Department of Computer Science

Dual Degree(B.Tech + M.Tech) Computer Science

Academic Year 2021 - 2022

Prediction of Dissolved Oxygen in water

Shiv Shankar
(CED17I026)
Under
Dr. Venkatesh Pandiri

This report wish to describe my work done in Design Project

IIITDM Kancheepuram
Department of Computer Science
Vandalur-Kelambakkam Road
Chennai-127
Tamil Nadu
India
T: +91-44-2747 6300
F: +91-44-2747 6301

Abstract

This report wish to describes my Design Project work at the Indian Information of Information Technology Design and Manufacturing in Kancheepuram , and it is divided into four chapters.

Chapter One summarizes the projects and It gives us the introduction and need of the project.

Chapter Two provides idea about preparing the data for machine learning (Machine Learning or Cleaning data).

Chapter Three Dealt with the Back end or hosting server and make prediction online using API the work done in Nodejs.

Chapter Four In this chapter we will finally develop a web App for end user.

Acknowledgements

Working on this project is very good experience of me. During this i learned a lot and all my courses made equal and big contribution in completing this project, Programming, Data cleaning and all maths courses helped me a lot.

Further on I want to thank stackoverflow, github and all programming site because these community helped me a lot while i stuck in any error.

I certify that the work presented in the dissertation is my own unless referenced.

Signature: 

Date: 4/Dec/2021

Contents

Chapter 1

Introduction

A crucial activity in prawn farming is monitoring prawn pond water quality. Variables such as dissolved oxygen (DO), pH, temperature, and salinity are commonly monitored using sensors Temp sensor pH sensor and many other sensor are widely used by small farmers and pond owner to monitor water quality for fishes and agriculture related work. Sensor monitoring comes with challenges such as: purchasing and main- taining sensors is costly, gathering readings with hand-held sensors over many ponds is time consuming, readings can be incorrectly logged, and sensors can fail. Such problems can be mitigated by reducing the number of sensors required. A sensor can be made redundant if its associated variable can be predicted from other sensor data In this study a recurrent neural network (RNN) , a linear regression model,are used to predict dissolved oxygen from pH and water temperature sensor readings.

Aquaculture consists of the set of activities, knowledge and techniques for the breeding of aquatic plants and some species of animals. This activity has a great importance in economic development and food production. Continuous monitoring of the physical, chemical and biological parameters of pond water helps not only to predict and control the negative conditions of aquaculture, but also to avoid environmental damage and the collapse of the production process. The monitoring of physical and chemical variables such as: oxygen, temperature and pH in water are vital to maintain adequate conditions and avoid undesirable situations that may lead to the collapse of aquaculture systems

1.1 Aims and Objectives

The aim of this project is to get the main factor of water Dissolved Oxygen get predict by some given properties of water using Machine Learning Model and made them accessible to end user without any hustle so that they can access in Mobile device like Application of iOS and Android and web App.

Here is the list of the necessary and complete set of objectives that we will need to achieve in order to satisfy our above listed aim using modern day technologies :.

1. Undertook a relevant background study to identify existing work in the area, and to identify appropriate techniques which can be adopted to produce a

solution in this project. like is there any model which can help us to do thing easily

2. In the task first we have and we want data of temp and ph so that we can make a prediction of dissolved oxygen but before making prediction data cleaning we need to done change format like json and other compatible with mobile app or web app.

3. We have to implement one Tensorflow model by which we can make prediction and calculate approximated output given any particular input we have to use Tensorflow Lite (For Mobile App Development) and Tensorflow.js for web app Development.

4. After getting model ready and data uploaded to cloud we have to make data accessible to end user like farmer and fish farmer so that they can use prediction to make aquatic and agriculture life better. So for this purpose we have to make one Web App and Mobile App with very easy interface so everybody can use it.

5. In Last all the blocks are ready now we have to combine them all and make one final product which can make good predictions.

1.2 Project Approach

The Project will used many new and advanced technology which will make good and better product finally and which will make significant impact on society so we will follow all the above aim and make final product good as all the technology we will use are new so we have to handle thing by learning and doing. Tensorflow is very good for advanced Machine Learning thing and Java Script is good for web and cross platform native mobile application. The problem which we will face is that the data which we are using is having the temperature between 25-31°C and ph is of normal water (6-8) so, making prediction out of this bound will give us error in prediction like for very cold water (o-10 °C) it will give error as we have to collect more data in this range and train the model.

1.3 Dissertation Outline

Below is the task and the chapter in which detailed description of each task is included chapter wise and summary of the chapter included in Dissertation Outline.

Chapter 1, already talk about what the project is all about where Machine Learning and Web Model will be used.

