A Project Report

(part-1)

On

PREDICTION OF DISSOLVED OXYGEN IN AQUACULTURE PONDS USING LSTM AND GRU (HYBRID MODEL).

Submitted in partial fulfilment of the requirements for the award of the degree of

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE AND ENGINEERING

By

S. LAVANYA	**(20NN1A05A4)**
R. HEMA LATHA	**(20NN1A0598)**
R. SRIVANI	**(20NN1A05A1)**
R. RAJYA LAKSHMI	**(20NN1A0599)**

Under the Esteemed Guidance of

Mr. Arepalli Peda Gopi

(Assistant Professor)

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

VIGNAN'S NIRULA INSTITUE OF TECHNOLOGY AND SCIENCE FORWOMEN

PEDAPALAKALURU, GUNTUR-522005

(Approved by AICTE, NEW DELHI and Affiliated to JNTUK,Kakinada)

(2019-2023)

VIGNAN'S NIRULA INSTITUE OF TECHNOLOGY AND SCIENCE FORWOMEN

(Approved by AICTE, NEW DELHI and Affiliated to JNTUK, Kakinada)
PEDAPALAKALURU, GUNTUR-522005
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

This is to certify that the project entitled **"PREDICTION OF DISSOLVED OXYGEN IN AQUACULTURE PONDS USING LSTM AND GRU (HYBRID MODEL)"**, is a Bonafide work of S.LAVANYA(20NN1A05A4), R.HEMA LATHA(20NN1A0598), R.SRIVANI(20NN1A05A1), R.RAJYALAKSHMI (20NN1A0599), submitted to the department of Computer Science and Engineering, in the partial fulfilment of the requirements for the award of degree of **BACHELOR OF TECHNOLOGY** in **COMPUTER SCIENCE AND ENGINEERING** from **VIGNAN'S NIRULA INSTITUTEOF TECHNOLOGY AND SCIENCE FOR WOMEN, GUNTUR.**

Project Guide	**Head of the Department**
A. PEDA GOPI	**Dr. V. Sujatha**
Assistant profess or	

EXTERNAL EXAMINER

DECLARATION

We hereby declare that the work described in this project work, entitled **"PREDICTION OF DISSOLVED OXYGEN IN AQUACULTURE PONDS USING LSTM AND GRU (HYBRID MODEL)"** which is submitted by us in partial fulfilment for the award of Bachelor of Technology in the department of **Computer Science and Engineering** to the **Vignan's Nirula Institute of Technology and Science for women**, affiliated to Jawaharlal Nehru Technological University Kakinada, Andhra Pradesh, is the result of work done by us under the guidance of **Arepalli Peda Gopi.**

The work is original and has not been submitted for any Degree/Diploma of this or any other university.

Place: Guntur

Date:

S. LAVANYA (20NN1A05A4)

R. HEMA LATHA (20NN1A0598)

R. SRIVANI (20NN1A05A1)

R. RAJYA LAKSHMI (20NN1A0599)

ACKNOWLEDGEMENT

We express our heartfelt gratitude to our beloved principal **Dr. P. Radhika, Professor** for giving a chance to study in our esteemed institution and providing us all the required resources.

We would like to thank **Dr. V. Sujatha, Associate Professor, Head of the Department of Computer science and Engineering**, for her extended and continuous support, valuable guidance and timely advices in the completion of this project thesis.

We wish to express our profound sense of sincere gratitude to our Project Guide, **A. Peda Gopi, Assistant Professor, Department of Computer Science and Engineering**, without whose help, guidance and motivation this project thesis could not have been completed the project successfully.

We also thank all the faculty of the Department of Computer Science and Engineering for their help and guidance of numerous occasions, which has given us the cogency to build-up adamant aspiration over the completion of our project thesis.

Finally, we thank one and all who directly or indirectly helped us to complete our project thesis successfully.

S. LAVANYA (20NN1A05A4)

R. HEMA LATHA (20NN1A0598)

R. SRIVANI (20NN1A05A1)

R. RAJYA LAKSHMI (20NN1A0599)

PREDICTION OF DISSOLVED OXYGEN

IN AQUACULTURE PONDS USING LSTM AND GRU (HYBRID MODEL)

ABSTRACT

An essential ecological component in promoting the development of aquatic products is dissolved oxygen in water. The content of Dissolved Oxygen is an important indicator of control in recirculating aquaculture, it's content and dynamic changes have great impact on healthy growth of fish. The physical, chemical, and biological parameters that affect the Dissolved oxygen in water are PH, Temperature, Nitrate, and Biological Oxygen Demand, Mg, EC, Turbidity, Na, PO4, PH, Ca, NO3 and NO2. Hence it is important to detect Dissolved Oxygen in Water. In the earlier days people used traditional methods like Laboratory analysis for water quality prediction. However Traditional methods cannot capture Non-linear and Non-Stationary data for water quality. Additionally Laboratory methods requires a lot of time and manual work. Researches are being done using ML and DL models such as Artificial neural network(ANN), Long short term memory(LSTM), Gated recurrent unit(GRU) are often used. ANN requires large amount of data and is easy to fall into local minima. but it is unclear which one of them is more suitable for prediction of DO in Aquaculture. In order to improve the accuracy and effectiveness of dissolved oxygen (DO) prediction, a combined forecasting model based on LSTM - Gated Recurrent Unit (GRU) hybrid Model was used to predict the content of DO in water for growth and Survival of fishes. Proposed Hybrid model yields more accurate values than the existing models.

TABLE OF CONTENTS

TABLE OF CONTENTS

List of Figures

List of Tables

List of Acronyms

S.No	Acronym	Full Form
1	ANN	Artificial Neural Networks
2	RNN	Recurrent Neural Networks
3	LSTM	Long Short Term Memory
4	GRU	Gated Recurrent Unit
5	DO	Dissolved Oxygen
6	BOD	Biological Oxygen Demand
7	pH	Potential of Hydrogen
8	MSE	Mean Squared Error
9	RMSE	Root Mean Squared Error

CHAPTER – 1

INTRODUCTION

CHAPTER – 1

INTRODUCTION

1.1 Introduction

In aquaculture, dissolved oxygen (DO) has become an important parameter to predict water quality [1]. Water quality (WQ) is usually determined by the general composition of water in relation to its physical, biological, and chemical properties [2]. DO (dissolved oxygen) is one of the freshwater properties and undoubtedly one of the most important components for the survival of the aquatic life. DO concentration is an important WQ indicator of water pollution in the aquaculture ecosystem [3]. In the aquaculture farms, the required concentration of dissolved oxygen typically depends on the fish species and the water temperature. However, concentrations below 3 mg/L are related to stress in the aquatic species, increasing mortality and disease in most of the species. Dissolved Oxygen in Water is affected by many physical, chemical and biological parameters. These parameters include pH, Temperature, BOD, DO, Nitrate, Sulphates, salinity, Turbidity, Ammonia, and many more [5]. The changes in any one parameter will affect the levels of other parameters. The characteristics of dissolved oxygen in pond aquaculture water vary with meteorological and diurnal changes. Under better weather conditions, the dissolved oxygen content in water is higher during the day, but decreases significantly at night and in the early morning. To prevent water quality deterioration and aquatic product death caused by hypoxia, intelligent algorithms are used to predict the dissolved oxygen.

Temperature is a predominant controlling factor for fish growth and effect the Dissolved oxygen [6]. All chemical and biological mechanisms are influenced by temperature [7]. The optimal range of temperature for which the Fishes metabolism is not affected is 6.0oC - 22.5oC. The temperatures outlying this range can affect the productivity of Fishes [8-10]. PH (Potential of Hydrogen) is an effective measure of the concentration of hydrogen ion (H+)in a substance. The pH range that is acceptable for Fishes is 5.7- 8.5 (both inclusive). Higher pH levels cause denaturing cellular membranes that leads to death of fish. Also, lower ranges of pH can accelerate the inability of fish to breathe [11].

Biological oxygen demand (BOD) is the measure of dissolved oxygen required to decompose the organic compounds by aerobic organisms in water. If BOD is high, it generates obnoxious smell and creates unhealthy environment. It is usually measured in mg/l (milli grams per litre) [12]. Lower ranges of BOD are appropriate for healthy growth of fish normally between 1.0 mg/l – 5.0 mg/l. Biological oxygen demand and dissolved oxygen are inversely related to each other, BOD also depends on water temperature. DO (Dissolved Oxygen) is the quantity of oxygen that is present in water usually measured in mg/l units [13]. The DO and temperature are inversely related to each other that is, when there in an increase in water temperature, the amount of oxygen that gets dissolved into water is low. If DO percentage is too low then the Fishes cannot survive [14]. The optimal range of DO id 6.5 mg/l – 9.0 mg/l.

Nitrate: Nitrate is a combination of two products which involves oxidized NH3 or NH4+ during this conversion of the intermediate product NH3 into NO3.

The desired concentration of nitrate in the water is among 0.06 mg/l to 0.5 mg/l. So, these parameters are very important In Fish growth and do not effect the Dissolved Oxygen.

By using above factors, we can estimate the Rate of Dissolved Oxygen in water required for healthy growth of Fishes. In the existing system ANN algorithm was used for evaluating the Rate of Dissolved Oxygen in water for the growth of fishes. The parameters used in ANN are pH, Nitrate, BOD, temperature. This algorithm gave better results in predicting rate of Dissolved Oxygen in Water for Fishes, but the accuracy for these algorithms is not so optimal. This is becauseANN gives a probing solution that leads to loss of trust in the network. To overcome thisproblem, LSTM and GRU algorithms (HYBRID MODEL) was adopted. The Model gave better results when compared to ANN algorithm. Table 1 describe the parameters that were taken into consideration for estimating rate of Dissolved Oxygen in water. It also describes the units of every parameter and their maximum and minimum values that are suitable for Fish Survival.

Table -1: Ranges of Factors Affecting the Dissolved Oxygen in Water.

