EFFICIENT WATER QUALITY ANALYSIS & PREDICTION USING MACHINE LEARNING

A Project report submitted in partial fulfilment of 7th semester in degree of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING

Submitted by

Team ID: PNT2022TMID46078

AKALYA.A	814819104004
RAJALAKSHMI. K	814819104022
SANTHIYA.P	814819104024
SOWMIYA.A	814819104029
SRIVIDHYA.R	814819104031

UNIVERSITY COLLEGE OF ENGINEERING, ARIYALUR

ANNA UNIVERSITY: CHENNAI 600025

DEPARTMENT OF COMPUTER SCIENCE ANDENGINEERING

NOVEMBER/DECEMBER-2022

(REGULATIONS-2017)

UNIVERSITY COLLEGE OF ENGINEERING, ARIYALUR

(A Constituent College of Anna University, Chennai)

BONAFIDE CERTIFICATE

Certified that this project report **"EFFICIENT WATER QUALITY ANALYSIS & PREDICTION USING MACHINE LEARNING"** is the bonafide record work done by **AKALYA.A**(**814819104004**),**RAJALAKSHMI.K**(**814819104022**),**SANTHIYA.P** (**814819104024**),**SOWMIYA.A**(**814819104029**)and **SRIVIDHYA.R(814819104031)** for **"HX 8001 PROFESSIONAL READINESS FOR INNOVATION, EMPLOYABILITY AND ENTREPRENEURSHIP"** in **VII** semester of **B.E.,**degree course **in ComputerScience and Engineering** branch during the academic year of 2022 - 2023.

Staff-In charge
MR.S.SURESH

Evaluator
Dr.S.Jayanthi

Head of the Department
Dr.S.Jayanthi

EFFICIENT WATER QUALITY ANALYSIS & PREDICTION USING MACHINE LEARNING

ABSTARCT

The main objective of this project is to measure water quality using machine learning methods. An evaluation of a body of water's quality is done using a numerical term called potability. The overall water quality in terms of potability was evaluated in this study using the following water quality measures. The factors were ph, Hardness, Solids, Chloromines, Sulfate, Conductivity, Organic Carbon, Trihalomethanes, and Turbidity. These factors are utilised as a feature vector to represent the water quality. Both a genuine dataset containing data from various sites a synthetic dataset produced randomly using parameters were used in the experiments. The Linear regression algorithm classifier outperforms other classifiers, according on the outcomes of two different kinds of classifiers. The results show that potability may be predicted with accuracy by machine learning techniques. Index terms include classification, data mining, potability, and water quality parameters.

1. INTRODUCTION

1.1 PROJECT OVERVIEW

Access to safe drinking-water is essential to health, a basic human right and a component of effective policy for health protection. This is important as a health and development issue at a national, regional and local level. In some regions, it has been shown that investments in water supply and sanitation can yield a net economic benefit, since the reductions in adverse health effects and health care costs outweigh the costs of undertaking the interventions. Analyzing water quality is a complicated subject because of the many variables that affect it. The different ways that water is used are intricately tied to this idea. Different standards are required by various needs. Water quality prediction is a topic of extensive research. The physical and chemical characteristics of water are typically assessed in relation to the intended use of the water. The values for each variable must then be determined to be acceptable and unacceptable. Water is deemed suitable for a given purpose when it complies with the established requirements. Numerous physical and chemical characteristics can be used to evaluate the quality of water. As a result, it is not practical to adequately define water quality on a spatial or temporal basis by investigating the behaviour of each individual variable alone. Combining the values of several physical and chemical factors into one value is the trickier approach. Each variable's index had a quality value function (often linear) that represented the equivalency between the variable and its quality level. These calculations were made using physical parameter values obtained from water sample studies or direct measurements of a substance's concentration.

1.2 PURPOSE

The major goal of this project is to use machine learning techniques to measure water quality. Portability is a numerical phrase that is used to assess the quality of a body of water. The proposed methodology achieves reasonable accuracy using a minimal number of parameters to validate the possibility of its use in real time water quality detection systems. It will be using a dataset that contains data on all of the major factors that affect the portability of water. All of the factors that affect water quality are very important, so we need to briefly explore each feature of this dataset before training a machine learning model to predict whether a water sample is safe or unsuitable for consumption.

2. LITERATURE REVIEW

2.1 EXISTING PROBLEM

2.1.1 TITLE: Efficient Water Quality Prediction Using Supervised Machine Learning

AUTHOR: Umair Ahmed 1 , Rafia Mumtaz

About 70% of the earth's surface is covered by water, which is one of the most critical resources for maintaining life. Rapid industrialization and urbanisation have caused an alarming rate of water quality degradation, which has resulted in terrible diseases. Traditional methods for estimating water quality involve costly and time-consuming statistical and laboratory tests, making the idea of real-time monitoring irrelevant today. There must be a quicker, more practical solution because of the dire effects of bad water quality. In order to estimate the water quality index (WQI), a unique index to characterise the general quality of water, and the water quality class (WQC), a distinct class established on the basis of the WQI, this research investigates a number of supervised machine learning techniques. The suggested methodology uses temperature, turbidity, pH, and total dissolved solids as its four input parameters. With mean absolute errors (MAE) of 1.9642 and 2.7273, respectively, gradient boosting with a learning rate of 0.1 and polynomial regression with a degree of 2 predict the WQI the most accurately. A multi-layer perceptron (MLP), on the other hand, classifies the WQC most accurately (0.8507), with a configuration of (3, 7). The suggested methodology validates the viability of its usage in real-time water by achieving reasonable accuracy with a small number of parameters. systems for detecting quality.

2.1.2 TITLE: Hybrid decision tree-based machine learning models for short-term water quality prediction

AUTHOR: Hongfang Lu a, c, * , Xin Ma

The foundation of human existence and economic growth, as well as a factor in environmental sustainability, are water resources. The secret to bettering pollution control and water management is accurate water quality forecasting. Two new hybrid decision tree-based machine learning techniques are presented in this research. To produce findings for short-term water quality prediction that are more precise, models are offered. Extreme gradient boosting (XGBoost) and random forest (RF), the two hybrid models' fundamental models, introduce, respectively, complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN). Taking the water resources of Gales Creek site in Tualatin River (one of the most polluted rivers in the world) Basin as an example, a total of 1875 data (hourly data) from May 1, 2019 to July 20, 2019 are collected. Two hybrid models are used to predict six water quality indicators, including water temperature, dissolved oxygen, pH value, specific conductance, turbidity, and fluorescent dissolved organic matter. Six error metrics are introduced as the basis of performance evaluation, and the results of the two models are compared with the other four conventional models. In addition, this paper discusses the prediction model's stability. The investigation demonstrates that CEEMDAN-RF and CEEMDAN-XGBoost have stronger prediction stability than other benchmark models.