Chapter 2, Data Computation, cleaning and make inference from data for predictions.

Chapter 3, One website Users.

Chapter 4, It shows about out testing and App and web App for are all the things working fine ?

Chapter 5, Conclusion of our Work and Future.

Chapter 2

Machine Learning and Data Cleaning (Or Approach)

In this section we will talk about cleaning data using pandas and convert clean csv data into json using online csv to json converter and we will use sci-kit and tensorflow for two machine learning model Linear Regression and Artificial Neural Network in both Tensorflow and Sci-kit and later we will decide which model to use. In below subsection. first we will write little description of Data Cleaning and ML.

2.1 Data Cleaning

Data cleaning is the process of preparing data for analysis by removing or modifying data that is incorrect, incomplete, irrelevant, duplicated, or improperly formatted. This data is usually not necessary or helpful when it comes to analyzing data because it may hinder the process or provide inaccurate results. There are several methods for cleaning data depending on how it is stored along with the answers being sought. Data cleaning is not simply about erasing information to make space for new data, but rather finding a way to maximize a data set's accuracy without necessarily deleting information. For one, data cleaning includes more actions than removing data, such as fixing spelling and syntax errors, standardizing data sets, and correcting mistakes such as empty fields, missing codes, and identifying duplicate data points. Data cleaning is considered a foundational element of the data science basics, as it plays an important role in the analytical process and uncovering reliable answers. Most importantly, the goal of data cleaning is to create data sets that are standardized and uniform to allow business intelligence and data analytics tools to easily access and find the right data for each query.

2.1.1 Pandas and Numpy

For the purpose of data cleaning we will use Pandas and Numpy as it is very easy to handle data in pandas dataframes and Numpy lib in python is very good library to handle array and manipulate array using advanced array methods.

In this Project we have csv data of water with following properties

STATION CODE,LOCATIONS,STATE,Temp,D.O. (mg/l),PH,CONDUCTIVITY (µmhos/cm),B.O.D. (mg/l),NITRATENAN N+ NITRITENANN (mg/l),FECAL COLIFORM (MPN/100ml),TOTAL COLIFORM (MPN/100ml)Mean,year.

But the problem with this data is that the temp, ph and D.o is having some values which is not in scale and for example ph above 14 and and have value NAN in the temp, ph etc and the which we will need only temperature, ph and dissolved oxygen for our project so we will get all these clean data in new csv.

2.1.2 Machine Learning and Prediction

Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves.

The process of learning begins with observations or data, such as examples, direct experience, or instruction, in order to look for patterns in data and make better decisions in the future based on the examples that we provide. The primary aim is to allow the computers learn automatically without human intervention or assistance and adjust actions accordingly.

Below is the code of Predictions using Our data using Tensorflow and Sci-kit Learn and data clearning in pandas also done in same code.

* Note - full code for machine learning model training can be found https://github.com/sivsnkr/design-project/blob/main/SequentialApi.ipynb

2.1.3 Tensorflow or Sci-kit Which is Best ?

When it's come to which model is best the answer is might complicated because which one is suitable for our App or website is the main reason to think about.

Sci-kit learn is a great library when it comes to simplicity and handling data sets on the other hand the tensorflow is slightly complicated because in tensorflow we have to deal with session to print and make calculation and also it deals with tensors so it is little bit difficult.

But Tensorflow is a great framework for machine learning enthusiast beacuse it provide vast variety of api which handle machine learning task easily. Tensorflow is easy can be used in mobile apps ans Website because of its great community and it is handled by google Below are the framework which are developed by tensorflow community

- Tensorflow (Main Python Framework Widely used)
- Tensorflow.js (Use in Javascript web and app and node application)

- Tensorflow Lite (For Mobile compatibility)

So, Tensorflow is a great option, as it provide great API which we will need in this project in Mobile App and in Web App.

In this chapter we finished all the data handling cleaning, trained the machine learning model and exported them to make inferences from them in real time arduino data. Now we have to include them in React and iOS and Andriod Application.

Chapter 3

Web App

3.1 WebApp

A web application is a computer program that utilizes web browsers and web technology to perform tasks over the Internet.

Web applications use a combination of server-side scripts (PHP and ASP) to handle the storage and retrieval of the information, and client-side scripts (JavaScript and HTML) to present information to users. This allows users to interact with the company using online forms, content management systems, shopping carts and more. In addition, the applications allow employees to create documents, share information, collaborate on projects, and work on common documents regardless of location or device.

3.1.1 How it works?

Web applications are usually coded in browser-supported language such as JavaScript and HTML as these languages rely on the browser to render the program executable. Some of the applications are dynamic, requiring server-side processing. Others are completely static with no processing required at the server.