PARAMETERS	UNITS	MINIMUM	MAXIMUM
pH	-------	5.7	8.5
Temperature	^{o}C	6.0	22.5
Total Alkalinity	mg/l	50	150
Biological Oxygen Demand	mg/l	1.0	5.0
Nitrate	mg/l	0.06	0.5

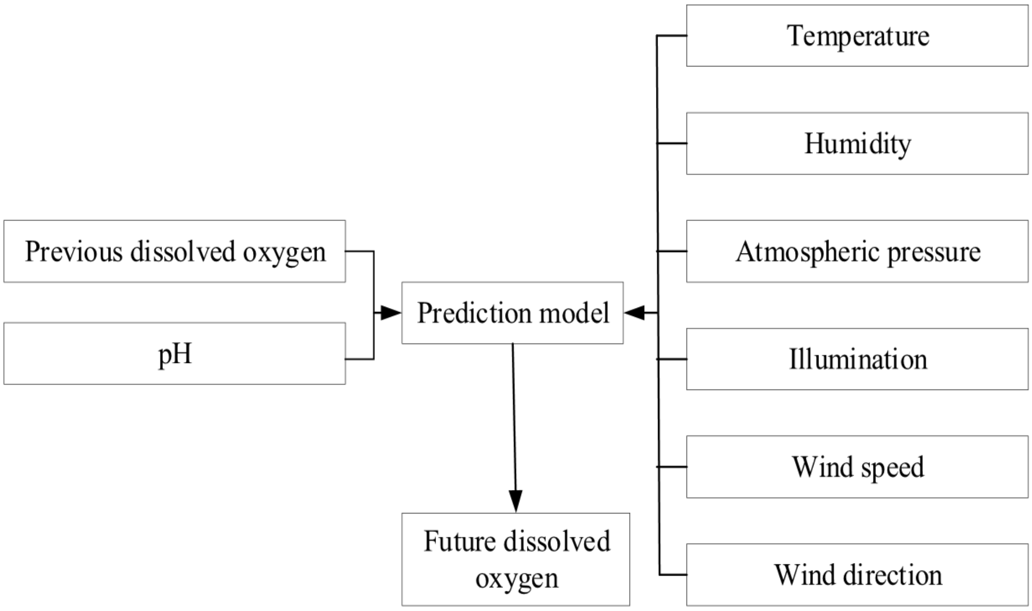


Fig: 1.1 Input output structure for forecasting model.

The above figure1.1 depicts how various atmospheric elements effect the rate of Dissolved oxygen in water and how it effects the survival of fishes in Aquaculture.

1.1 PROBLEM STATEMENT

The goal is to predict Rate of Dissolved Oxygen in water that accurately and efficiently in an hourly manner, and to alert the user in advance. Regression analysis and prediction method through a large number of data samples for correlation analysis to obtain the correlation, and establish a regression equation, on the basis of considering the prediction error to determine the future Rate of Dissolved Oxygen in Water prediction value. The model is complex, and the distribution of data and samples requires higher requirements. In spite of the fact that ANN models are very adaptable to time series, their major impediment is that no non- linear can be captured by the ANN model. LSTM and GRU hybrid model is picked so as to build up a far- reaching approach for effective Dissolved Oxygen prediction and analysis. LSTM and GRU hybrid model can dynamically memorize and retain the historical water quality parameters information while learning new information.

1.2 OBJECTIVES

The main objective is to predict the rate of Dissolved Oxygen in water for growth and Survival of Fish.

1. To Maintain Dissolved Oxygen rate efficiently.
2. To predict the rate of Dissolved Oxygen in water accurately, DL Algorithms are used.
3. Predict the water parameters in an hourly manner. i.e., predict next hour values based on previous data.
4. When predicted rate of Dissolved Oxygen seems to exceed the critical conditions an alert is
5. sent to user immediately to take the precautionary actions required.
6. Build a web application using Flask to show alerts to the users as well as data visualization of the various water parameters considered and Rate of Dissolved Oxygen

1.3 SCOPE OF THE PROJECT

1. The project focusses on the importance of rate of Dissolved Oxygen in water for aquaculture. It helps the aqua farmers to produce good quality of fishes, which in turn helps the economy of aquaculture sector.
2. Deep Learning helps provide a better, accurate and faster prediction of rate of Dissolved Oxygen in water based on the previous data collected.
3. Predictive analysis can help to capture relationships among numerous variables that can help to assess risk with a particular set of conditions. LSTM and GRU hybrid model is considered better and gives increasingly precise data to foresee and assess the rate of Dissolved Oxygen in water for survival of fish.
4. In Asian/American Cultures, freshwater fishes like are predominantly grown through Aquaculture.

1.4 METHODOLOGY

To complete this project successfully agile methodology as shown in Table 1.2 is followed.

- **Sprint 1:** Literature survey is done to understand other related works in this domain andhence do the analysis of problem in the existing model.

- **Sprint 2:** Identifying the project goals, project requirements and formulating theproblem statement. Collection of dataset and preprocessing.
- **Sprint 3:** Developing the models with various DL algorithms. These models are trained, tested and evaluated.
- **Sprint 4:** Choosing the best model and predicting the water parameters, rate of Dissolved Oxygen in hourlymanner.
- **Sprint 5:** Developing web application for displaying the predicted values and alert.
- **Sprint 6:** Implementation of data visualization for better understanding of the parameterchances over the past hours.

Table 1.2 Software Agile Technology

Story Id	Requirement Description	User stories/Taks	description
1	Collection of waterdata with various Parameters	Data collection	In .csv format
2	Determing the relevant parameters needed for predicting the rate of DO in water	Cleaning the dataset	Removal of noisy data
3	Selection of algorithm	Designing the model	Identifying the best algorithmwith data available
4	Training the data sets	Training, testingand evaluating the model	Training the model to predictthe water quality
5	Predict rate of DO in water on hourlybasis	Predict rate of DO in water.	Rate of DO in Water is predicted for next hour based past six hours
6	Testing the results		Accuracy calculation
7		Alert system and data visualiz ation	To display alert message and graph

Calculating the accuracy of algorithm Web application	showing the predicted values	and appropriatemessage. Changes of the parameters are shown in the form of the graphs

1.5 SUMMARY

In this chapter, we have discussed about water quality parameters and their impact on Rate of Dissolved Oxygen in water in aquaculture, its importance in the economy according to research. We have also discussed about the scope and the relevance of our project in the current trends. The main objectives of predicting Rate of Dissolved Oxygen in water for growth and survival of fish is discussed in this chapter.

CHAPTER – 2

LITERATURE SURVEY

CHAPTER - 2
LITERATURE SURVEY

“Prediction of dissolved oxygen content in aquaculture based on clustering and improved ELM” is written by Shouqi Cao, Lixin Zhou, Zhang. To improve the accuracy and effectiveness of dissolved oxygen (DO) prediction, a combined forecasting model based on ensemble empirical mode decomposition (EEMD) and least squares support vector machine(LSSVM) is used. .EEMD-LSSVM model Shows the higher prediction accuracy in terms of all the evaluation indexes.

“Prediction of dissolved oxygen content in aquaculture using clustering based soft plus extreme learning machine” is written by Pei Shi, Guanghui Li , Yongmingyuan , Guangyan huan , Liang Kuang . Clustering-based Soft plus Extreme Learning Machine method (CSELM) is used to accurately and efficiently predict dissolved oxygen change from time series data.The proposed CSELM is obviously the best model for predicting the water quality changes in terms of accuracy and efficiency, compared to the two counterparts (PLS-ELM and ELM).

“Prediction of dissolved oxygen in pond culture water based on K-means clustering and gated recurrent unit neutral network” is written by Xinkai Cao, Yiran Liu, Jianbing Wang, Chunhong Liu, Qingling Duan. The prediction model of dissolved oxygen in pond culture was proposed based on K-means clustering and Gated Recurrent Unit (GRU) neural network. proposed method can predict the dissolved oxygen content of aquaculture water overdifferent time intervals according to the demands of real-world scenarios.

“Prediction of dissolved oxygen in a fishery pond based on gated recurrent unit” is written by Wuyan Li, Hao Wu, Nanyang Zhu, Yongnian Jiang, Jinglu Tya, Guo. RNN model, LSTM model, and GRU model were used to build three DO predicting models.The performance of RNN is worse result than LSTM and GRU. Among all but the time cost and number of parameters used for GRU is much lower than LSTM. It is concluded that the GRU has overall better performance and can be applied to practical applications.

“Dissolved oxygen forecasting in Aquaculture a hybrid model approach” is written by Elias Eze, Tahmina Ajmal. we present a novel hybrid DO concentration forecasting method with ensemble empirical mode decomposition (EEMD)-based LSTM (long short-term memory)neural network (NN). The performance of this proposed model in training and validation set was compared with the observed real sensor data.

“Prediction of dissolved oxygen in pond culture water based on K-means clustering and gated recurrent unit neural network” is written by Xinkai Cao, Yiran Liu, Jianbing Wang ,Chunhong

Liu, Qingling Duan. The prediction model was proposed based on K-meansclustering and Gated Recurrent Unit (GRU) neural network. proposed method can predictthe dissolved oxygen content of aquaculture water over different time intervals according to the demands of real-world scenarios.

“Dissolved oxygen prediction using a new ensemble method” is written by Ozgur Kisi and Meysam Aliza Mir and AliReza Docheshmeh Gorgij. The potential of the BMA investigated and compared with five data-driven methods, extreme leaning machine (ELM), artificial neural networks (ANNs), adaptive neuro-fuzzy inference system (ANFIS), classification and regression tree (CART), and multilinear regression (MLR), by considering hourly temperature. the Bayesian model averaging considerably improved the estimation accuracy.

“The dissolved oxygen prediction method based on neural network” is written by Zhong Xiao, Lingzi Peng, Yi Chen, Haohuai Liu, Jiaqing Wang, and Yangang Nie. The methods that are used for predicting DO in aquaculture based on BP neural network. This method fully meets the needs of practical applications and is suitable for a wide range of promotion. The prediction model can help to improve the water quality monitoring level of aquaculture which will prevent the deterioration of water quality and the outbreak of disease.