2.1.3 TITLE: Comparative analysis of surface water quality prediction performance and identification of key water parameters using different machine learning models based on big data

AUTHOR: Kangyang Chen a, 1 , Hexia Chen

The effectiveness of machine learning models for predicting water quality may depend on both the models themselves and the parameters in the data set that were selected for training the learning models. In order to further lower prediction costs and increase prediction efficiency, the learning models should additionally identify the essential water characteristics. Using big data (33,612 observations) from the major rivers and lakes in China from 2012 to 2018, we attempted to compare the water quality prediction performances of 10 learning models for the first time (7 traditional and 3 ensemble models), based on the precision, recall, F1-score, weighted F1-score, and explore the potential key water parameters for future model prediction. Our findings indicated that learning models could forecast water quality more accurately when given larger amounts of data. Decision tree (DT), random forest (RF), and deep cascade forest (DCF) models that were trained using data sets of pH, DO, CODMn, and NH3 eN performed much better than the other 7 models in terms of predicting all 6 levels of water quality that the Chinese government recommends. Furthermore, DT, RF, and DCF found and verified two important water parameter sets (DO, CODMn, and NH3eN; CODMn, and NH3eN) as having high specificities for perdition water quality. Therefore, for future water quality monitoring and fast water quality warning, DT, RF, and DCF with specified important water parameters might be prioritised.

2.1.4 TITLE: Machine learning methods for better water quality prediction

AUTHOR: Ali Najah Ahmeda , Faridah Binti Othman

The modelling water quality factors play a big role in any analysis of an aquatic system. The conventional modelling approaches typically entail time-consuming procedures and rely on datasets with significant amounts of uncertain or unspecified input data. Artificial intelligence (AI) implementation produces a flexible mathematical structure with the ability to recognise complicated and non-linear correlations between input and output data. The Johor River Basin has seen significant deterioration as a result of many human and development-related activities. Therefore, developing a water quality prediction model is crucial and will be an effective tool for improved management of water resources. Radial Basis Function Neural Networks (RBF-ANN), Multi-Layer Perceptron Neural Networks, and Adaptive Neuro-Fuzzy Inference System (ANFIS) are a few of the several modelling strategies that have been used (MLP-ANN). However, as a result of random and intentional errors, data collected from monitoring stations and experiments may be contaminated by noise signals. Making an accurate prediction is rather challenging because of the noise in the data. As a result, an enhanced wavelet de-noising technique based on the WDT-ANFIS Neuro-Fuzzy Inference System, which uses historical data for the water quality parameter, has been suggested. Due to objective and/or subjective errors, the data received from experimentation and examination may be tainted by noise signals. For instance, measuring, recording, reading, and environmental circumstances can result in experimental errors. Prior to using this data, the noise must be removed (also known as signal de-noising) because it may affect the model results.

2.1.5 TITLE: Novel PSO Optimized Voting Classifier Approach for Predicting water Quality

AUTHOR: Shweta Agrawal , 1 Sanjiv Kumar Jain

Diverse contaminants have been a threat to the water's quality during the past few years. As a result, predicting and modelling water quality are crucial for managing water contamination. In this research, a machine learning ensemble model for evaluating water quality is proposed. The findings of the suggested model are contrasted with those of k-nearest neighbour, Nave Bayes, support vector machine, and decision tree machine learning models. Seven statistically significant variables are present in the dataset under consideration: pH, conductivity, dissolved oxygen, biochemical oxygen demand, nitrate, total coliform, and faecal coliform. To evaluate the quality of the water, the water quality index is computed. A voting classifier with hard voting has been created to use an ensemble approach. The voting classifier offers the highest prediction accuracy of 99.5% of the water quality index, compared to the predictions of 99.2%, 90%, 79%, and 99% offered by k-nearest neighbors, Naive Bayes, support vector machines, and decision trees, respectively. Using particle swarm based optimization, it was further improved to 99.74%. We propose a homogenous ensemble learning technique that is both effective and efficient for enhancing classification performance. The suggested method splits the dataset into smaller groups at random using a mean-based splitting strategy, and then uses the classification and regression tree (CART) algorithm to model each division. It is compared to certain well-known machine learning techniques, including decision trees (DT), Naive Bayes, support vector machines (SVM), and k-nearest neighbour (KNN).

2.2 REFERENCES

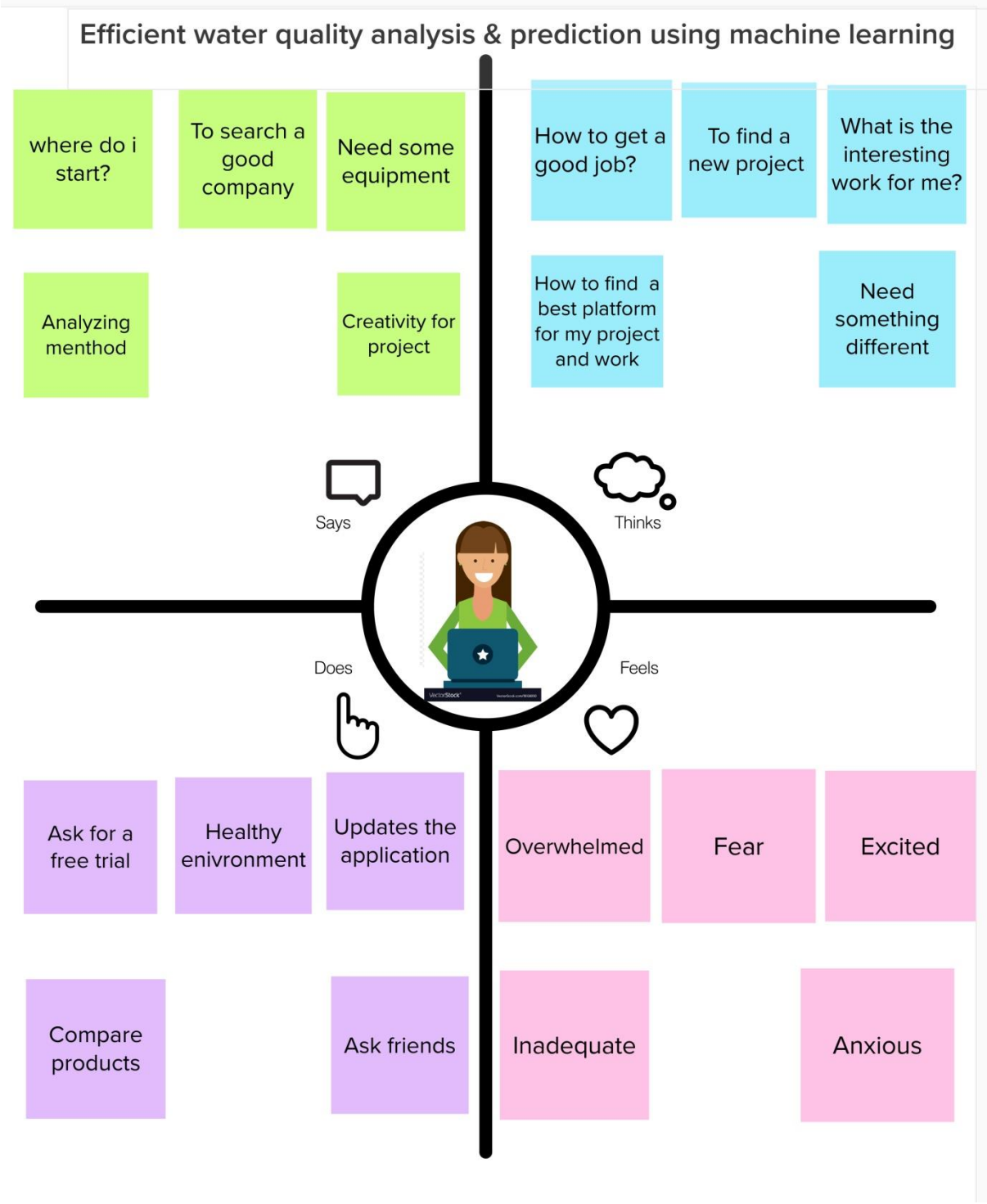
1. Umair Ahmed 1 , Rafia Mumtaz,Efficient Water Quality Prediction Using Supervised Machine Learning, 2019.
2. Hongfang Lu a, c, Xin Ma,Hybrid decision tree-based machine learning models for short-term water quality prediction, 2020.
3. Kangyang Chen a, 1 , Hexia Chen, Comparative analysis of surface water quality prediction performance and identification of key water parameters using different machine learning models based on big data, 2020.
4. Ali Najah Ahmeda , Faridah Binti Othman, Machine learning methods for better water quality prediction, 2019.
5. Shweta Agrawal , 1 Sanjiv Kumar Jain, Novel PSO Optimized Voting Classifier Approach for Predicting Water Quality, 2022.

