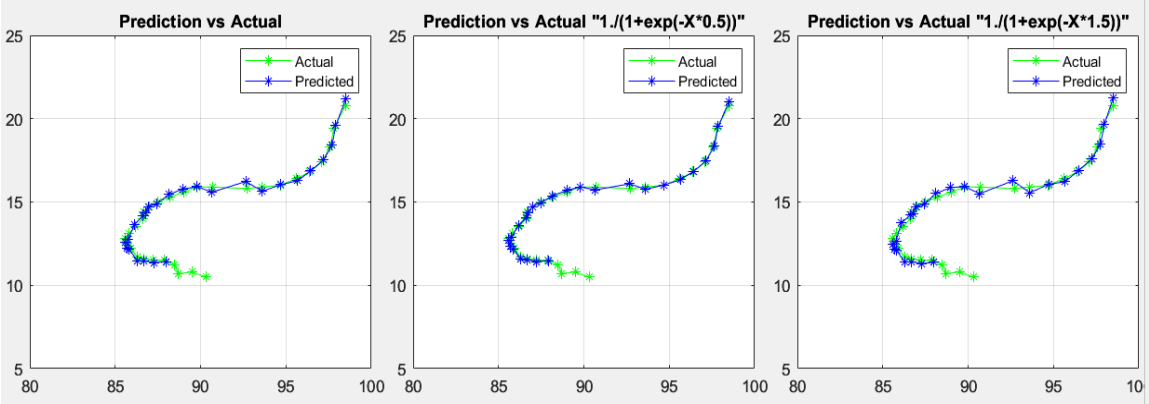


Figure 5.10 Predict Cyclone Nargis's Track by Using Three Threshold Functions

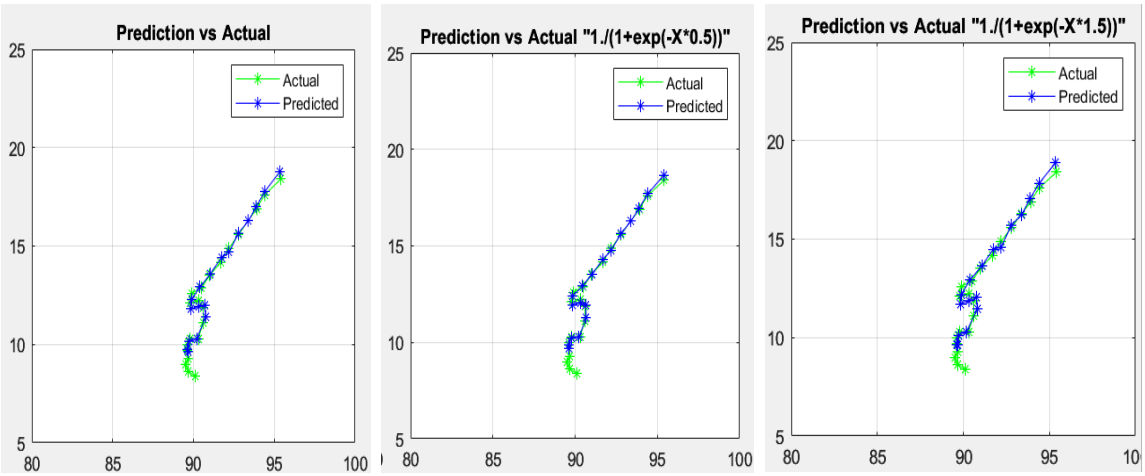


Figure 5.11 Predict Cyclone Mala's Track by Using Three Threshold Functions

5.1 Performance Measurement

To evaluate each method's performance between the forecast and actual tracks, the following metrics were selected: mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean squared error (RMSE). The MAE, which is determined by Equation (5.1), is the mean of all absolute deviations between all values that had been expected and values that were observed.. Poor model performance is shown by larger errors.

$$MAE = \frac{\sum_{i=1}^{h} |F_i - A_i|}{h} \tag{5.1}$$

where h is the number of historical data, Fi is the forecast value and Ai is the actual vale. The average error rate for actual values is denoted by MAPE. It considers the proportion of the error to the actual tangency as well as the error between the forecast and actual tracks. Equation (5.2) provides the computation. A perfect model is indicated by a MAPE value of 0, and a poor model is indicated by a number greater than 1.