The below Table 2.1 depicts the various research papers, that are related to prediction of Dissolved oxygen in Aquaculture ponds using various machine learning and deep learning techniques.

Table-2.1: Compression of different existing models

s.no	Authors	Description of papers	year	Advantages	Limitations
1	Juan Huan, weijian Cao, Yilin Qin.	To improve the accuracy and effectiveness of dissolved oxygen (DO) prediction, a combined forecasting model based on ensemble empirical mode decomposition (EEMD) and least squares support vector machine (LSSVM) is proposed. This paper used the single point iterative method to achieve multi-step prediction in order to obtain forecasting results for 24 h into the future. The EEMD-LSSVM model has high forecast accuracy and generalization ability.	2018	1. EEMD-LSSVM model Shows the higher prediction accuracy in terms of all the evaluation indexes. 2. LSSVM network has a faster convergence rate and higher accuracy of prediction compared with the conventional methods.	1. EEMD-LSSVM model Shows the higher prediction accuracy in terms of all the evaluation indexes. 2. LSSVM network has a faster convergence rate and higher accuracy of prediction compared with the conventional methods.
2	Qin Rena, Xuanyu wang, Wenzhou Li, Yaoguang Wei, Dong An.	The content of dissolved oxygen is an important indicator of control in recirculating aquaculture, its content and dynamic changes have great impact on the healthy growth of fish. The prediction model based on deep	2020	1. To solve the problems of traditional prediction model that low accuracy, poor stability, noise in the local characteristic data	The VMD-DBN model performs better than the DBN according to the MAE, RMSE, MAPE and R^2.

		belief network has been proposed in this paper to realize the dissolved oxygen content prediction. The VMDDBN model produces higher prediction accuracy and stability.		collected from water quality, the model which based on VMD-DBN has been proposed in this paper. 2.This model can quickly and accurately predict the dissolved oxygen content in time series, and the prediction performance meets the needs of actual production.	
3	Shouqi Cao, Lixin Zhou, Zhang.	In this paper, a hybrid method is proposed to predict the change of dissolved oxygen from the perspective of time series in aquaculture, which based on k-means clustering and improved Soft plus extreme learning machine (SELM) with particle swarm optimization (PSO).	2021	1. The experimental results show that PLS-SELM model can achieve better prediction performance and accuracy of prediction compared with other single models. 2. It provides an accurate predictive model framework for researchers to track dissolved oxygen.	The performance of the model doesn't give more in-depth analysis.
4	Pei Shi, Guanghui Li, Yongmingyuan, Guangyanhuan , Liang Kuang	This paper provides a novel Clustering-based Soft plus Extreme Learning Machine method (CSELM) to accurately and efficiently predict dissolved oxygen change from time series data. The CSELM adopts k-medoids clustering to group the dataset into different clusters based on Dynamic Time Warping (DTW) distance, and uses a new Soft plus ELM algorithm to discover a common trend in a cluster of time series pieces (within the same period) and then predict the future trend.	2019	1. The proposed CSELM is obviously the best model for predicting the water quality changes in terms of accuracy and efficiency, compared to the two counterparts (PLS-ELM and ELM). 2. CSELM is able to discover the periodic change patterns and trend changes of water quality and meteorological time series data (in aquaculture) and thus provide better prediction results.	The CELM achieves worse precision (RMSE, MAPE and MAE); this demonstrate that Soft plus in both CSELM and SELM improves the accuracy.

5	Xinkai Cao, Yiran Liu, Jianbing Wang , Chunhong Liu, Qingling Duan.	The accurate monitoring and prediction of dissolved oxygen is the key to precise regulation and control of pond aquaculture water quality. The prediction model of dissolved oxygen in pond culture was proposed based on K-means clustering and Gated Recurrent Unit (GRU) neural network.	2020	1. proposed method can predict the dissolved oxygen content of aquaculture water over different time intervals according to the demands of real-world scenarios. 2. current dissolved oxygen prediction model has some limitations, such as a short prediction period and inadequate prediction accuracy for actual production demands. Therefore, a prediction model of dissolved oxygen in pond culture was proposed based on K-means clustering and Gated Recurrent Unit (GRU) neural network.	This model does not provide the better optimized parameter selection of the GRU model in an attempt to improve the experimental accuracy.
6	Wuyan Li, Hao Wu, Nanyang Zhu, Yongnian Jiang, Jinglu Tya, Guo.	In this paper, the RNN model, LSTM model, and GRU model were used to build three DO predicting models. The performance of the three models were compared by mean absolute error (MAE), mean square error (MSE), mean absolute percentage error (MAPE), and the coefficient of determination (R^2). The performance of RNN is worse result than LSTM and GRU. Among all but the time cost and number of parameters used for GRU is much lower than LSTM. It is concluded that the GRU has overall better performance and can be applied to practical applications.	2020	In this study, the RNN, LSTM, and GRU are compared which is suitable for prediction of DO. The fitting results and the evaluation values show that the predictive performance of GRU and LSTM have a better performance than RNN	This model takes more time For prediction of DO
7	Elias Eze , Tahmina Ajmal.	In this article, we present a novel hybrid DO concentration forecasting method with ensemble empirical mode decomposition (EEMD)-based LSTM (long short-term memory) neural network (NN). This can provide the basis of data support for an early warning system, for an improved management of the aquaculture farm. hybrid EEMD-based LSTM forecasting model is then constructed. The performance of this proposed model in training and validation sets was compared with the observed real sensor data.	2020	Theactual experimental WQ data from a fish farm show that the hybrid model provides good results and outperforms related models with high accuracy.	a hybrid EEMD-based multi-variate prediction model can be explored to propose a more comprehensive water quality forecasting and analysis. it doesn't used more WQ measuring sites considered to expand this model.
8	Ozgur Kisi and Meysam Aliza Mir and AliReza Docheshmeh Gorgij.	, pH, and specific conductivity data as inputs. This study proposes a new ensemble method, Bayesian model averaging (BMA), for estimating hourly dissolved oxygen. The potential of the BMA was investigated and compared with five data-driven methods, extreme leaning machine (ELM), artificial neural networks (ANNs), adaptive neuro-fuzzy inference system (ANFIS), classification and regression tree	2019	The Bayesian model averaging considerably improved the estimation accuracy; by about 5–8%, 13–12%, 7–9%, 18–27%, and 7–32%, compared to the best single ELM, ANN, ANFIS, CART, and MLR models, respectively.	Uncertainty analysis of the new ensemble method showed the high complexity and nonlinearity of the investigated phenomenon. Compliance with it

		(CART), and multilinear regression (MLR), by considering hourly temperature			
9	Zhong Xiao, Lingzi Peng, Yi Chen, Haohuai Liu, Jiaqing Wang, and Yangang Nie.	The paper presents the methods for predicting DO in aquaculture based on BP neural network. Using BP neural network method with the combination of purlin, losing, and tansig activation functions is proposed for the prediction of aquaculture's dissolved oxygen.	2017	The prediction model can help to improve the water quality monitoring level of aquaculture which will prevent the deterioration of water quality and the outbreak of disease. This method fully meets the needs of practical applications and is suitable for a wide range of promotion. The future work for us to do is to apply the neural network method in aquaculture	It takes more time for prediction and that much faster compared to other
10	Reza Dehghani, Hassan Torabi Poudeh, Zohreh Izadi·	In this study, hybrid models based on support vector regression were used to simulate dissolved oxygen in river water. Support vector regression (SVR) was employed for prediction of the model in both standalone and hybrid forms. The employed hybrid models consisted in SVR combined with metaheuristic algorithms of chicken swarm optimization (CSO), social ski driver (SSD) optimization, Black widow optimization (BWO), and the Algorithm of the innovative gunner (AIG). Pearson correlation coefficient was utilized to select the best input combination.	2021	The results showed that the higher the number of effective parameters (dependent variables) in hand, the better the network performance. Also, the higher the input to the network, the higher the efficiency and accuracy of the model. Overall, this study showed that the use of SVR–AIG model could be effective in estimating dissolved oxygen in river water.	The SVR–BWO, SVR–CSO, and SVR–SKI have more error rate than SVR-AIG.
11	Wuyan Li, Hao Wu, Nanyang Zhu, Yongnian Jiang, Jingle Tan ,Yaa Guo, Key Laboratory.	In This learning models, such as recurrent neural network (RNN), long short-term memory (LSTM), and gated recurrent unit (GRU), are often used to predict the trend of time series, but it is unclear which one of them is more suitable for prediction of DO in fishery ponds.	2020	The fitting results and the evaluation values show that the predictive performance of GRU and LSTM have a better performance than RNN.	This model takes long time for evaluation of results.

CHAPTER – 3

SOFTWARE REQUIREMENT SPECIFICATION

CHAPTER - 3

SOFTWARE REQUIREMENTS SPECIFICATION

3.1 Requirement Analysis

Software Requirement Specification (SRS) is the starting point of the software developing activity. As system grew more complex it became evident that the goal of the entire system cannot be easily comprehended. Hence the need for the requirement phase arose. The software project is initiated by the client needs. The SRS is the means of translating the ideas of the minds of clients (the input) into a formal document (the output of the requirement phase.) Under requirement specification, the focus is on specifying what has been found giving analysis such as representation, specification languages and tools, and checking the specifications are addressed during this activity. The purpose of the Software Requirement Specification is to reduce the communication gap between the clients and the developers. Software Requirement Specification is the medium though which the client and user needs are accurately specified. A good SRS should satisfy all the parties involved in the system.The software used are open source and easy to install. This is an independent application which can be easily run on to any system whichhas Python installed and Jupyter Notebook. Product features the application is developed in such a way that the Rate of Dissolved Oxygen in Aquaculture Ponds is predicted using LSTM and GRU hybrid Model. We can compare the accuracy for the implemented algorithms. User Requirements in which user can decide on the prediction accuracy to decide on which algorithm can be used in real-time predictions. Non-Functional Requirements why Dataset collected should be in the CSV format why The column values should be numerical values why Training set and test set are stored as CSV files why Error rates can be calculated for prediction algorithms Product Requirements Efficiency: Reliability: Maturity, fault tolerance and recoverability. Portability: can the software easily be transferred to another environment, including install ability. Usability: How easy it is to understand, learn and operate the software system Organizational Requirements: Do not block some available ports through the windows firewall.