2.3 PROBLEM STATEMENT DEFINITION

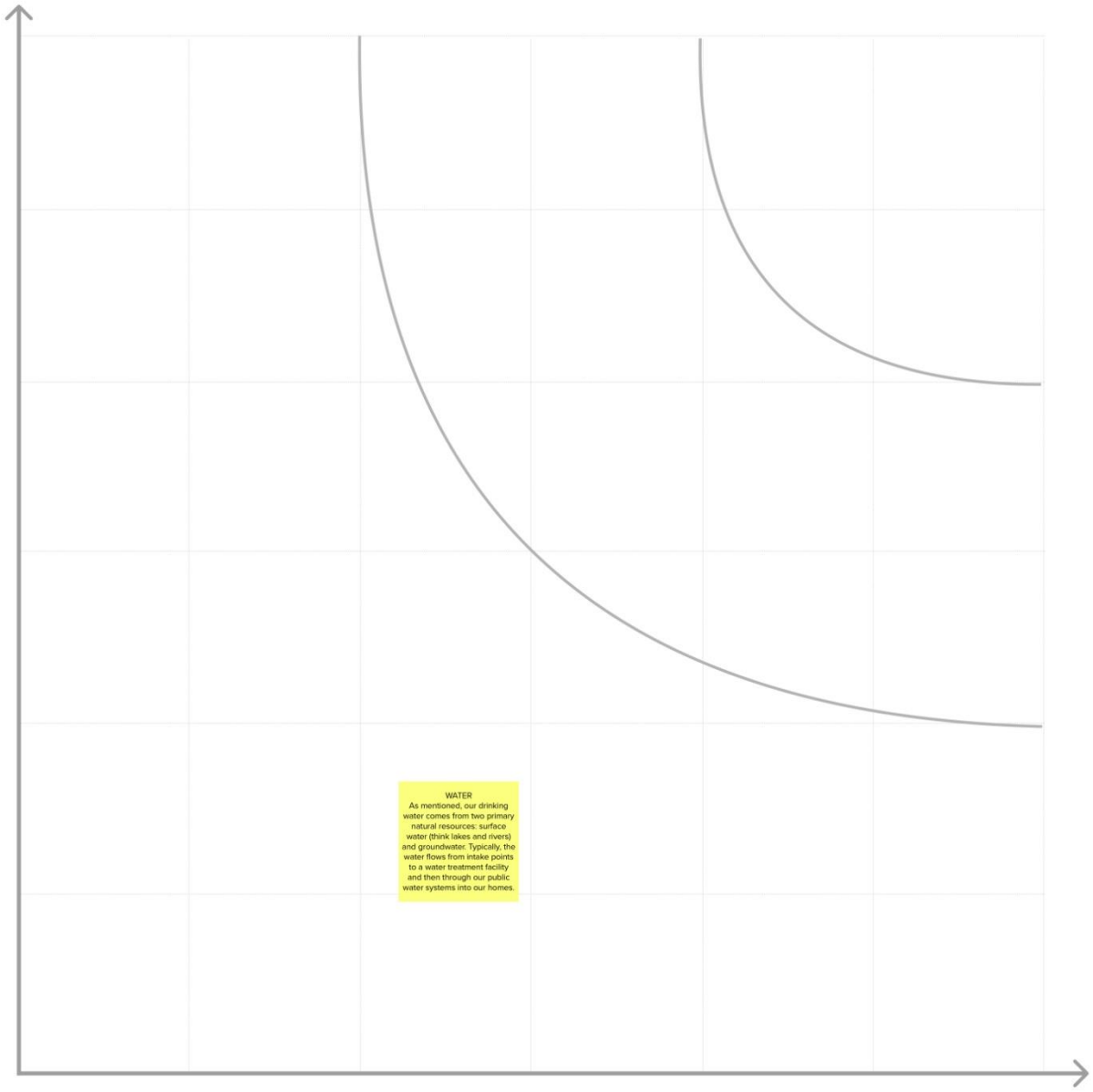
Since several decades ago, water contamination has been a significant issue in many developing nations. Another significant problem that might result from contaminated water is water scarcity. The quality of water may be impacted by a number of variables. Examples include Contamination, acidification, and thermal pollution. Additionally, dissolved oxygen levels are impacted by surface water temperatures, which are influenced by urbanisation and industrial effluents. The metabolism of the organisms residing in the reservoir depends on dissolved oxygen. The effluents that are introduced by industry or by people living nearby the water body might affect the pH of the water. Contamination, on the other hand, can destroy aquatic life and diminish the supply of clear drinking water. Water's turbidity is a characteristic that indicates how pure it is. Greater transparency in aquatic bodies indicates high water quality. In the monsoon, when the water is slightly murkier and also influenced by suspended particulates, reservoirs have significant turbidity. Water quality is greatly influenced by temperature. Both the aquatic metabolism and the pace of chemical reactions are impacted. Aquatic metabolism has a limited range of temperature tolerance, therefore even a small shift in the temperature of the surface water can be harmful to the organisms that live in the water basin. The components of nutrition are necessary for living. Poor water quality can have an impact on aquatic metabolism and nutrients. Using the information collected from monitoring the quality of the water, predictions about its quality can be produced in a few stages to enhance water management.