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{F_i - A_i}{A_i} \right| \tag{5.2}$$

The square root of the difference between actual and forecast values as well as the number of data is known as the root mean square error, or RMSE. The following is the mathematical formulation.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (F_i - A_i)^2}{n}} \tag{5.3}$$

The experimental results were forecasted by using the past twelve-hour locations of the tropical storm that occurs in the North Indian Ocean. In this result, a threshold value 0.5 in using the sigmoid function is the best prediction result. MAE, MAPE, and RMSE are used to measure the performance of the proposed system.

The accuracy of the result also depends on the number of similar tropical cyclones' tracks. Two, three, and five similar tropical cyclones are tested with three threshold functions by using three measurement formulations as shown in Table 5.1 and Table 5.2. According to the result, three similar tropical cyclones are the best accuracy for this system and more similar cyclones can get less accuracy of the result.

Table 5.1 Prediction results of Cyclone Nargis using MAE, MAPE and RMSE

Sigmoid Function	Longitude			Latitude		
	MAE	MAPE	RMSE	MAE	MAPE	RMSE
Threshold 1	0.002757	0.00336	0.00397	0.003752	0.007548	0.001375
Threshold 0.5	0.002585	0.00336	0.003945	0.00358	0.007452	0.001348
Threshold 1.5	0.002627	0.00337	0.004949	0.003696	0.0076	0.001359

Table 5.2 Prediction results of Cyclone Nargis using Two, Three and Five Similar Cyclones Tracks

Similar Cyclones	Sigmoid Function	Longitude			Latitude		
		MAE	MAPE	RMSE	MAE	MAPE	RMSE
Two Similar Cyclones' Tracks	Threshold 1	0.002728	0.003624	0.004357	0.003948	0.007828	0.001994
	Threshold 0.5	0.002635	0.003442	0.00420	0.003729	0.007739	0.001639
	Threshold 1.5	0.002836	0.003555	0.005382	0.003886	0.008362	0.001561
Three Similar Cyclones' Tracks	Threshold 1	0.002757	0.003368	0.00397	0.003752	0.007548	0.001375
	Threshold 0.5	0.002585	0.003365	0.003945	0.00358	0.007452	0.001348
	Threshold 1.5	0.002627	0.003370	0.004949	0.003696	0.0076	0.001359
Five Similar Cyclones' Tracks	Threshold 1	0.003286	0.004027	0.005836	0.004481	0.008824	0.001949
	Threshold 0.5	0.003629	0.003820	0.005183	0.003259	0.008272	0.002107
	Threshold 1.5	0.003925	0.004464	0.005982	0.004028	0.008661	0.001849

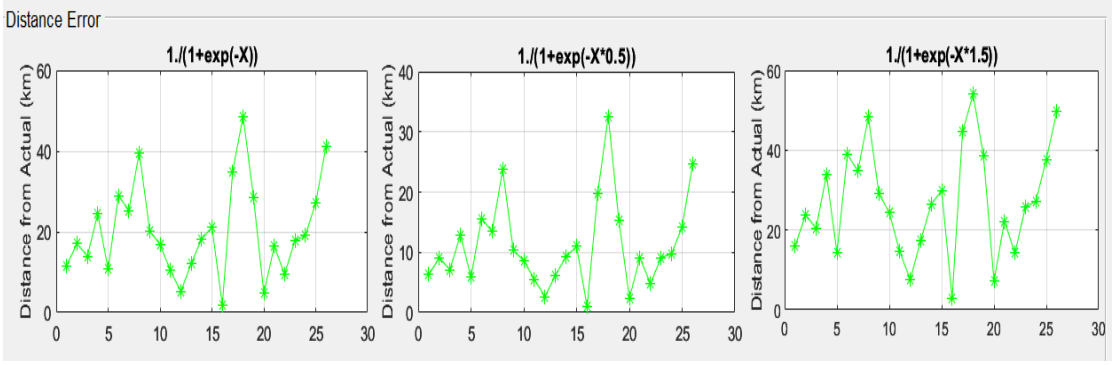


Figure 5.12 Graph of Distance Error using Three Threshold Functions (km)

Figure 5.12 illustrates the distance error between the historical data track and tested cyclone data in kilometers (km).

