Turkish Online Journal of Qualitative Inquiry (TOJQI)
Volume 12, Issue 6, June 2021:321-328

Research Article

Prediction of Water Quality System for Aquaculture using Machine Learning

Kiran babu T S[1], Manoj Challa[2]

Abstract

Indian aquaculture and fishing is one of the key stages of food production that provides healthy food security which also provides employment, livelihoods and contribute to the growth of Indian agriculture. Like all other aquatic organisms they also have a wide range of tolerance, so fish farming of any species requires certain conditions that must be met. Water quality is a critical factor in the processing of aquatic organisms. In this paper, we created an application using ML to predict water boundaries and monitor fish ponds. The water level is predicted in an hourly manner to ensure the growth and survival of aquatic life. The web application is built using Flask to alert the user to critical situations. The impact of water parameter changes can be effectively treated if the information is analyzed and water quality is expected ahead of time.

Keywords: *Aquaculture, Machine Learning, LSTM, Square Error*

[1] Department of CSE, CMR Institute of Technology, Bengaluru India, kiran.ts@cmrit.ac.in
[2] Department of CSE, CMR Institute of Technology, Bengaluru India, manoj.c@cmrit.ac.in

Introduction

Studies have shown that over the past three decades the use of fish has been widely used. This study serves as a contribution to increasing and stabilizing fish production. Fish farming means a growing variety of marine species such as fishing fish, sport fish, ornamental fish, mussels, mussels, crustaceans, seaweed, algae, and fish eggs to harvest, breed, and raise in various places such as rivers, lakes, sea, and lakes.

There are various restrictions on marine life. One parameter depends on the other parameter and indirectly affects aquatic organisms. In fish farming, acid and alkaline water should be kept between 6 and 8. If the pH value is below 4.5, the fish will die. Aqua farmers rely on self-assessment by knowing the limits of water, time-consuming and inaccurate as boundaries can change over time. To overcome this problem, modern technology must be introduced into agriculture. Therefore, it increases productivity and reduces losses by constantly monitoring water quality parameters. The water parameters that need to be continuously checked are Temperature, Melted Oxygen, PH, Turbidity, Salt, etc.

Our main goal is to increase the productivity of marine fish and sustainably save their lives. To accurately predict water quality, ML Algorithms such as LSTM and ARIMA models are analyzed and used as the most appropriate. In this project water boundary prediction is done on an hourly basis. e.g., predict next hour values based on previous data. When the predicted water quality appears to exceed the critical conditions it is immediately sent to the user to take the necessary precautionary measures. The web application is built using Flask to display alerts to users and data detection of various targeted water restrictions.

This paper focuses on the importance of water quality for aquaculture. It helps aqua farmers to produce high-quality fish, which in turn helps the economy of the agricultural sector. Machine learning helps to provide better, more accurate, and faster forecasts of water quality based on previously collected data. Predictability analysis can help catch up relationships between multiple variables that can help assess risk with a specific set of conditions. LSTM NN is regarded as the best and provides very clear information for detecting and monitoring water quality.

Related Work

Juntao Liu et al [1] are using the Simple Recurrent Unit (SRU) to create an accurate and automated system for predicting cultural water quality. It focuses on predicting water quality based on pH and temperature. The advanced progression method is used where the abstract data is removed using a method called a fixed limit. Finally, it is presented with the SRU model and compared with the RNN model which shows that SRU has high accuracy in simultaneous difficulty.

A prediction model based on Smooth Support Vector Machine (SSVM) was proposed by Wijayanti Nurul Khotimah [2] to predict water quality in water. The RMSE value obtained is 0.0275 indicating that SSVM is an effective means of predicting water quality in water.

J. Wang et al. [3] investigates the characteristics of strong interactions with the correction of water quality parameter information and the disappearance of gradient and gradient

explosions caused by data training of the traditional RNN network model, etc. The structure is shown on this page.

Dong Yao, Lei Cheng, QiuXuan Wu, Gong Zhang, Bei Wu, YuQing He [4] investigates how to analyze and predict the quality of fishing using an electrochemical sensor array such as melted oxygen, pH ammonia, and nitrogen carried by an unauthorized air vehicle.

(UAV) proposed. A floating T-shaped floating structure of a UAV can see the rise and fall of water. The sensors in the structure enable real-time detection in the floating UAV system. Sensor data is transmitted backstage via the wireless network. Thereafter a series of time analyses are used to evaluate and predict water quality data. Finally, the conclusion is shown in the application process. The results show that the system can assess the quality of fishing water efficiently and has high accuracy (95.1%).