3. IDEATION & PROPOSED SOLUTION

3.1 EMPATHY MAP CANVAS



3.2 IDEATION & BRAINSTORMING



AKALYA

SANTHIYA

SOWMIYA

SRIVIDHYA

RAJAYALAKSHMI

3.3 PROPOSED SOLUTION

The suggested system's main goal is to assess portability. It is split into two sections: a testing phase and a training phase. The subsequent steps are completed in both sections. data on training Data from pH and hardness tests There are several words that can be used to describe something, including solids, chloramines, sulphate, conductivity, organic carbon, trihalomethanes, turbidity, and portability. This is how the data set was selected: The selection of the water quality data set, which is a requirement for model construction, is based on the collection of crucial parameters that affect water quality, the estimation of the number of data samples, and the definition of the class labels for each data sample that is present in the data. The data sets used in this investigation consist of ten indicator parameters. Hardness and pH value are two examples of these variables. A substance's characteristics can be described using the phrases solids, chloramines, sulphate, conductivity, organic carbon, trihalomethanes, turbidity, and portability, among others.

3.4 PROBLEM SOLUTION FIT

Water is one of the most important natural resources for all living organisms on earth. The monitoring of treated wastewater discharge quality is vitally important for the stability and protection of the ecosystem. Collecting and analyzing water samples in the laboratory consumes much time and resources.

4. REQUIREMENT ANALYSIS

4.1 FUNCTIONAL REQUIREMENT

Framework Creation

Users can register in this module and provide input for the water's chemical values, such as its pH and chemical composition, in order to estimate the quality of the water.

Data Collection

The process of gathering data and identifying its patterns and behaviours is known as data collection. This was accomplished by obtaining secondary data from a third party for various water parameter values. The information include elements such as temperature, pH, dissolved oxygen, turbidity, chloride, etc. The study was conducted by extracting variables from the dataset, such as pH, index value, dissolved oxygen, and turbidity. These factors directly affect living things, making them the impactful characteristics of water quality that can be used to quantify water quality accurately.

User Inputs

The user enters the pre-defined water dataset in this module, such as the water's temperature, pH, dissolved oxygen content, turbidity, chloride content, etc. When providing input, the water quality index should be high (80), with high denoting good quality.

Data Pre-processing

Pre-processing data enhances the effectiveness and quality of the data. The quality of the data is impacted by the inconsistent and noisy nature of raw data. Data preparation and transformation are part of the process known as data preprocessing.

Forecasting Water Quality

This process comprises creating, testing, and validating a predictive model. Techniques based on machine learning using the Linear Regression Algorithm were employed to create the prediction model. In this study, the model was created using denoising prediction techniques. Both of these techniques train the model via unsupervised learning.

4.2 NON FUNCTIONAL REQUIREMENTS

Usability

The system shall allow the users to access the system with pc using web application. The system uses a web application as an interface. The system is user friendly which makes the system easy

Availability

The system is available 100% for the user and is used 24 hrs a day and 365 days a year. The system shall be operational 24 hours a day and 7 days a week.

Scalability

Scalability is the measure of a system's ability to increase or decrease in performance and cost in response to changes in application and system processing demands.

Security

A security requirement is a statement of needed security functionality that ensures one of many different security properties of software is being satisfied.

Performance

The information is refreshed depending upon whether some updates have occurred or not in the application. The system shall respond to the member in not less than two seconds from the time of the request submittal. The system shall be allowed to take more time when doing large processing jobs. Responses to view information shall take no longer than 5 seconds to appear on the screen.

Reliability

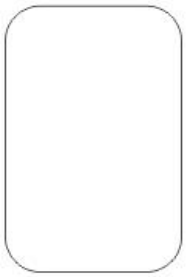
The system has to be 100% reliable due to the importance of data and the damages that can be caused by incorrect or incomplete data. The system will run 7 days a week. 24 hours a day.

5. PROJECT DESIGN

5.1 DATA FLOW DIAGRAMS

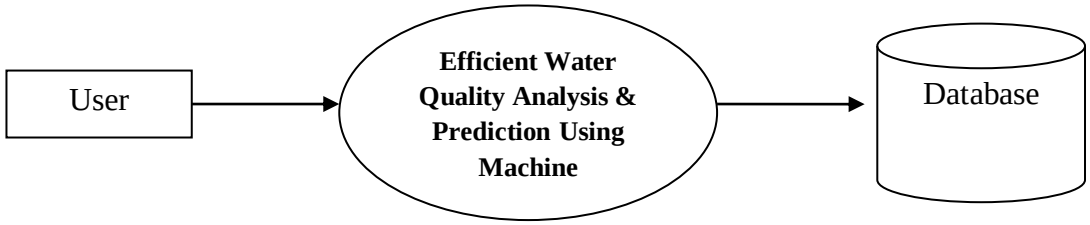
A two-dimensional diagram explains how data is processed and transferred in a system. The graphical depiction identifies each source of data and how it interacts with other data sources to reach a common output. Individuals seeking to draft a data flow diagram must identify external inputs and outputs, determine how the inputs and outputs relate to each other, and explain with graphics how these connections relate and what they result in. This type of diagram helps business development and design teams visualize how data is processed and identify or improve certain aspects.

Data flow Symbols:

Symbol	Description
	An **entity**. A source of data or a destination for data.
	A **process** or task that is performed by the system.
	A **data store**, a place where data is held between processes.
	A **data flow**.