The web application requires a web server to manage requests from the client, an application server to perform the tasks requested, and, sometimes, a database to store the information. Application server technology ranges from ASP.NET, ASP and ColdFusion, to PHP and JSP.
Here's what a typical web application flow looks like:

- User triggers a request to the web server over the Internet, either through a web browser or the application's user interface
- Web server forwards this request to the appropriate web application server
- Web application server performs the requested task – such as querying the database or processing the data – then generates the results of the requested data

- Web application server sends results to the web server with the requested information or processed data
- Web server responds back to the client with the requested information that then appears on the user's display

3.1.2 Let's Create our App

The technology we will use to build out web app is React app it's very fast and very scalable web framework managed and developed by Facebook. and once it got loaded then everything is stored and work easily.

Now Lets talk about Code We Will only talk about code about using machine learning model, how react works and more about html, css and JS is not in the scope of this document (optional*)

This is the main App.js Code which is starting point of our app

Listing 3.1: using trained model using TensorFlow.js

```
1  import * as tf from "@tensorflow/tfjs"
2
3  const model = async (temp, ph, setAns) => {
4      return await tf.loadLayersModel("https://raw.githubusercontent.com/s
5          if (parseFloat(temp) && parseFloat(ph)) {
6              if (temp === 0 || ph === 0) {
7                  alert("Please enter valid values")
8                  return
9              }
10             else {
11                 var dat = [parseFloat(temp), parseFloat(ph)]
12                 var shap = [1, 2]
13                 var results = m.predict(tf.tensor2d(dat, shap));
14                 Promise.resolve(results.dataSync()).then(s => {
15                     setAns(s)
16                 }).catch(e => {
17                     return
18                 })
19             }
20         }
21         else{
22             alert("Enter numeric value")
23             return
24         }
25     }).catch(e => {
26         alert(e)
27     })
28 }
```

Listing 3.2: accepting the user input and showing result

```
<div style={{ marginTop: 20 }} className="card-form">
    <form className="signup">
        <div className="form-title">Predictions of D.O(mg/L)</div>
        <div className="form-body">
            <div className="row">
                <input onChange={(e) => setTemp(e.target.value)} type="t
            </div>
            <div className="row">
                <input onChange={(e) => setPH(e.target.value)} type="tex
            </div>
        </div>
        <div className="rule"></div>
        <div className="form-footer" style={{ display: "flex", flexDirec
            <a href="/#" onClick={async () => { await model(temp, PH, se
            <div style={{ color: 'black' }}>{parseFloat(ans).toFixed(4)}
        </div>
    </form>
</div>
```

3.1.3 Let's Jump to Code

Let's understand the code behind very simple yet complex machine learning app. We are using binary of a trained machine learning model that we saved to our file in previous step.
some more data filtering been added by checking weather the input is correct format or not.

Every time user provide the input and press Enter, a async call to **model** function get called and then in turn the **model** uses the already trained ML model to make the prediction and return us the result.

Chapter 4

Testing and Evaluation

By Far our App and web app are working perfectly but we need to take care of some error and make app development and web much more efficient

Below are the Point which we need to be rectify

4.1 Limited data

These App uses water data which is very limited available and the data which is used in this whole project is downloaded from kaggle and that water data has some drawbacks

- The Temperature data is in between normal temperature for example (23-30°C) and if we make prediction below or above these temperature then we for sure will not get good predictions so we need more data of wide variety of temperature.
- And another point is of pH as we know all the water bodies which are pond lake all are not too much acidic and too much basic so when you make prediction for pH like 1-4 or basic solution like 10-14 then it will surely give error the trained model is not train for these temperature and pH ranges so we have to collect more data, to make good predictions

4.2 Data Scaling and Best Model

As the title suggest Data scaling and good model we already using best model for our mobile and web app but as we all know there are chances to minimize test error in machine learning we can improve model and make error minimum but we can't make error go away so we can always do work to reduce error and make our model accuracy very good

And Data scaling says while training data we can scale input to make predictions good this process is also the subset of above paragraph but these are two different task but interlinked so we have to take care of it

Chapter 5

Conclusion

So here we are finishing all our work make a very vast machine learning Model and deployed them in Web App. So here are the things I achieve so far in this project

1. I have completely met my aim and solved the problem
2. The solution is although best for the certain most common temperature and pH range but for wide variety of data it is getting failed. It will require more data which is either have to observed manually or find some good resources where it can be found.

5.1 Future Work

Gather more data over wide range of temperature and pH value and train model of it

Develop an small prototype

ask user about their permission if we can use their data to train our model(app feature)