5.2 Chapter Summary

This chapter illustrates the experimental and evaluation of the system. Some results are expressed in the user interface (UI) and it is easy to use this application. According to the data of the evaluation results, three types of metrics are selected to test the accuracy of the result. Sigmoid functions with threshold values are tested and a threshold value 0.5 gets the best accuracy rate. Error is shown not only by the error of the latitude and longitude value but also by the distance from ground truth data and test data in kilometers (km). This proposed system is very easy for end users to predict the short-term tropical cyclones' track in real-time forecasting.

CHAPTER 6

CONCLUSION AND FUTURE WORK

This chapter presents three parts: (1) Conclusion of the research (2) Advantages and Limitation of the system (3) Further extension of the system. Conclusion of the system presents the whole system for feature extraction and tracking system for tropical cyclones. Some limitations are still challenging for the research. It needs to describe the problems of the system that can't be solved in this research. Prediction of the tropical cyclone track is the one piece of a weather forecasting system. Future extensions present the problems that should be solved in the future.

The tropical cyclone is one of the most destructive storms in a natural disaster. Natural disasters cannot be obstructed by human beings but we can prevent to decrease the loss of life and damages. Every year, people living in the coastal region suffer from tropical cyclones, flooding, and destructive wind. Statistical techniques that can automatically extract pertinent rules from massive amounts of data for purposes like detection, analysis, and prediction are the foundation of machine learning. In this research, predicting tropical cyclone tracks in the North Indian Ocean are a big challenge for the cyclone track. Historical data are obtained from the Joint Typhoon Warning Center (JTWC) and the previous 24-hour tropical cyclone track is predicted using simple multiple logistic regression. To reduce data redundancy and complexity of the model, the main contribution is considering the tested cyclone and correlated cyclones. The second idea is changing the location of the cyclone to consider direction and magnitude. Changing three threshold values of sigmoid function were used to compare the results. Another benefit of this proposed system provides a quick response within a few seconds when a new tropical cyclone track occurs in the ocean for short-term prediction.

6.1 Advantages and Limitations

The proposed system increases many benefits such as quality, accuracy, and efficiency of the proper tropical cyclone track predicting system. The first advantage is the processing time, this system only takes a few seconds to predict the real-time cyclone's track. The second is the accuracy of the result, errors are tested by three

matrix such as RMSE, MAE, and MAPE in distance kilometer (km). This error rate is acceptable for this system by comparing other deep learning techniques.

Existing research on machine learning usually only offers short-lead-time predictions, or when it comes to long-lead-time predictions, the accuracy falls short of expectations. Shortly, one bottleneck that needs to be addressed is how to create predictions that are longer and more accurate. Because TCs are extreme weather occurrences, most machine learning-based TC forecasting models now in use require supervised learning approaches and cannot be used as labels directly because they cannot be quantitatively defined in the real world. How to build training datasets and label data appropriately so that machine learning models can be trained to accomplish predicted objectives.

6.2 Future Extension

For a very long time, cyclone prediction has been difficult in weather forecasting. Some features such as wind pressure, wind direction, surface storm surge, storm direction and should be considered. Infrared satellite images can also be used as input data. Determining the intensity of the tropical cyclone can be extended for this research. A tropical cyclone track lasting no more than 24 hours could be predicted by the proposed system, not a longer duration. To make the research better in the future, more datasets will be added to predict more significant variables over longer periods and incorporate tropical cyclone intensities into the deep learning model. Mobile applications for weather forecasting systems can also be extended and it is more familiar for the mobile user. Sending SMS or making an alarm system to the end user when a cyclone occurs is also a prevention system for natural disasters.