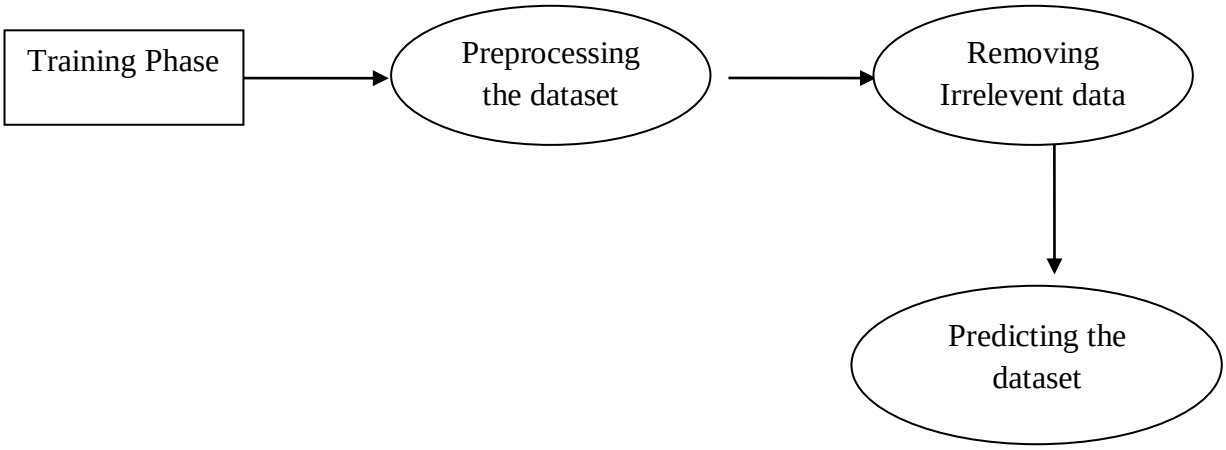
LEVEL 0

The Level 0 DFD shows how the system is divided into 'sub-systems' (processes), each of which deals with one or more of the data flows to or from an external agent, and which together provide all of the functionality of the system as a whole. It also identifies internal data stores that must be present in order for the system to do its job, and shows the flow of data between the various parts of the system.



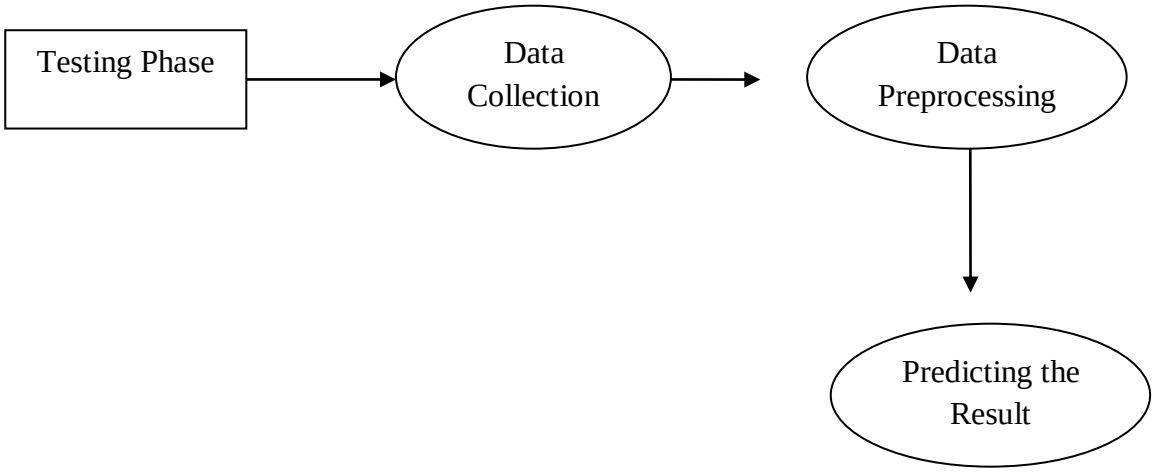
LEVEL 1

The next stage is to create the Level 1 Data Flow Diagram. This highlights the main functions carried out by the system. As a rule, to describe the system was using between two and seven functions - two being a simple system and seven being a complicated system. This enables us to keep the model manageable on screen or paper.

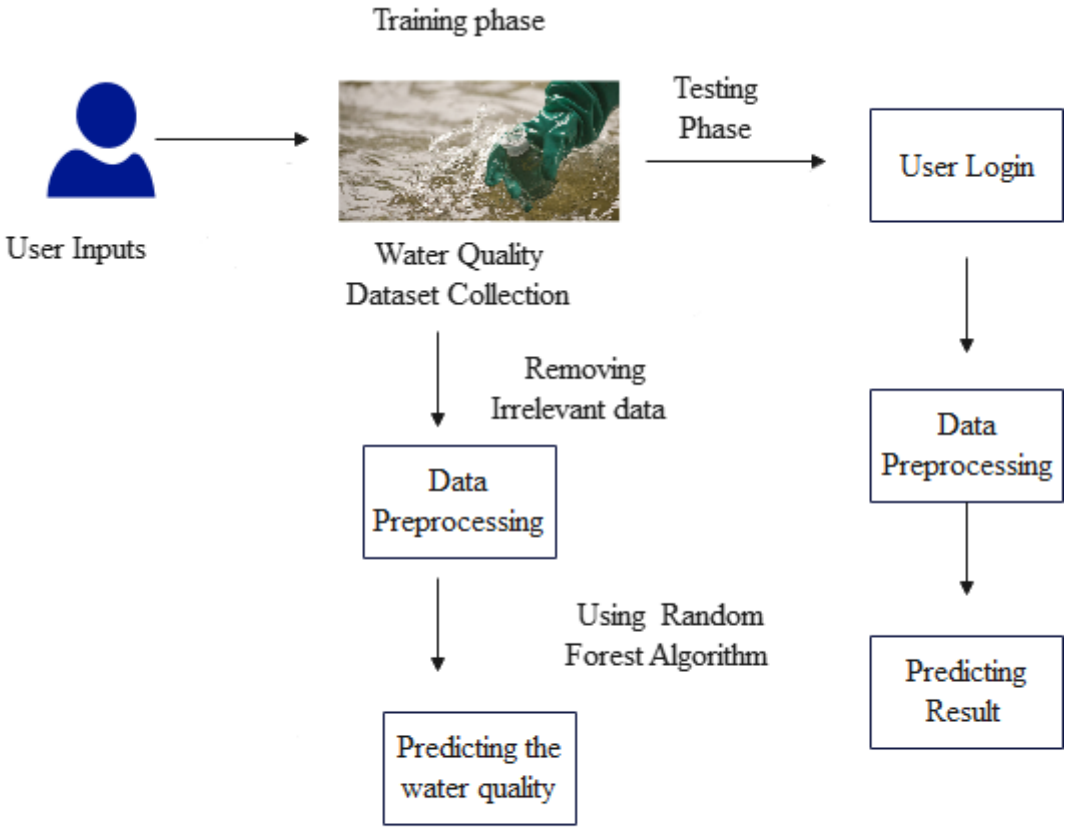


LEVEL 2

A Data Flow Diagram (DFD) tracks processes and their data paths within the business or system boundary under investigation. A DFD defines each domain boundary and illustrates the logical movement and transformation of data within the defined boundary. The diagram shows 'what' input data enters the domain, 'what' logical processes the domain applies to that data, and 'what' output data leaves the domain. Essentially, a DFD is a tool for process modeling and one of the oldest.