Proposed System

The proposed system architecture shows the complete working of the system which can be divided into three main stages i.e. Data Pre-processing, Model development, and prediction of results.The Figure 1 show the proposed system architecture for water quality predicition for aquaculture.

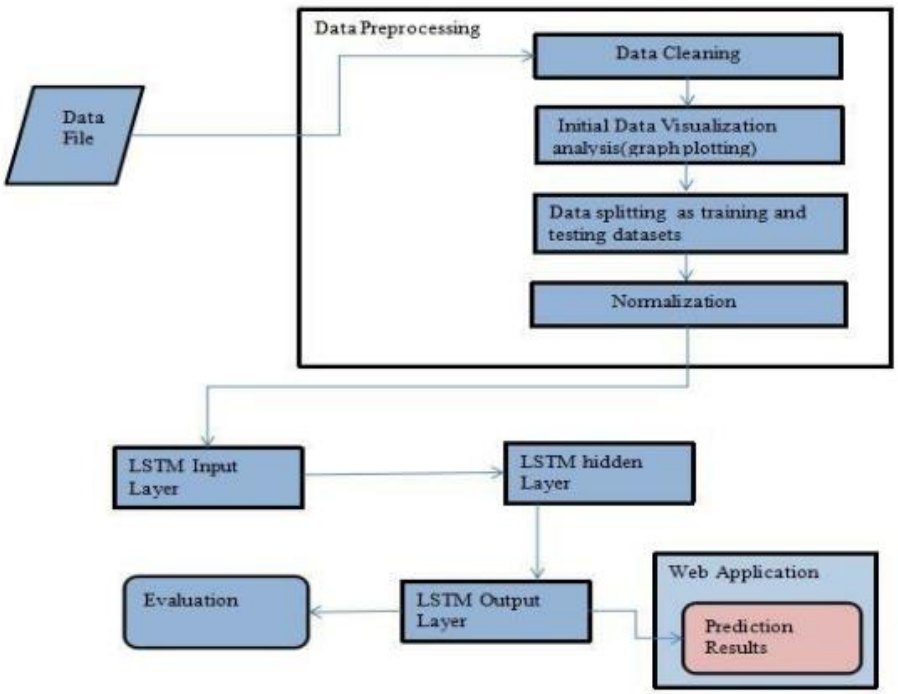

Figure:1 Proposed System Architecture

Data Preprocessing:

In this stage first, Data cleaning has been done, meaning data is filtered and altered such that unwanted data will be parted out. Then the filtered data is analyzed and rifted into training and testing data sets. After that data is normalized to bring the values of all variables to the same scale without deforming differences in the ranges of values.

Model Development Process:

LSTM(Long Short Term Memory) model is used to design this prediction system. Development of the system has occurred in two phases. The 1st phase model is built using the training data set. In the 2nd phase, the model is evaluated to assess the performance and accuracy of the system.

Prediction of Result:

In this stage, the model is run for many different unseen sets of values to assess the likely future behavior of the system. And the result is shown on a web application for easy understanding.

Implementatıon

Various tools and modules of python are used to create the proposed model. Jupyter Notebook was used to work with our project. Many libraries of Python are used to create our water quality prediction model and produce some visualizations for the same such as Pandas, Numpy, Keras, Scikit-learn. Flask framework is used to build the web application. The various modules involved in the project are:

1) Data Cleaning and Preparation

The dataset is stored in a data frame. The rows with NaN values are dropped from the data frame. The data has been resampled and the daily average has been calculated. Then, the resampled data is saved in another CSV file. The resultant dataset is used for the model training.

2) Data Pre-processing

The dataset has been split into training and testing datasets. The training set is 80%, whereas the testing set is 20% of the original dataset. MinMax normalization is used to transform the values of both training and testing datasets in the range of -1 and 1 as the ranges of the parameters are different.

3) Creating the model

To create the LSTM model, first, we create a Sequential model. Next, add a visible layer with6 input, a hidden layer with 16 LSTM neurons, and an output layer that makes a four values prediction. ReLU activation function is used for the LSTM neurons to help in accomplishing better execution, helps to overcome the vanishing gradient problem, enabling models to adapt faster and perform better.

4) Compiling the model

The model is compiled using the loss function and the optimizer before training the model. Mean Squared Error is used as the loss function and Adam as the optimizer which is anoptimization technique that can be utilized instead of the classical stochastic gradient process to update network weights repetitively based on our training data.