\

AUTHOR'S PUBLICATIONS

[P1] Thu Zar Hsan, Myint Myint Sein, "Detecting Tropical Cyclone Using Infrared Satellite Images", 17th International Conference on Computer Applications (ICCA), Yangon, Myanmar, 27th-28th February, 2019. **Pg.** 83-89.

[P2] Thu Zar Hsan, Myint Myint Sein, "Combining Support Vector Machine and Polynomial Regressing to Predict Tropical Cyclone Tracking", IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech 2021), Nara, JAPAN, **Pg.** 220-221, 09-11 March 2021. **DOI:** 10.1109/LifeTech52111.2021.9391780.

[P3] Thu Zar Hsan, Thin Lai Lai Thein, "Feature Extraction and Tropical Cyclone Prediction System Using Correlation Coefficient and Logistic Regression", in Indian Journal of Computer Science and Engineering, Volume 8, Number 6, pp. 11-23, June 2024, (Q4).

BIBLIOGRAPHY

[1] W. Alomoush, A. Alrosan, N. Norwawi, Y. Alomari, D. Albashish, A. Almomani and M. Alqathani, "A Survey: Challenges of Image Segmentation Based Fuzzy C-Means", Journal of Theoretical and Applied Information Technology, Volume. 96, No. 16, pp. 5153- 5170, August 2018.

[2] MM. Ali, C. Kishtawal , "Predicting cyclone tracks in the north Indian Ocean: An artificial neural network approach", Geophysical Research Letters, Vol, 34, L04603, 2007.

[3] S. Alemany, J,.Beltran, A. Perez, S.Ganzfried, "Predicting hurricane trajectories using a recurrent neural network", arXiv 2018, arXiv:1802.02548, 2018.

[4] A. Asuero. G.Sayago. A. Gonzalez. A. G, "The Correlation Coefficient: An Overview", Critical Reviews in Analytical Chemistry, 36:41–59, 2006.

[5] L. Chen, X.Pan, Y.H. Zhang, M. Liu, T. Huang, Y.D.Cai," Classification of Widely and Rarely Expressed Genes with Recurrent Neural Network", Comput. Struct. Biotechnol.J, 17, 49–60, 2019.

[6] B. B. Chaudhuri and N. Sarkar, "Texture Segmentation Using Fractal Dimension", IEEE Trans on Pattern Analysis and Machine Intelligence Volume.17, No.1, pp. 72-77, January 1995.

[7] D. B. David and D. Rangaswamy, "Forecasting of Cyclone Using Multi-Temporal Change Detected Satellite Images", IEEE International Conference on Computational Intelligence and Computing Research, 2014.

[8] V. F. Dvorak, "Tropical Cyclone Intensity Analysis Using Satellite Data", NOAA Technical Report, September 1984.

[9] G.E. Dallal, Correlation coefficient (2003), http://www.tufts. edu/~gdallal/corr.htm.

[10] S.Gao, P.Zhao, B.Pan, Y.Li, M.Zhou,J. Xu "Zhong, S.; Shi, Z. A nowcasting model for the prediction of typhoon tracks based on a long short term memory neural network", Acta Oceanol. Sin, 37, 8–12, 2018.

[11] J.W.Hardin, J.M. Hilbe , "Generalized Linear Models and Extensions", 2nd edition, College Station: Stata Press, 2007.

[12] W. G. Hopkins, A new view of statistics, correlation coefficient, 2004. http://www.sportsci.org/resource/stats/correl.html .

[13] M. W. Khan, “A survey: Image Segmentation Techniques”, International Journal of Future Computer and Communication, Volume. 3, pp. 89, 2014.

[14] K. Kim, “Improvement of Tropical Cyclone Track Forecast over the Western North Pacific Using a Machine Learning Method”, master thesis, Department of Urban and Environmental Engineering, Graduate School of UNIST, 2020.

[15] H-J.Kim, I-J. Moon, M .Kim, “Statistical prediction of typhoon induced accumulated rainfall over the Korean Peninsula based on storm and rainfall data”, research article, Metrological Application, 16, October, 2019.