5.2 SOLUTION & TECHNICAL ARCHITECTURE



6. PROJECT PLANNING & SCHEDULING

6.1 SPRINT PLANNING & ESTIMATION

Project Planning Phase

Project Planning Template (Product Backlog, Sprint Planning, Stories, Story points)

Date	26 October 2022
Team ID	PNT2022TMID46078
Project Name	Efficient water quality analysis and prediction using machine learning

Product Backlog, Sprint Schedule, and Estimation

Use the below template to create product backlog and sprint schedule

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registration	USN-1	As a user ,I can register to the Application by using my Email and password.	10	High	Akalya,Santhiya, Sowmiya
Sprint-1	Confirmation mail	USN-2	Confirmation Mail will be received by the user, that he is registered for the Application.	10	High	Akalya,Santhiya, Sowmiya
Sprint-1	Login	USN-3	As a user I need to login to my registered mail to access the Application.	10	Low	Akalya,Santhiya, Sowmiya
Sprint-2	Search box	USN-4	User need search box to search the water quality details	20	Medium	Srividhya, Rajalakshmi
Sprint-3	News gathering	USN-5	User need to get result for his search.	20	High	Akalya,Srividhya, Rajalakshmi
Sprint-4	Chatbot(IBM Watson)	USN-6	If user have any doubt in using the Application , Chat bot is available in the Application to help the user.	20	High	Akalya,Santhiya, Sowmiya, Srividhya, Rajalakshmi
Sprint-4	IBM Db2	USN-7	Application is integrated with DB2	10	High	Akalya, Sowmiya

6.2 SPRINT DELIVERY SCHEDULE

Project Planning Phase
Milestone and Activity List

Date	27 October 2022
Team ID	PNT2022TMID46078
Project Name	Efficient Water Quality Analysis and PredictionUsing Machine Learning

TITLE	DESCRIPTION	DATE
Literature Survey & Information Gathering	Literature survey on the selected project & gathering information by referring the, technical papers, research publications etc.	2 Sept 2022
Prepare Empathy Map	Prepare Empathy Map Canvas to capture the user Pains & Gains, Prepare list of problem statements	10 Sept 2022
Ideation Brain Storming	List the by organizing the brainstorming session and prioritize the top 3 ideas based on the feasibility & importance.	15 Sept 2022
Proposed Solution	Prepare the proposed solution document, which includes the novelty, feasibility of idea, business model, social impact, scalability of solution, etc.	23 Sept 2022
Problem Solution Fit	Prepare problem solution fit document	28 Sept 2022
Solution Architecture	Prepare solution architecture document.	5 October 2022
Customer Journey	Prepare the customer journey maps to	7 October 2022

	understand the user interactions & experiences with the application (entry to exit).	
Solution Requirement	Prepare the functional requirement document	8 October 2022
Data Flow Diagrams	Draw the data flow diagrams and submit for review.	10 October 2022
Technology Architecture	Prepare the technology architecture diagram.	15 October 2022
Prepare Milestone & Activity List	Prepare the milestones & activity list of the project.	26 October 2022
Project Development - Delivery of Sprint-1, 2, 3 & 4	Develop & submit the developed code by testing it.	16 &17 November 2022

7. TESTING

7.1 TEST CASES

A test case has components that describe input, action and an expected response, in order to determine if a feature of an application is working correctly. A test case is a set of instructions on “HOW” to validate a particular test objective/target, which when followed will tell us if the expected behavior of the system is satisfied or not.
Characteristics of a good test case:

- Accurate: Exacts the purpose.
- Economical: No unnecessary steps or words.
- Traceable: Capable of being traced to requirements.
- Repeatable: Can be used to perform the test over and over.
- Reusable: Can be reused if necessary.

S.NO	**Scenario**	**Input**	**Excepted output**	**Actual output**
1	User Login Form	User name and password	Login	Login success.
2	Chemical components	Chemical components details	Added successfully	Chemical component stored in database.
3	Pre processing	Chemical values	Removing irrelevant data	Predicting water quality.

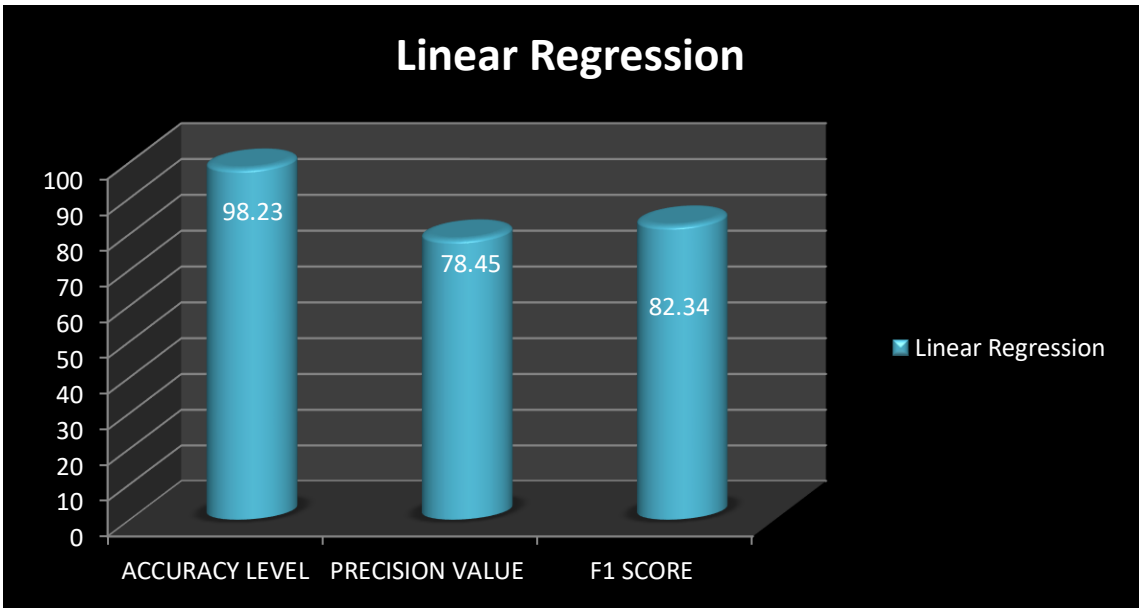
7.2 USER ACCEPTANCE TESTING

This is a type of testing done by users, customers, or other authorised entities to determine application/software needs and business processes. Acceptance testing is the most important phase of testing as this decides whether the client approves the application/software or not. It may involve functionality, usability, performance, and U.I of the application. It is also known as user acceptance testing (UAT), operational acceptance testing (OAT), and end-user testing.