5) Training the model

The model is trained with 100 epochs or iterations. A batch size of 1 has been used while training. EarlyStopping callback is used so that the model can terminate itself when the monitored feature has stopped improving.

6) Model Evaluation

MSE, RMSE, and R2 score between the actual values and the predicted values are calculated for evaluation and these values are the best for the LSTM model than other models for our application. To improve the accuracy, the LSTM model can be re-trained.

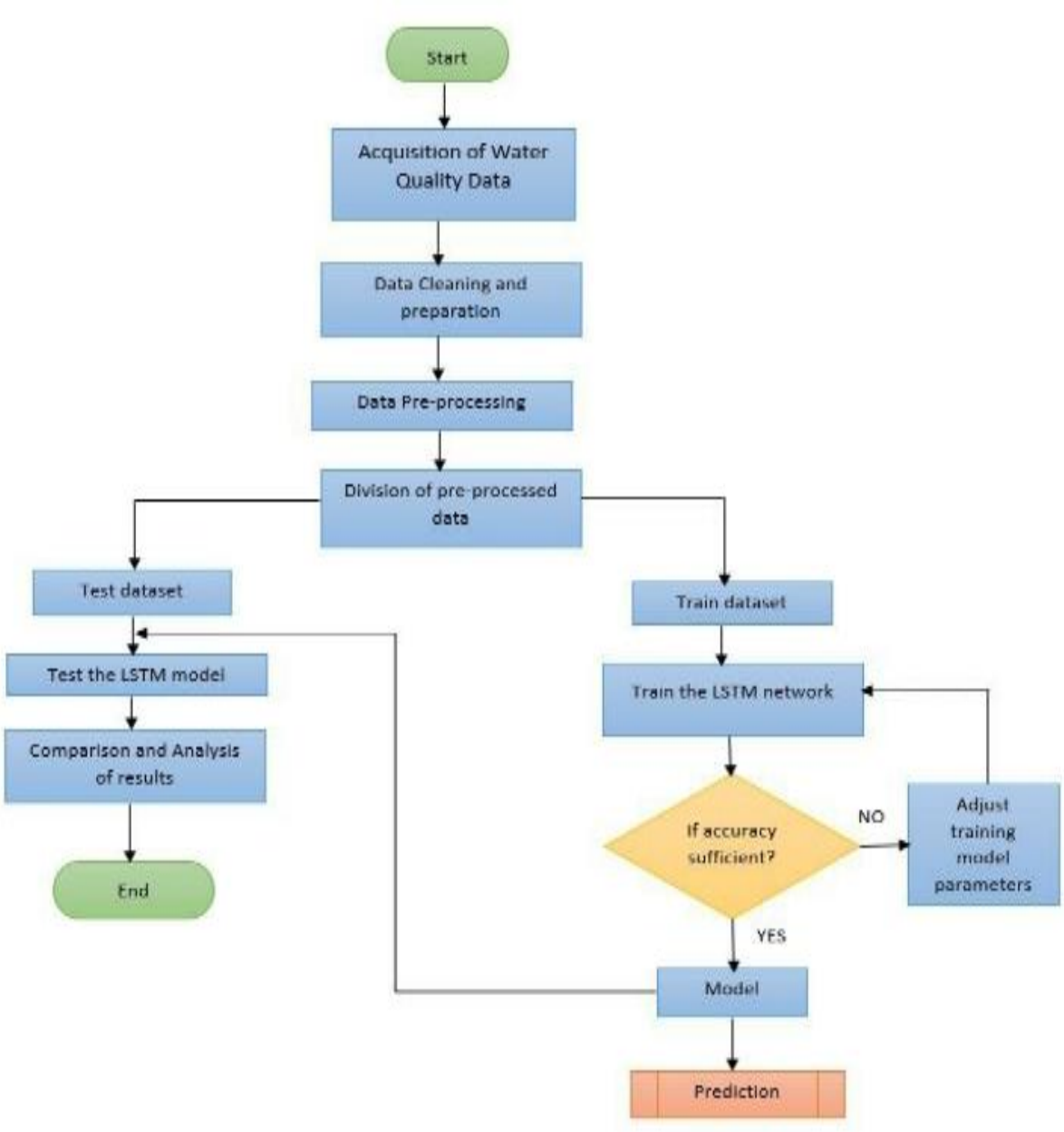

Figure: 2 Flowchart of the Water Quality Prediction Model

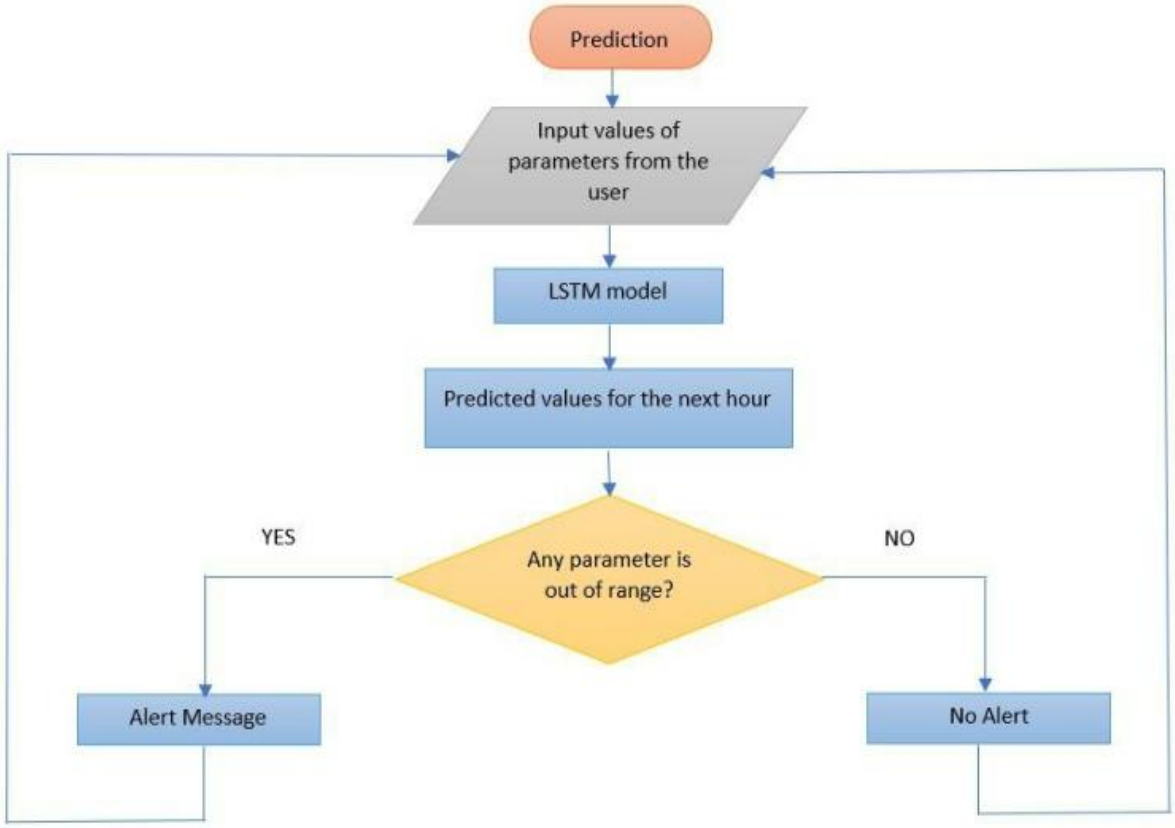

Figure: 3 Flowchart of Prediction Subroutine

7) Prediction

The LSTM model is used to predict the values of water parameters such as pH, Dissolved Oxygen, Temperature, and Turbidity.

The below flow charts (Figure 2 and Figure 3) show the step-by-step execution implemented at the backend and frontend of the system.

Results And Dıscussıon

Discussion

An application is developed were the parameter values can be entered to the system. The Dataset is collected from kaggle.com for predicition water for an Aquaculture.

Once the values of parameters are entered, the predicted values for the next hour are shown to the user using the backend algorithm for prediction and an appropriate message is displayed. When all the predicted parameters go out of range alert message is shown, When the visualization button is clicked, the graphs for the various parameters are shown as in Figure 4, which can help the user understand the changes in parameters over time.