Hardware Interfaces Ethernet on the AS/400 supports TCP/IP, Advanced Peer-to-Peer Networking (APPN) and advanced program-to-program communications (APPC). ISDN to connect AS/400 to an Integrated Services Digital Network (ISDN) for faster, more accurate data transmission. An ISDN is a public or private digital communications network that can support data, fax, image, and other services over the same physical interface. The customers are those that perform the eight primary functions of systems engineering, with special emphasis on the operator as the key customer Health and Safety. The software may be safety-critical. If so, there are issues associated with its integrity level. The software may not be safety-critical although it forms part of a safety-critical system. For example, software may simply log transactions. If a system must be of a high integrity level and if the software is shown to be of that integrity level, then the hardware must be at least of the same integrity level. If a computer system is to run software of a high integrity level, then that system should not at the same time accommodate software of a lower integrity level. Systems with different requirements for safety levels must be separated. Otherwise, the highest level of integrity required must beapplied to all systems in the same environment.

3.2 Software Requirements

1 OS: Windows or Linux
2 Python IDE: python 2.7.x and above
3 Pycharm IDE Required, jupyter notebook
4 Setup tools and pip to be installed for 3.6 and above
5 Language: Python Scripting

3.3 Hardware Requirements

1 RAM : 4GB and Higher
2 Processor : Intel i3 and above
3 Hard Disk : 500GB: Minimum

3.4 Functional Requirements

Outputs from computer systems are required primarily to communicate the results of processing to users. They are also used to provide a permanent copy of the results for later consultation. There are various types of outputs in general are:

1 External Outputs, whose destination is outside the organization.
2 Internal Outputs whose destination is within organization and they are the users main interface with the computer.
3 Operational outputs whose use is purely within the computer department.
4 Interface outputs, which involve the user in communicating directly.

5 Understanding user's preferences, expertise level and his business requirements through a friendly questionnaire.
6 Input data can be in four different forms – Relational DB, text files, .xls and xml files.

3.4.1 Output Definition

The outputs should be defined in terms of the following points:

1. Type of the output
2. Content of the output
3. Format of the output
4. Location of the output
5. Frequency of the output
6. Volume of the output
7. Sequence of the output

It is not always desirable to print or display data as it is held on a computer. It should be decided as which form of the output is the most suitable.

3.4.2 Input design

Input design is a part of overall system design. The main objective during the input design is as given below:

1. To produce a cost-effective method of input.
2. To achieve the highest possible level of accuracy.
3. To ensure that the input is acceptable and understood by the user.

3.4.3 Input types

It is necessary to determine the various types of inputs. Inputs can be categorized as follows:

1. External inputs, which are prime inputs for the system.
2. Internal inputs, which are user communications with the system.
3. Operational, which are computer department's communications to the system?
4. Interactive, which are inputs entered during a dialogue.

3.4.4 Input stages

The main input stages can be listed as below:

1. Data recording

2. Data transcription
3. Data conversion
4. Data verification
5. Data control
6. Data transmission
7. Data validation
8. Data correction

3.4.5 Input media

At this stage choice has to be made about the input media. To conclude about the input media consideration has to be given to;

1. Type of input
2. Flexibility of format
3. Speed
4. Accuracy
5. Verification methods
6. Rejection rates
7. Ease of correction
8. Storage and handling requirements
9. Security
10. Easy to use
11. Portability

Keeping in view the above description of the input types and input media, it can be said that most of the inputs are of the form of internal and interactive. As input data is to be the directly keyed in by the user, the keyboard can be considered to be the most suitable input device.

Error avoidance

At this stage care is to be taken to ensure that input data remains accurate form the stage at which it is recorded up to the stage in which the data is accepted by the system. This can be achieved only by means of careful control each time the data is handled.

Error detection

Even though every effort is made to avoid the occurrence of errors, still a small proportion of errors is always likely to occur, these types of errors can be discovered by using validations to check the input data.

Data validation

Procedures are designed to detect errors in data at a lower level of detail. Data validations have been included in the system in almost every area where there is a possibility for the user to commit errors. The system will not accept invalid data. Whenever an invalid datais keyed in, the system immediately prompts the user and the user has to again key in thedata and the system will accept the data only if the data is correct. Validations have been included where necessary.

The system is designed to be a user friendly one. In other words, the system has been designed to communicate effectively with the user. The system has been designed with popup menus.

3.5 Non-Functional Requirements

All the other requirements which do not form a part of the above specification are categorize as non-functional requirements. A system may be required to present the user with a display of the number of records in a database. This is a functional requirement. How up to date this number needs to be is a non-functional requirement. If the number needs to be updated in real time, the system architects must ensure that the system is capable of updating the displayed record count with in a acceptably short interval of the number of the records changing. Sufficient network bandwidth may also be a non-functional requirement of a system.

CHAPTER – 4

SYSTEM ANALYSIS

CHAPTER - 4
SYSTEM ANALYSIS

4.1 Existing System

In existing system, different machine learning as well as deep learning algorithms were used to predict the Rate of Dissolved Oxygen in Aquaculture ponds that is suitable for fish. The dataset considered in some of the existing system includes images of the water in the pond. Through feature extraction of the images, rate of DO in Water was predicted. Very few systems considered the percentage presence of water quality parameters, rate of Dissolved Oxygen in Water and its impact on fishes. In the existing system, ANN algorithm was used to predict the quality of water by onesidering only two or three parameters.

4.1.1 Disadvantages

1) Hardware Dependence:
   a) Artificial Neural Networks require processors with parallel processing power, by their structure.
   b) For this reason, the realization of the equipment is dependent.
2) Unexplained functioning of the network:
   a) This the most important problem of ANN.
   b) When ANN gives a probing solution, it does not give a clue as to why and how.
   c) This reduces trust in the network.
3) Assurance of proper network structure:
   a) There is no specific rule for determining the structure of artificial neural networks.
   b) The appropriate network structure is achieved through experience and trial and error.
4) The difficulty of showing the problem to the network:
   a) ANNs can work with numerical information.
   b) Problems have to be translated into numerical values before being introduced to ANN.
   c) The display mechanism to be determined will directly influence the performance of the network.
   d) This is dependent on the user's ability.
5) The duration of the network is unknown:
   a) The network is reduced to a certain value of the error on the sample means that

the raining has been completed.

b) The value does not give us optimum results.

4.2 Proposed System

Our goal is to implement deep learning model in order to predict the rate of Dissolved Oxygen in water, to the highest possible degree of accuracy, The dataset that consists of different water quality parameters of water is gathered from Kaggle. After initial data exploration, we knew we could implement a Neural Network model for best accuracy reports.

As mentioned earlier, rate of Dissolved Oxygen in water is an important role in determining the growth of fish. In Aquaculture farms, the farmers usually grow the fishes in freshwater environments. Harvest and health of fishes are affected by many factors/parameters liketemperature, PH, BOD and nitrate present in fresh water. So, in our proposed mechanism LSTM and Gated Recurrent unit network (GRU) hybrid model was used to estimate the rate of Dissolved Oxygen in Aquaculture ponds.

4.2.1 Algorithm

Input: input all the parameters like ph,turbidity,temperature,….etc

Output: predict the rate of dissolved oxygen.

Step1 : Take the input values and declare h(t-1) and current input Xt.

Declare σ () = sigmoid & tanh()=hyperbolic tangent are activation Functions.

Step 2 : Intialize the net parameters LSTM net, calculate the forgetgate as following.

Ft= σ [(Wf *Xt) + (Wf*h(t-1))+ bf]

Now calculate the output of forgetgate as follows:

Ctf= c(t-1)*ft

Step 3: calculate the inputgate ,it has two parts as shown below

it= σ [(Wi *Xt)+(Wi*h(t-1))+ bi]

gt=tanh[(Wg*Xt)+(Wg*h(t-1))+ bg]

Now calculate the output of inputgate as follows:

Cti= it*gt

Ct=Cti+Ctf.

Step 4 : Calculate the outputgate ,as below ot= σ [(Wo *Xt)+(Wo*h(t-1))+ bo]

Step 5 : now calculate :ht = tanh(Ct)+ot.

// GRU algorithm

Step 6 : take the past information as input , Calculate the Updategate Zt

As Follows:

Zt= σ [WzXt + Wz h(t-1)]

Step 7 : Calculate the resetgate denoted by Rt as follows:

Rt= σ [WrXt + Wr h(t-1)]

Step 8 : Calculate the memory content which will use the resetgate to

Store the Relevant information from the past:

ht =tanh[WXt+Rt (xnor) Wh(t-1)]

Step 9: Now finally calculate htt—Vector which holds the information,

Update-gate is Required.

ht=Zt (xnor) h(t-1) +(1- Zt) (xnor) h't .

Step 10: END.

4.2.2 Advantages

1. LSTM and GRU hybrid model is less complex than any other neural networks because it has less numberof gates.
2. The Accuracy obtained was almost equal to cent percent which proves using of LSTM and GRU hybrid model gives best results.
3. The plots that were plotted according to the proper data that is processed during their implementation.
4. LSTM and GRU hybrid model uses less training parameter and therefore uses less memory and executesfaster.

CHAPTER –5

FEASIBILITY STUDY

CHAPTER-5

FEASABILITY STUDY

A Feasibility Study is a preliminary study undertaken before the real work of a project starts to ascertain the likely hood of the project's success. It is an analysis of possible alternative solutions to a problem and a recommendation on the best alternative.

1. Economic Feasibility
2. Technical Feasibility
3. Operational Feasibility

1.1 Economic Feasibility

It is defined as the process of assessing the benefits and costs associated with the development of project. A proposed system, which is both operationally and technically feasible, must be a good investment for the organization. With the proposed system the users are greatly benefited as the users are able to detect rate of Dissolved Oxygen in water that is useful for Aquaculture basing on the physical, chemical and biological parameters of water. This proposed system does not need any additional software and high system configuration. Hence the proposed system is economically feasible.