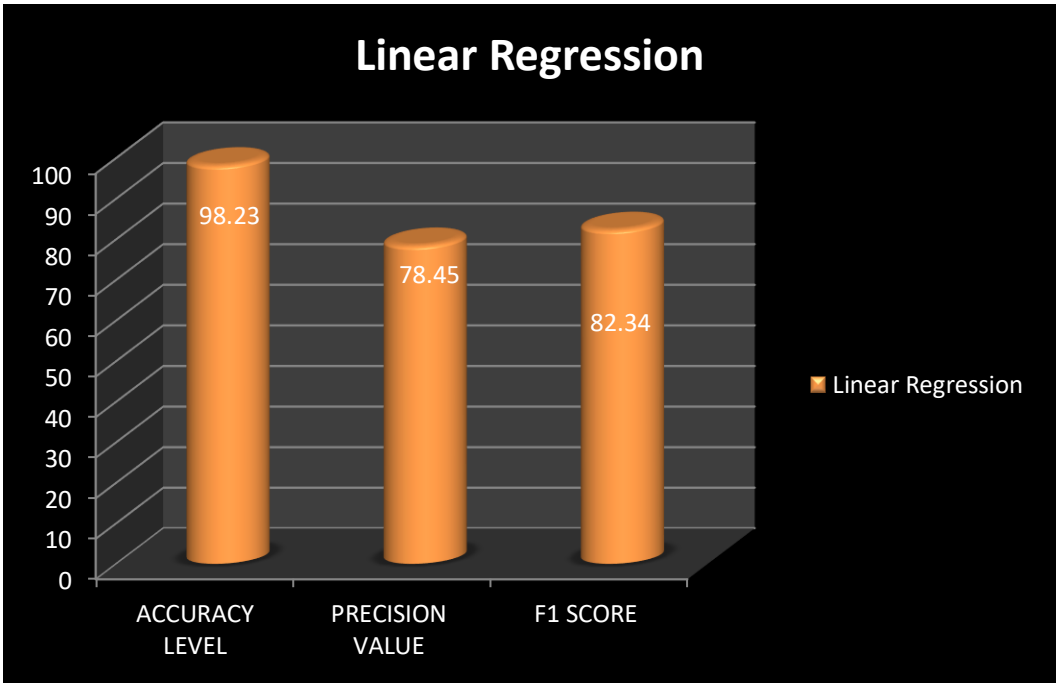
8. RESULTS

8.1 PERFORMANCE METRICS

Existing system



Proposed system



9. ADVANTAGES & DISADVANTAGES

ADVANTAGES

- Control is simple; testing simple; and it is not an expensive method.
- Proposed models can accurately predict and classify the water quality.
- In addition, it has been reported that machine learning methods showed high performance in predicting the WQ.
- To protect the environment and human health.

DISADVANTAGES

- Need additional sensors to predict the water quality
- Time complexity is high.
- Technically Challenging But Fails To Account For The Variation In Water Quality
- Automatic control systems are expensive

10. CONCLUSION

Water quality, one of the most crucial resources for life, is determined by potability. Traditionally, a costly and drawn-out lab analysis was needed to test the quality of water. This study investigated a different machine learning approach for forecasting water quality based on just a few basic water quality variables. A group of representative supervised machine learning methods were utilised to estimate. Before water was made available for consumption, it would find water of poor quality and alert the proper authorities. By decreasing the number of people who consume poor water, it should help prevent illnesses like typhoid and diarrhoea. The application of a prescriptive analysis based on expected values in this situation would lead to the development of future tools to support decision- and policy-makers.

11. FUTURE SCOPE

This project describes crop yield prediction ability of the algorithm. In future we can determine the efficient algorithm based on their accuracy metrics that will helps to choose an efficient algorithm for crop yield prediction

12. APPENDIX

SOURCE CODE:

//Model building

```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
import warnings
import pickle
data = pd.read_csv('./Data/water_dataX.csv', encoding ='Latin-1', low_memory=
False)
data.head()
```

```
data.describe()

data.info()

data['Temp'] = pd.to_numeric(data['Temp'], errors ='coerce')
data ['D.O. (mg/l)'] =pd.to_numeric(data['D.O. (mg/l)'],errors ='coerce')
data ['PH'] =pd.to_numeric(data['PH'],errors ='coerce')
data ['B.O.D. (mg/l)'] =pd.to_numeric(data['B.O.D. (mg/l)'],errors ='coerce')
data ['CONDUCTIVITY (µmhos/cm)'] =pd.to_numeric(data['CONDUCTIVITY
(µmhos/cm)'],errors ='coerce')
data ['NITRATENAN N+ NITRITENANN (mg/l)']
=pd.to_numeric(data['NITRATENAN N+ NITRITENANN (mg/l)'],errors
='coerce')
data ['TOTAL COLIFORM (MPN/100ml)Mean']
=pd.to_numeric(data['TOTAL COLIFORM (MPN/100ml)Mean'],errors
='coerce')
print(data.dtypes)

data.isnull().sum()

data['Temp'].fillna(data['Temp'].mean(),inplace =True)
data['D.O. (mg/l)'].fillna(data['D.O. (mg/l)'].mean(),inplace =True)
data['PH'].fillna(data['PH'].mean(),inplace =True)
data['CONDUCTIVITY (µmhos/cm)'].fillna(data['CONDUCTIVITY
(µmhos/cm)'].mean(),inplace =True)
data['B.O.D. (mg/l)'].fillna(data['B.O.D. (mg/l)'].mean(),inplace =True)
data['NITRATENAN N+ NITRITENANN (mg/l)'].fillna(data['NITRATENAN
N+ NITRITENANN (mg/l)'].mean(),inplace =True)
data['TOTAL COLIFORM (MPN/100ml)Mean'].fillna(data['TOTAL
COLIFORM (MPN/100ml)Mean'].mean(),inplace =True)

data.drop(["FECAL COLIFORM (MPN/100ml)"], axis =1, inplace= True)

data = data.rename (columns = {'D.O. (mg/l)':'do'})
```

```
data = data.rename (columns = {'CONDUCTIVITY (µmhos/cm)':'co'})
data = data.rename (columns = {'B.O.D. (mg/l)':'bod'})
data = data.rename (columns = {'NITRATENAN N+ NITRITENANN
(mg/l)':'na'})
data = data.rename (columns = {'TOTAL COLIFORM
(MPN/100ml)Mean':'tc'})
data = data.rename (columns = {'STATION CODE':'station'})
data = data.rename (columns = {'LOCATIONS':'location'})
data = data.rename (columns = {'STATE':'state'})
data = data.rename (columns = {'PH':'ph'})

data.drop(["station"], axis =1, inplace= True)



data.drop(["location"], axis =1, inplace= True)
data.drop(["state"], axis =1, inplace= True)

data['npH']= data.ph.apply(lambda x:(100 if ( 8.5 >= x >= 7)
                          else (80 if (8.6 >=x>= 8.5) or (6.9>=x>=6.8)
                              else (60 if (8.8>= x >=8.6) or (6.8 >= x>=6.7)
                                  else (40 if (9>=x>=8.8) or (6.7 >=x>=6.4)
                                      else 0)))))
data['ndo']= data.do.apply(lambda x:(100 if (x>=6)
                          else (80 if (6>=x>=5.1)
                              else (60 if (5>=x>=4.1)
                                  else(40 if (4>=x>=3)
                                      else 0)))))
data['nco']= data.co.apply(lambda x:(100 if (5>=x>=0)
                          else (80 if (50>=x>=5)
                              else (60 if (500>=x>=50)
                                  else(40 if (10000>=x>=500)
                                      else 0)))))
data['nbdo']= data.bod.apply(lambda x:(100 if (3>=x>=0)
                          else (80 if (6>=x>=3)
```

```
                                else (60 if (80>=x>=6)
                                    else(40 if (125>=x>=80)
                                        else 0)))))
data['nec']= data.co.apply(lambda x:(100 if (75>=x>=0)
                                else (80 if (150>=x>=75)
                                    else (60 if (225>=x>=150)
                                        else(40 if (300>=x>=225)
                                            else 0)))))


data['nna']= data.na.apply(lambda x:(100 if (20>=x>=0)
                                else (80 if (50>=x>=20)
                                    else (60 if (100>=x>=50)
                                        else(40 if (200>=x>=100)
                                            else 0)))))

data['wph']= data.npH * 0.165
data['wdo']= data.ndo * 0.281
data['wbdo']= data.nbdo * 0.234
data['wec']= data.nec * 0.009
data['wna']= data.nna * 0.028
data['wco']= data.nco * 0.281
data['wqi']= data.wph+data.wdo+data.wbdo+data.wec+data.wna+data.wco
data

average= data.groupby('year')['wqi'].mean()

average.head()

x= data.iloc[:,0:7].values
y = data.iloc[:,7:].values

print(x.shape)
print(y.shape)
```

```
//Train the model
from sklearn import linear_model
from sklearn.model_selection import train_test_split
reg= linear_model.LinearRegression()
X_train, X_test, y_train, y_test = train_test_split (x,y, test_size = 0.2,
random_state =10)
reg.fit(X_train, y_train)

from sklearn import metrics
y_pred = reg.predict(X_test)
print ('MAE :', metrics.mean_absolute_error(y_test, y_pred))
print ('MSE :', metrics.mean_squared_error(y_test, y_pred))
print ('RMSE :', np.sqrt(metrics.mean_squared_error(y_test, y_pred)))

metrics.r2_score(y_test, y_pred)

import pickle


//save the model
pickle.dump(reg,open('reg_rf.pkl','wb'))
print("Training process is complete Model File Saved!")
```