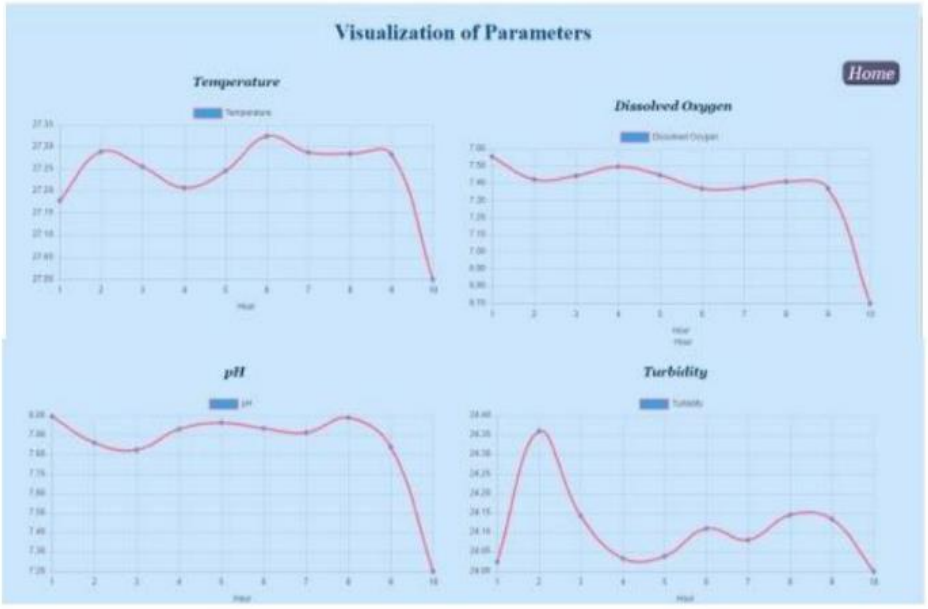

Figure:4 Visualisation of Parameter

Conclusıon And Future Scope

Water quality is an essential factor for aquaculture. In this project LSTM model is developed to predict the value of pH, Temperature, Turbidity, Dissolved Oxygen which are the main parameters for water quality in determining the aquatic species habitations.

The dataset fed into the model is the recorded information of water quality parameters which are obtained from the USGS website which monitors the stations and maintains the database and continuous values can be generated by the USGS REST service URL tool.

The datasets show recorded values of the water quality parameters like temperature, DO, pH and turbidity from the year 2014 to 2019(till 18th Feb) for Georgia State with an hour of intervals. To remove the noisy and inconsistent data, the data cleaning, and data scaling (normalization) are carried out along with the graphical presentations and exploratory analysis. To improve the efficiency of the model, a few re-enactments and parameter determination are completed. The model shows that the RMSE for the LSTM is 0.03 and the prediction results are proper. Our model is fit for multivariate parameters and the predictions are done in an hourly manner. The web application is built to make it user-friendly. Visualization of water parameters gives good insight for understanding the parameter changes for past hours.

Future Scope

Developing a mobile app is more convenient for alerting the aquafarmers when critical conditions exceed water parameter values and the app is accessible through phones. Giving suggestions for aquafarmers based on seasons how to maintain water quality for the surveillance of fishes. Water quality monitoring for aquaculture using IoT and predicting the water quality using ML.

References

Juntao Liu, Chuang Yu, Zhuhua Hu," Automatic and Accurate Prediction of Key Water Quality Parameters Based on SRU Deep Learning in Mariculture" 2018 IEEE International Conference on Advanced Manufacturing (IEEE ICAM 2018)

Wijayanti Nurul Khotimah," Aquaculture Water quality prediction using Smooth SVM", IPTEK, Journal of Proceeding Series, Vol. 1, 2014 (eISSN: 2354-6026).

Sundarambal, P., Shie-Yui, L., & Pavel, T. (2009). Development of a neural network for dissolved oxygen in seawater. Indian Journal of Marine Science, 38 (2), 151-159.

Lee, Y. J., & Mangasarian. (2001). A smooth support vector machine. Journal of Computational Optimization and applications, 20: 5 – 22.

J. Wang, F. Xiang, F. Qiu, H. Wang, H. Liu. Research progress of "River Water Quality Parameters Prediction Method Based on LSTM-RNN Model" Guild of Environmental Sciences, Vol.37, No.4, 63-67, 2018.

H. B. Wang, Z. P. Yang, Q. Yu, T. J. Hong, X. Lin. Online reliability time series prediction via convolutional neural network and long short-term memory for service-oriented systems. Knowledge-Based Systems, Vol.159, 132-147, 2018.

Dong Yao, Lei Cheng, QiuXuan Wu, Gong Zhang, Bei Wu, YuQing He, "Assessment and Prediction of Fishery water quality using electrochemical sensor array carried by UAV**",** Engineering Research Center of Metallurgical Automation and Measurement Technology of Ministry of Education, Wuhan University of Science and Technology, Wuhan, China

Abe, S. (2010). Support Vector Machines for Pattern Classification 2nd edition. New York: Springer- Verlag.