1.2 Technical Feasibility

The technical feasibility infers whether the proposed system can be developed considering the technical issues like availability of the necessary technology, technical capacity, adequate response and extensibility. The project is decided to build using Python. Jupyter Note Book is designed for use in distributed environment of the internet and for the professional programmer it is easy to learn and use effectively. As the developing organization has all the resources available to build the system therefore the proposed system is technically feasible.

1.3 Operational Feasibility

Operational feasibility is defined as the process of assessing the degree to which a proposed system solves business problems or takes advantage of business opportunities. The system is self-explanatory and doesn't need any extra sophisticated training. The system has built in methods and classes which are required to produce the result. The overall time that a user needs to get trained is less than one hour. As the software that is used for developing this application is very economical and is readily available in the market. Therefore, the proposed system is operationally feasible.

1.4 Effort, Duration, and Cost Estimation using COCOMO Model

The COCOMO (Constructive Cost Model) model is the most complete and thoroughly documented model used in effort estimation. The model provides detailed formulas for determining the development time schedule, overall development effort, and effort breakdown by phase and activity as well as maintenance effort.

COCOMO estimates the effort in person months of direct labor. The primary effort factor is the number of source lines of code (SLOC) expressed in thousands of delivered source instructions (KDSI).

The model is developed in three versions of different level of detail basic, intermediate, and detailed. The overall modeling process takes into account three classes of systems.

Embedded: This class of system is characterized by tight constraints, changing environment, and unfamiliar surroundings. Projects of the embedded type are model to the company and usually exhibit temporal constraints.

Organic: This category encompasses all systems that are small relative to project size and team size, and have a stable environment, familiar surroundings and relaxed interfaces. These are simple business systems, data processing systems, and small software libraries.

Semidetached: The software systems falling under this category are a mix of those of organicand embedded in nature. Some examples of software of this class are operating systems, database management system,and inventory management systems

For Basic: COCOMOEffort= a * (KLOC) bType = c * (effort) d

For Intermediate and Detailed COCOMOEffort = a * (KLOC) b* EAF

(EAF = product of cost drivers)

Table 5.1: The values of a, b, c, d for Organic, Semidetached and Embedded stems.

Type of Product	**A**	**B**	**C**	**D**
Organic	2.4	1.02	2.5	0.38
Semi Detached	3.0	1.12	2.5	0.35
Embedded	3.6	1.20	2.5	0.32

Intermediate COCOMO model is a refinement of the basic model, which comes in the function of 15 attributes of the product. For each of the attributes the user of the model has to provide a rating using the following six-point scale.

VL (Very Low)	HI (High)
LO (Low)	VH (Very High)
NM (Nominal)	XH (Extra High)

The list of attributes is composed of several features of the software and includes product, computer, personal and project attributes as follows:

1.4.1 Product Attributes

Required reliability (RELY):

It is used to express an effect of software faults ranging from slight inconvenience (VL) to loss of life (VH). The nominal value (NM) denotes moderate recoverable losses.**Data bytes per DSI (DATA):**

The lower rating comes with lower size of a database.

Complexity (CPLX):

The attribute expresses code complexity again ranging from straight batch code (VL) to relate code with multiple resources scheduling (XH).

1.4.2 Computer Attributes

Execution time (TIME) and memory (STOR) constraints

This attribute identifies the percentage of computer resources used by the system. NM states that less than 50% is used;95% is indicated by XH.]

Virtual machine volatility (VIRT)

It is used to indicate the frequency of changes made tothe hardware, operating system, and overall software environment. More frequent and significant changes are indicated by higher ratings.

Development turnaround time (TURN)

This is a time from when a job is submitted untiloutput becomes received. LO indicated a highly interactive environment, VH quantifies a situation when this time is longer than 12 hours.

1.4.3 Personal attributes

Analyst capability (ACAP) and programmer capability (PCAP): describe skills of the developing team. The higher the skills the higher the rating. Application experience (AEXP),language experience (LEXP), and virtual machine experience (VEXP): These are used to quantify the number of experiences in each area by the development team; more experience, higher rating.

1.4.4 project attributes

Modern development practices (MODP)

MODP deals with the amount of use of modern software practices such as structural programming and object-oriented approach.

Use of software tools(TOOL)

TOOL is used to measure a level of sophistication of automated tools used in software development and a degree of integration among the tools being used. Higher rating describes levels in both aspects.

The below table 5.2 gives information about the Schedule effects of expansion or compression , through the below table we can easily calculate the project cost.

Table 5.2: Schedule Effects of expansion or compression

Schedule effects (SCED): concerns the amount of schedule compression (HI or VH),or schedule expansion (LO or VL) of the development schedule in comparison to a nominal (NM) schedule.	VL	NM	HI	VH	XH
RELY	0.77	1.00	1.15	1.40	
DATA		1.00	1.08	1.16	
CPLX	0.70	1.00	1.15	1.30	1.65
TIME		1.00	1.11	1.30	1.66
STOR		1.00	1.06	1.21	1.56
VIRT		1.00	1.15	1.30	

Table 5.3: Uneven effects of expansion or comparison

TURN		0.87	1.00	1.15	1.30
ACAP	1.46	1.19	1.00	0.86	0.71
AEXP	1.29	1.13	1.00	0.91	0.82
PCAP	1.42	1.17	1.00	0.86	0.70
LEXP	1.14	1.07	1.00	0.95	
VEXP	1.21	1.10	1.00	0.90	
MODP	1.24	1.10	1.00	0.91	0.82
TOOL	1.24	1.10	1.00	0.91	0.83
SCED	1.23	1.08	1.00	1.04	1.10

Our project is an organic system and for intermediate COCOMO Effort = a * (KLOC)b*EAF

KLOC = 115

For organic system, a =2.4, b = 1.02EAF = product of cost drivers Effort

= 2.4 * (0.115) ^ 1.02 * 1.30

= 1.034

Programmer Month’s Time for development = C * (Effort) d

= 2.5 * (1.034) ^0.38

= 2.71 months Cost of programmer = Effort * cost of programmer per month

= 1.034 * 20000

= 20680

Project cost = 20000 +20680

= 40680

CHAPTER – 6

SYSTEM DESIGN

CHAPTER – 6

SYSTEM DESIGN

6.1 System Architecture

As mentioned earlier, predicting the rate of Dissolved Oxygen in water is an important role in determining the growth of fish in Aquaculture. In Aquaculture farms, the farmers usually grow the fishes in freshwater environments. Harvest and health of fishes and rate of DO in Water are affected by many factors/parameters liketemperature, PH, BOD and nitrate present in fresh water. So, in our proposed mechanism LSTM and Gated Recurrent unit network (GRU) hybrid model was used to estimate the rate of Dissolved Oxygen in water.

The workflow of system is shown in the below figure 6.1. It depicts how the inputs are given to the system and how they are processed using the hidden layers of LSTM and GRU and how the output is predicted. To start with, the water quality parameters like pH, Temperature, B.O.D, and Nitrate are given as input, the next step includes processing of the data. In the proposed system we adopted LSTM and GRU hybrid model to predict the rate of Dissolved Oxygen in water. LSTM and GRU (Gated Recurrent Unit) are type of RNN networks that makes use of less tensor operations. GRUwas introduced in the year 2014 And LSTM was introduced in 1997.

The dataset we considered consists of water quality parameters in specific locations of India [6]. The information used in this study was acquired from many rivers located across India. Over the course of five years, from 2005 to 2014, 9624 samples were gathered from various Indian states. There are seven critical para meters in the dataset that must be taken into account. The Indian government gathered data to guarantee that the water given was of good quality.