[16] P. Komarek , AW. Moore, “ Making logistic regression a Core data mining tool with TR-IRLS”. In: Proceedings of the fifth IEEE international conference on data mining. USA: IEEE Computer Society, p. 685–8, 2005. https://doi.org/10. 1109/ICDM.2005.90.

[17] G. King G, L. Zeng, “Logistic regression in rare events data”, Polit Anal. 9:137–63, 2001.

[18] H. Liu,X. Mi, and Y.Li, “Smart multi-step deep learning model for wind speed forecasting based on variational mode decomposition, singular spectrum analysisz”, LSTM network and ELM, Energ. Convers. Manage., 159, 54–64, https://doi.org/10.1016/j.enconman.2018.01.010, 2018.

[19] R. de Levie, Advanced Excel for Scientific Data Analysis (Oxford: Oxford University Press, 91–92, 2001.

[20] J.Lian, P.Dong, Y. Zhnag, J.Pan, “A Novel Deep Learning Approach for Tropical Cyclone Track Prediction Based on Auto-Encoder and Gated Recurrent Unit Networks”, Article, Applied Sciences, 2020.

[21] M. Moradi Kordmahalleh, M.Gorji Sefidmazgi, A.Homaifar, “A sparse recurrent neural network for trajectory prediction of Atlantic hurricanes”, In Proceedings of the ACM Genetic and Evolutionary Computation Conference 2016, Denver, CO, USA, pp. 957–969, 20–24 July 2016.

[22] S. Naz, H. Majeed and H. Irshad, “Image Segmentation using Fuzzy Clustering: A Survey”, IEEE International Conference on Emerging Technologies (ICET), 2010.

[23] J. Nayak, B. Naik, and H. Behera, “Fuzzy C-means (FCM) clustering algorithm: a decade review from 2000 to 2014”, Computational Intelligence in Data Mining-Volume 2, ed: Springer, pp. 133-149, 2015.

[24] J.Nakamuar, “Western North Pacific tropical cyclone model tracks in present and future climates”, J Geophys Res Atmos 122(18): 9721-9744, 2017.

[25] M. Á. P. Ortiz and Cuba, “Visualization and Segmentation of Tropical Cyclones in Satellite images using the Dvorak technique”.

[26] S. Prabhakar, S. Kumar, “A Survey on Fuzzy C-means Clustering Techniques”, International Journal of Engineering Development and Research (IJEDR), Volume. 5, Issue. 4, 2017.

[27] A. Preeti and K. Ahuja, “A Survey on Image Segmentation Using Clustering Techniques Image Segmentation using Fuzzy Clustering: A Survey”, International Journal for Research in Applied Science and Engineering Technology, Volume.2, Issue. V, pp. 51-55, May 2014.

[28] J. C. H. Poon, C. P. Chau and M. Ghadiali, “Using Fuzzy Mathematical Morphology for Locating Tropical Cyclones in Satellite Imagery”, IEEE International Symposium on Circuits and Systems, pp. 1353-135, June 1997.

[29] C. Piani, J.O. Haerter, and E.Coppola, “Statistical bias correction for daily precipitation in regional climate models over Europe” , Theoretical and Applied Climatology, 99(1-2), 187-192, 21010.

[30] L. J. Quackenbush, “A Review of Techniques for Extracting Linear Features from Imagery”, Photogrammetric Engineering & Remote Sensing, Volume. 70, No. 12, pp. 1383–1392, December 2004.

[31] C. Roy and R. Kovordanyi, “Tropical cyclone track forecasting techniques: A review ”.

[32] MB.B.Richman,L.M. Lesile, H.A.Ramsay, P.J.Klotzbach, “Reducing Tropical Cyclone Prediction Errors Using Machine Learning Approaches”, Prodedia Comput, Sci. 114, 314-323,2017.

[33] V.Ravindra, S.Nag, A.Li, “Ensemble-Guided Tropical Cyclone Track Forecasting for Optimal Satellite Remote Sensing”, IEEE Transactions on Geoscience and Remote Sensing, 2020.