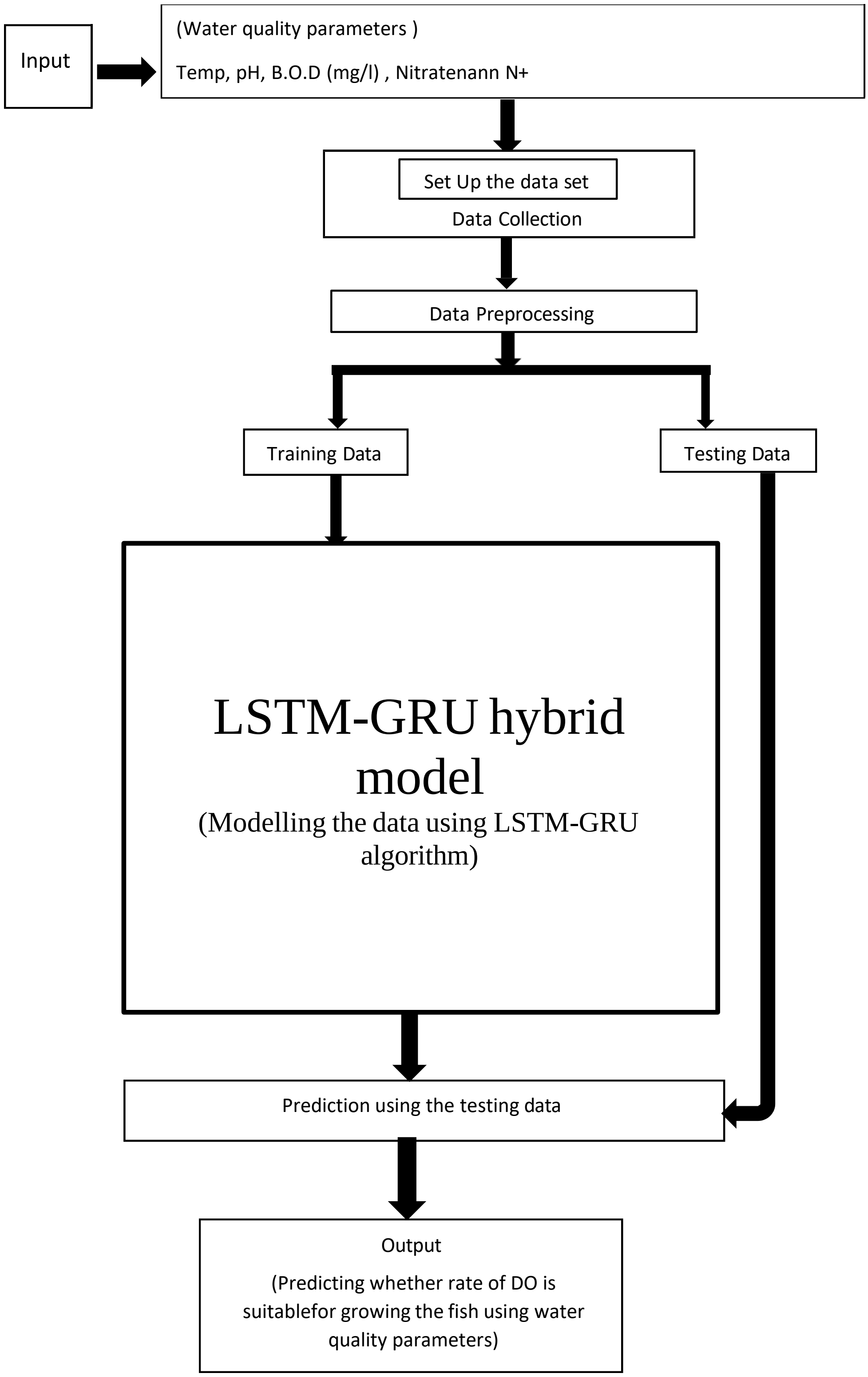
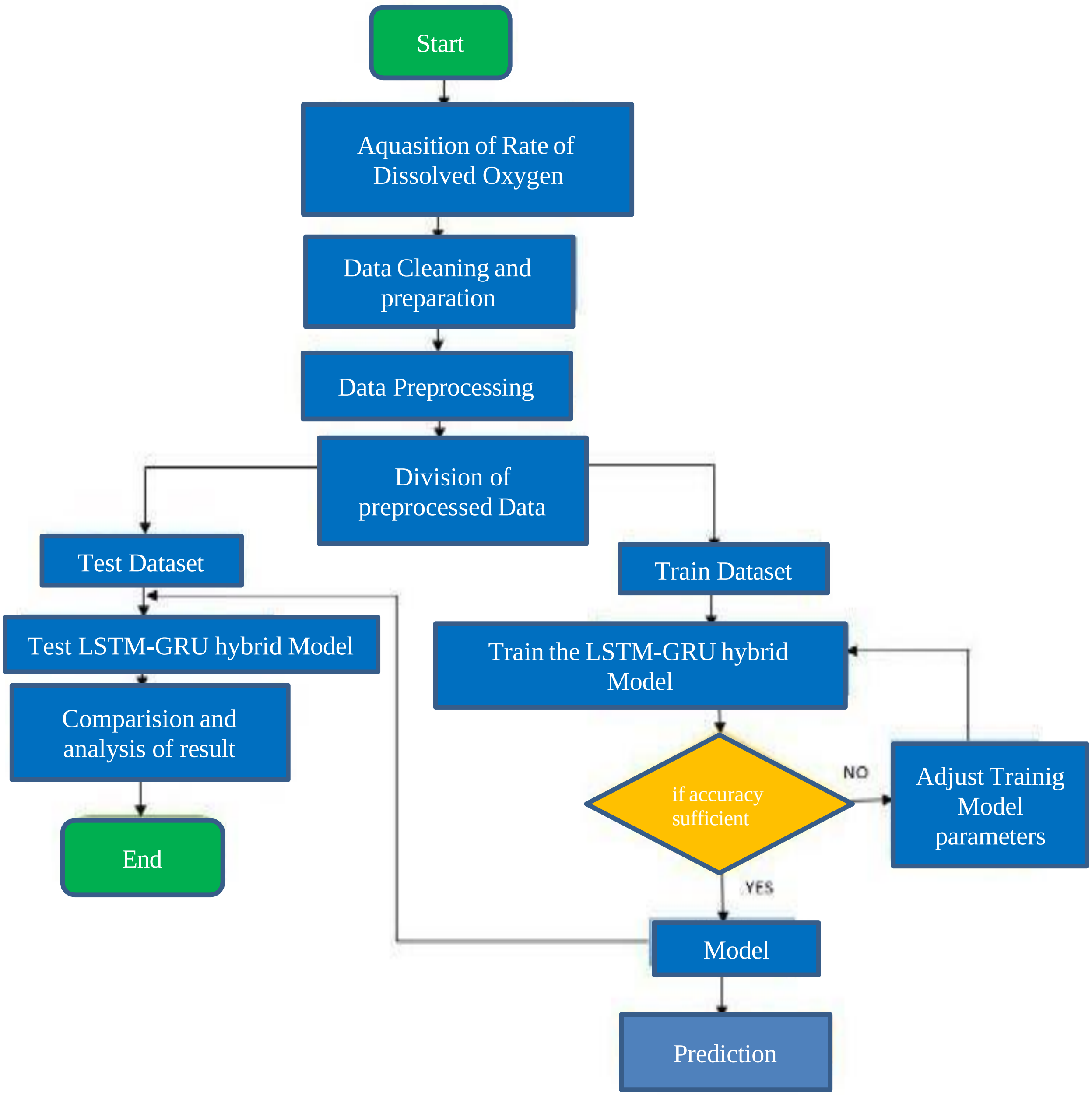


Fig 6.1: prediction of dissolved oxygen in aquaculture ponds using LSTM-GRU model

6.2 Flow Chart

Th process flow for the proposed methodology is depicted in figure 6.2 as shown below, as Initially the data preprocessing is done and split the data into train data and test data, with the help of train data our model is trained and tested for accuracy with test data.



.

Figure-6.2: Flow chart of Prediction of Dissolved oxygen using LSTM and GRU hybrid Model.

Data cleaning and preparation is performed and then the data preprocessing which involves normalization, scaling is done. The model is created using the preprocessed data. The preprocessed data is divided into train and test data. Fit/ Train the model using training data, matrix of input parameters and array of output, epoch and batch size. Next, the model is evaluated by using the test data and the performance indicators are calculated. The next hour values are predicted, and if they exceed the range of parameters specified, an alert is sent as in below figure 6.3

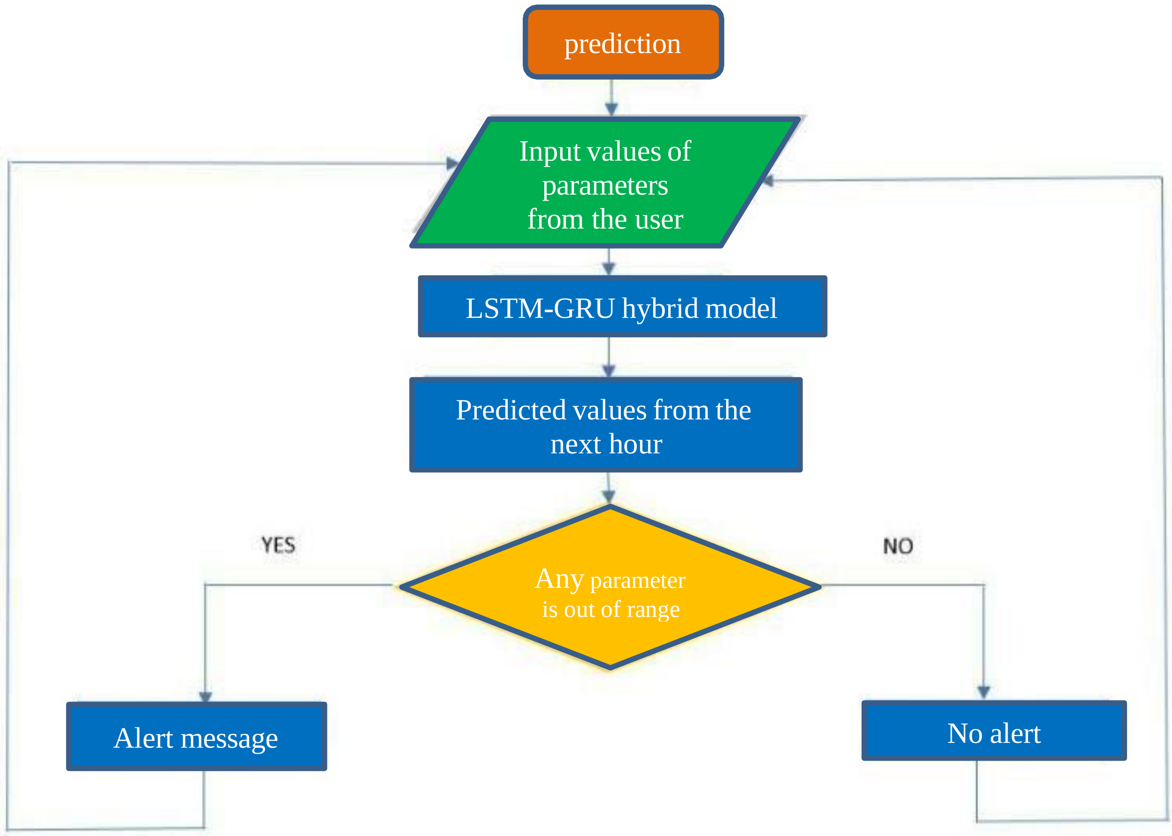


Figure-6.3: Flowchart of Prediction Subroutine

6.3 Design Overview

UML combines best techniques from data modeling (entity relationship diagrams), business modeling (work flows), object modeling, and component modeling. It can be used with all processes, throughout the software development life cycle, and across different implementation technologies UML has synthesized the notations of the Booch method, the Object modeling technique (OMT) and object-oriented software engineering (OOSE) by fusing them into a single, common and widely usable modeling language. UML aims to be a standard modeling language which can model concurrent and distributed systems.

The different types of UML diagrams are shown in figure 6.4

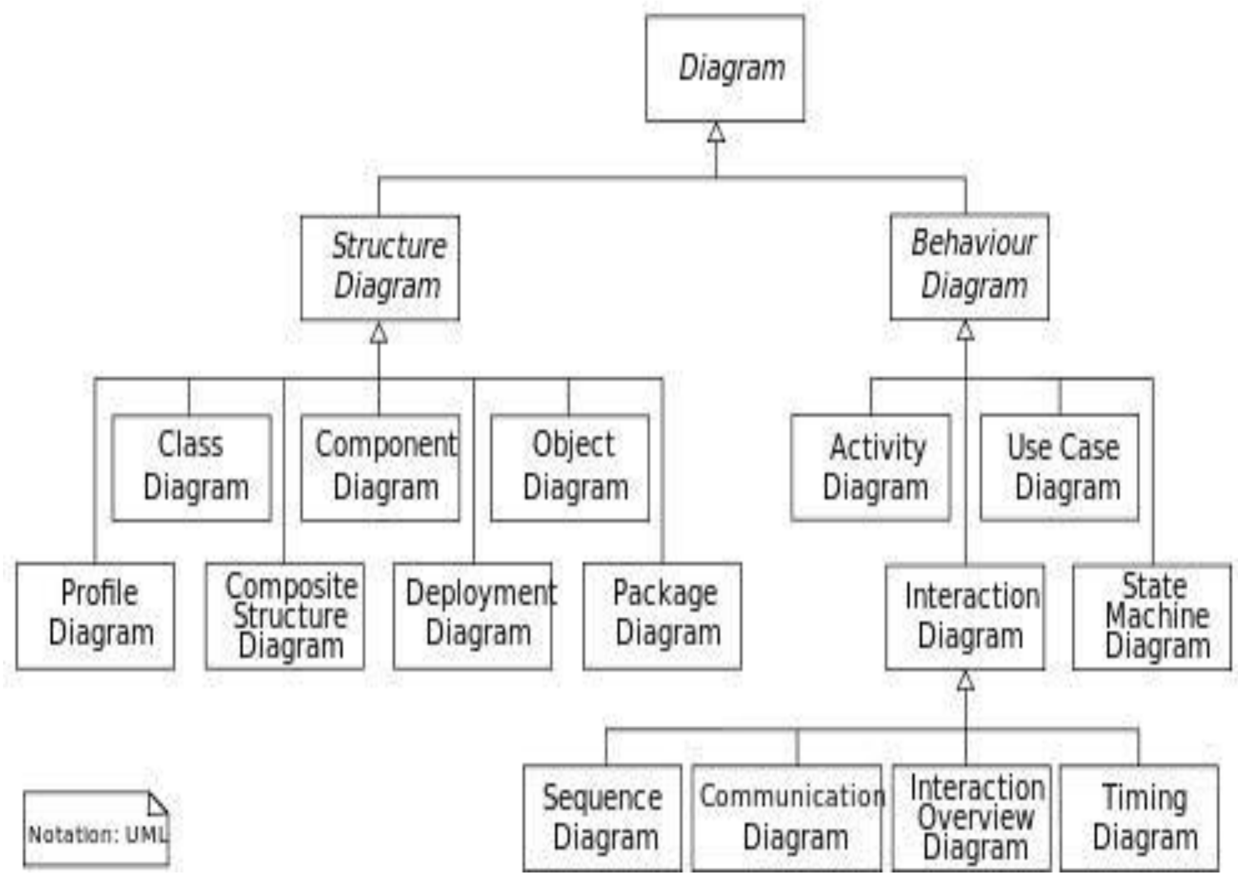


Figure-6.4: Types and categories of UML Diagram

6.4 UML Diagrams

The Unified Modeling Language (UML) is used to specify, visualize, modify, construct and document the articrafts of an object-oriented software intensive system under development. UML offers a standard way to visualize a system's architectural blue prints, including elements such as:

1. Actors
2. Business process
3. (Logical) Components
4. Activities
5. Programming language statements
6. Database schemas and reusable software components.

The unified modelling language allows the software engineer to express an analysis model using the modelling notation that is governed by a set of syntactic semantic and pragmatic rules. A UML system is represented using 5 different views that describe the system from distinctly different perspective. Each vie is defined by a set of diagrams, which is as follows:

User Model View:

1. This view represents the system from the user's perspective.
2. The analysis representation describes a usage scenario from the end-user's perspective.
3. The UML user model view encompasses the models which define a solution to a problem as understood by the client stakeholders.

Structural Model View:

1. In this model the and functionality are arrived from inside the system.
2. This model view models the static structures.

Behavioral Model View:

It represents the dynamic of behavioural as parts of the system, depicting the interactions of collection between various structural elements described in the user model and structural model view.

Implementation Model View:

Implementation view is also known as Architectural view which typically captures the enumeration of all the subsystems in the implementation model, the component diagrams illustrating how subsystems are organized in layers and hierarchies and illustrations of import dependencies between subsystems.

Environmental Model View:

These UML model describe both structural and behavioural dimensions of the domain or environment in which the solution is implemented. This view is often also referred to as the deployment or physical view.

6.4.1 Use case Diagram

A use case diagram at its simplest is a representation of a user's interaction with the system that shows the relationship between the user and the different use cases in which the user is involved. Here in the below figure 6.5 Supervisor is involved in the use cases View prediction result, View Visualization, View Alert and managing the application.

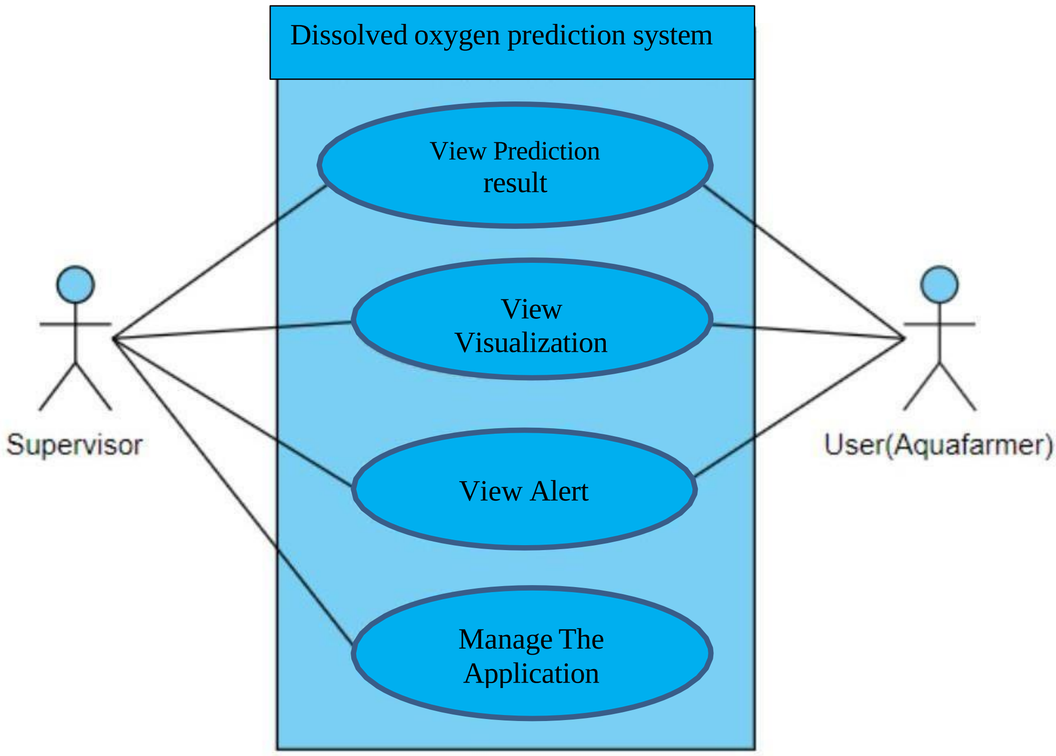


Figure-6.5: Use Case Diagram of prediction of Dissolved oxygen

6.4.2 State diagram

Below figure 6.6 represents the transition between various states of the prediction system. It gives an idea about the various states and the events involved from data collection to generating an alert when any parameter is not in range.