[34] M.Ruttgers, S.Lee, S, and D.You, “Prediction of typhoon tracks using a generative adversarial network with observational and meteorological data”, arXiv preprint arXiv:1812.01943, 2018.

[35] B.Szeląg, R. Suligowski, J.Studziński, F. De Paola, “ Application of logistic regression to simulate the influence of rainfall genesis on storm overflow operations: a probabilistic approach”, Hydrology and Earth System Sciences, 24(2), 595–614. https://doi.org/10.5194/hess-24-595-2020.

[36] A. Strzelecka, A .Kurdys-Kujawska, D.Zawadzka “Application of logistic regression models to assess household finicical decisions regarding debt”, Proceida computer science, 176, ppp.3418-3427,2020.

[37] J. Serra, “Introduction to Mathematical Morphology”, Computer Vision, Graphics and Processing, 1986, Volume 35, pp. 283-305.

[38] R. Senthikumar and K. Porkumaran, “A Fuzzy Logic Approach To Tropical Cyclone Eye Location (TCEL) Using Indian Geostationary Meteorological Satellite Imagery”, Journal of Theoretical and Applied Information Technology, Volume. 62, No.3, pp. 729-732, April 2014.

[39] J.Tan, S. Chen, J. Wang, “Western North Pacific tropical cyclone track forecasts by a machine learning model”, Springer-Verlag GmbH Germany, part of Springer Nature, 2020.

[40] C. Velden, B. Harper, F. Wells, J. L. Beven, r. Zehr, T, Olander, M. Mayfield, C. Guard, M. Lander, R. Edson, L, Avila, A. Burton, M.Turk, Ak Kikuchi, A, Christian, P. Caroff and P. Mccrone, “The Dvorak Tropical Cyclone Intensity Estimation Technique”, Color-enhanced IR image of Hurricane Katrina, viewed from GOES-12, August 2006.

[41] W. Wang and Y. Zhang, “On fuzzy cluster validity indices”, Fuzzy sets and systems, Volume. 158, pp. 2095-2117, 2007.

[42] L.Wang, B. Wan, S.Zhou,H. Sun, and Z. Gao, “Forecasting tropical cyclone tracks in the northwestern Pacific based on a deep-learning model”, Copernicus Publications, 16, 2167-2179, 2023.

[43] Y.Wang,W. Zhang, and W. Fu, “Back Propogation(BP) neural network for tropical cyclone track forecast”, Proceedings 2011 19th International Conference on Geoinformatics, Geoinformatics 2011, 24–26 June 2011,

Shanghai, China, 1–4, https://doi.org/10.1109/GeoInformatics.2011.5981095, 2011.

[44] J. Wang, J.C.Y. Guo, " An analytical stochastic approach for evaluating the performance of combined sewer overflow tanks", Water resources Research, 54(5), 3357–3375, 2018. https://doi.org/10.1029/2017WR022286.

[45] Y.Xu,C.J. Neumann, "A statistical model for the prediction of western North Pacific tropical cyclone motion (WPCLPR)", 1985.

[46] C. Zhang, Y. Chen and J. Lu, "Typhoon Center Location Algorithm Based on Fractal Feature and Gradient of Infrared Satellite Cloud Image", International Symposium on Optoelectronic Technology and Application 2014: Optical Remote Sensing Technology and Applications, Volume. 9299 92990F-1.

[47] Q. P. Zhang, L. L. Lai and W. C. Sun, "Intelligent Location of Tropical Cyclone Center", Proceeding of the Fourth International Conference on Machine Learning and Cybernetics, August 2005.

[48] Y.Zhang, "Cyclone Track Prediction with Matrix Neural Networks", International Joint Conference on Neural Networks (IJCNN),2018.

[49] M. F. Zady, Correlation and simple least squares regression, October 2000. http://www.westgard.com/lesson44.htm .

ACRONYMS

ANN	Artificial Neural Network
ADT	Advanced Dvorak Technique
JTWC	Joint Typhoon Warning Center
MLP	Multiple Perceptron
TC	Tropical Cyclone
PCA	Principal Component Analysis
SVM	Support Vector Machine