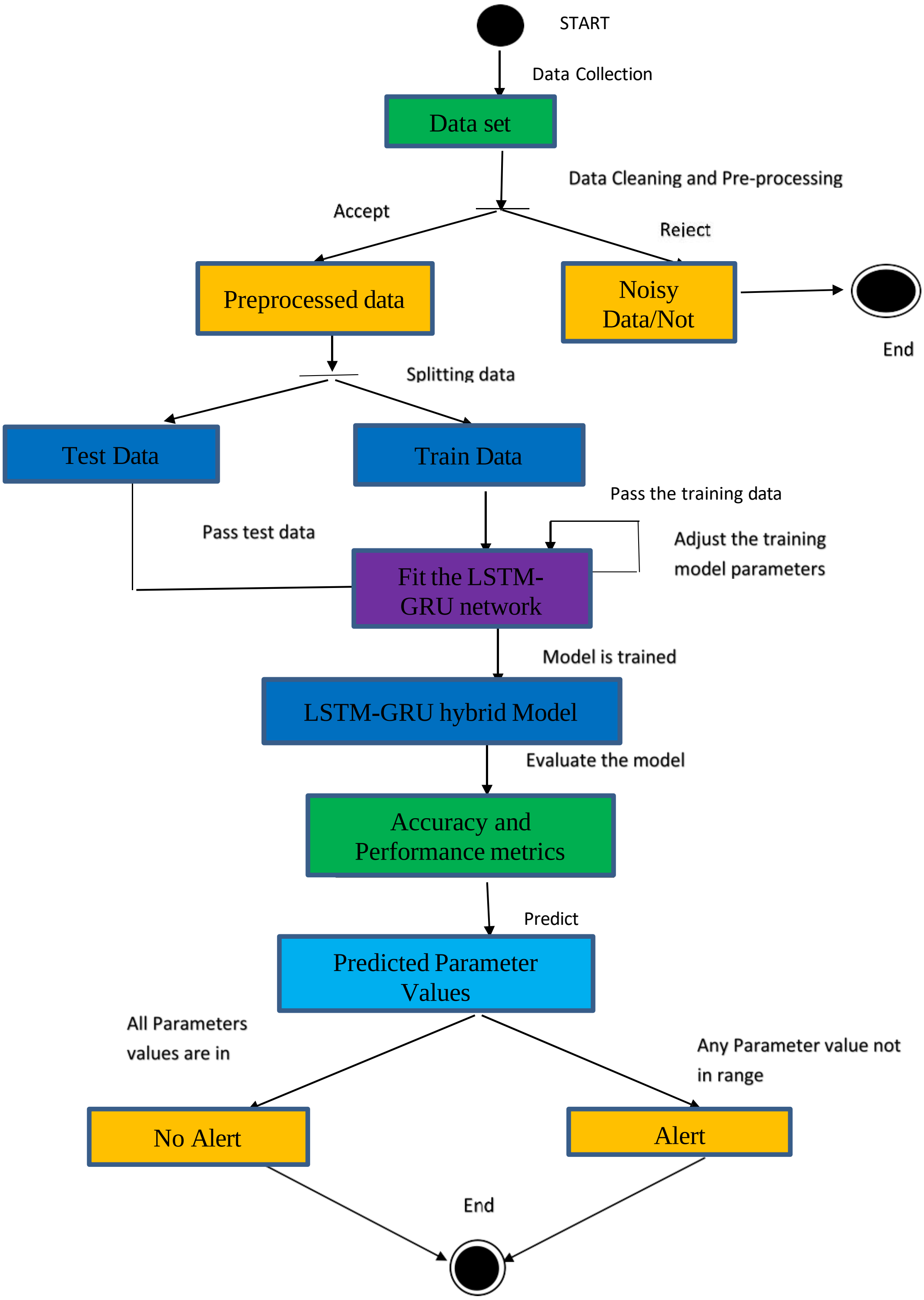


Figure-6.6: State Diagram of Prediction of Dissolved oxygen in Aquaculture System

6.4.2 Sequence Diagram

Sequence Diagram is interaction diagram that detail how operations are carried out. It shows how they interact over time and they are organized according to object (horizontally) and time(vertically). From the data set data is extracted and processed, respectively. Then DL model is trained, and input data is given by the user and future prediction is made. User can view the prediction results and visualization through the Web application. The sequence diagram for water quality prediction is shown in the below figure 6.7

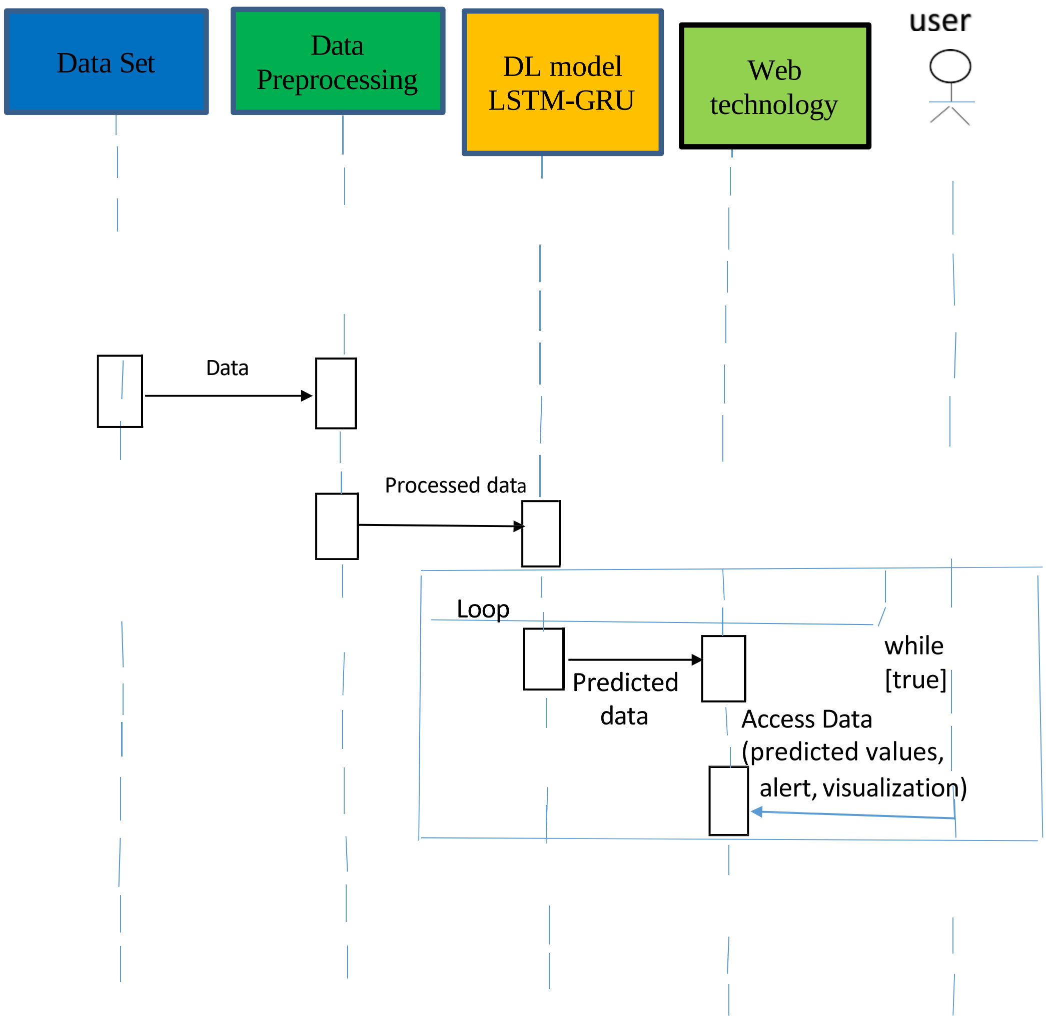


Figure-6.7: Sequence Diagram of Dissolved Oxygen Prediction System

CONCLUSION

Aquaculture plays an important role in food and nutrition industry. This proposes a hybrid prediction model to solve the challenge of poor DO concentration forecasting accuracy in aquaculture management. The hybrid model was designed by combining the LSTM with GRU. Real-time prediction of dissolved oxygen content is the premise and basis for achieving precise control of recirculating aquaculture. To solve the problems of traditional prediction model that low accuracy, poor stability noise in the local characteristic data collected from water quality, the model which based on LSTM-GRU has been proposed.

Its effectiveness has been verified in practical applications. It can predict the dissolved oxygen content, thereby reducing and avoiding unnecessary losses, such as fish mortality rate, caused by excessive or insufficient dissolved oxygen content. The experimental study shows that LSTM-GRU hybrid model performs better than other models.

REFERENCES

[1] Kisi, O., Alizamir, M., & Docheshmeh Gorgij, A. (2020). Dissolved oxygen prediction using a new ensemble method. *Environmental Science and Pollution Research*, *27*(9), 9589-9603.

[2] Ren, Q., Wang, X., Li, W., Wei, Y., & An, D. (2020). Research of dissolved oxygen prediction in recirculating aquaculture systems based on deep belief network. *Aquacultural Engineering*, *90*, 102085.

[3] Huan, J., Cao, W., & Qin, Y. (2018). Prediction of dissolved oxygen in aquaculture based on EEMD and LSSVM optimized by the Bayesian evidence framework. *Computers and electronics in agriculture*, *150*, 257-265.

[4] Eze, E., & Ajmal, T. (2020). Dissolved oxygen forecasting in aquaculture: a hybrid model approach. *Applied Sciences*, *10*(20), 7079.

[5] Moon, J., Lee, J., Lee, S., & Yun, H. (2022). Urban River Dissolved Oxygen Prediction Model Using Machine Learning. *Water*, *14*(12), 1899.

[6] Cao, S., Zhou, L., & Zhang, Z. (2021). Prediction of dissolved oxygen content in aquaculture based on clustering and improved ELM. *IEEE Access*, *9*, 40372-40387.

[7] Cao, S., Zhou, L., & Zhang, Z. (2021). Prediction of dissolved oxygen content in aquaculture based on clustering and improved ELM. *IEEE Access*, *9*, 40372-40387.

[8] Shi, P., Li, G., Yuan, Y., Huang, G., & Kuang, L. (2019). Prediction of dissolved oxygen content in aquaculture using Clustering-based Softplus Extreme Learning Machine. *Computers and electronics in agriculture*, *157*, 329-338.

[9] Cao, X., Liu, Y., Wang, J., Liu, C., & Duan, Q. (2020). Prediction of dissolved oxygen in pond culture water based on K-means clustering and gated recurrent unit neural network. *Aquacultural Engineering*, *91*, 102122.

[10] Eze, E., & Ajmal, T. (2020). Dissolved oxygen forecasting in aquaculture: a hybrid model approach. *Applied Sciences*, *10*(20), 7079.

[11] Li, W., Wu, H., Zhu, N., Jiang, Y., Tan, J., & Guo, Y. (2021). Prediction of dissolved oxygen in a fishery pond based on gated recurrent unit (GRU). *Information Processing in Agriculture*, *8*(1), 185-193.

[12] Feng, Huanhuan, et al. "Evaluation of dynamic stress level and physiological change for live salmon in waterless and lowtemperature transportation." Aquaculture 544 (2021): 737128.

[13] https://data.gov.in/catalog/water-quality-data-river-ganga Accessed on May 5,2021.

[14] Gopi, Arepalli Peda, and K. Jairam Naik. "A Model for Analysis of IoT based Aquarium Water Quality Data using CNN Model." 2021 International Conference on Decision Aid Sciences and Application (DASA). IEEE, 2021.

[15] K Jairam Naik (2021),"An Adaptive Push-Pull for Disseminating Dynamic Workload and Virtual Machine Live Migration in Cloud: Load Balancing & VM Live Migration in Cloud Computing", International Journal of Grid and High-Performance Computing