//application building

```
from flask import Flask, render_template, flash, request,session
from cloudant.client import  Cloudant
import pickle


client = Cloudant.iam("0e9cd905-134b-416c-bcf5-340b8fb90718-
bluemix","GR4agWMCjVG8qJzP5CvfAmkRajhP9iDr3FfwgFLR-
8pA",connect=True)
```

```
my_database = client.create_database("database-dharan")


app = Flask(__name__)
app.config.from_object(__name__)
app.config['SECRET_KEY'] = '7d441f27d441f27567d441f2b6176a'



@app.route("/")
def homepage():

    return render_template('index.html')



@app.route("/userhome")
def userhome():

    return render_template('userhome.html')
@app.route("/addamount")

@app.route("/NewUser")
def NewUser():

    return render_template('NewUser.html')

@app.route("/user")
def user():

    return render_template('user.html')


@app.route("/newuse",methods=['GET','POST'])
```

```
def newuse():
    if request.method == 'POST':#

        x = [x for x in request.form.values()]
        print(x)
        data = {
            '_id': x[1],
            'name': x[0],
            'psw': x[2]
        }
        print(data)
        query = {'_id': {'Seq': data['_id']}}
        docs = my_database.get_query_result(query)
        print(docs)
        print(len(docs.all()))
        if (len(docs.all()) == 0):
            url = my_database.create_document(data)
            return render_template('goback.html', data="Register, please login using
your details")
        else:
            return render_template('goback.html', data="You are already a member,
please login using your details")

@app.route("/userlog", methods=['GET', 'POST'])
def userlog():
        if request.method == 'POST':

            user = request.form['_id']
            passw = request.form['psw']
            print(user, passw)

            query = {'_id': {'$eq': user}}
            docs = my_database.get_query_result(query)
            print(docs)
```

```
        print(len(docs.all()))
        if (len(docs.all()) == 0):
            return render_template('goback.html', pred="The username is not found.")
        else:
            if ((user == docs[0][0]['_id'] and passw == docs[0][0]['psw'])):

                return render_template("userhome.html")
            else:
                return render_template('goback.html',data="user name and password incorrect")


@app.route("/predict", methods=['GET', 'POST'])
def predict():
    if request.method == 'POST':
        outttt =""

        year = request.form["year"]
        do = request.form["do"]
        ph = request.form["ph"]
        co = request.form["co"]
        bod = request.form["bod"]
        na = request.form["na"]
        tc = request.form["tc"]

        model = pickle.load(open('reg_rf.pkl','rb'))



         total = [[int(year), float(do), float(ph), float(co), float(bod), float(na), float(tc)]]
        #total = int(year)+ float(do)+ float(ph)+ float(co)+float(bod)+float(na)+ float(tc)
```

```
//Test the model
        y_pred = model.predict(total)
        print(y_pred)

        y_pred1 = y_pred[[0][0]]

        y_pred2 = y_pred1[[10][0]]


        print(y_pred2)

        if (y_pred2 >= 95 and y_pred2 <= 100):

           outttt ="Excellent, the Predicted value is " + str(y_pred2)


        elif (y_pred2 >= 89 and y_pred2 <= 94):
           outttt = "Very good, the Predicted value is " + str(y_pred2)

        elif (y_pred2 >= 80 and y_pred2 <= 88):

           outttt="Good, the Predicted value is " + str(y_pred2)

        elif(y_pred2 >= 65 and y_pred2 <= 79):
           outttt = "Fair, the Predicted value is " + str(y_pred2)

        elif (y_pred2 >= 45 and y_pred2 <= 64):
           outttt ="Marginal, the Predicted value is " + str(y_pred2)
        else:
           outttt="Poor, the Predicted value is " + str(y_pred2)

        return render_template('userhome.html', prediction=outttt)
```

```
        if __name__ == '__main__':
    app.run(debug=True, use_reloader=True)
```

//goback.html

```
        <!DOCTYPE html>
        <html>
        <body>


        <p>{{data}}</p>

        <button onclick="goBack()">Go Back</button>

        <script>
        function goBack() {
          window.history.back();
        }
        </script>

        </body>
</html>
```

//index

```
        <html xmlns="">
        <head>
        <meta http-equiv="content-type" content="text/html; charset=utf-8" />
        <title>Water Quality Prediction</title>
        <meta name="keywords" content="" />
        <meta name="description" content="" />
        <link href="default.css" rel="stylesheet" type="text/css" />
        <style>

        /*
        Design by Free CSS Templates
```

```
http://www.freecsstemplates.org
Released for free under a Creative Commons Attribution 2.5 License
*/

body {
	margin: 0;
	padding: 0;
	background: #FFFFFF url(static/images/img01.gif) repeat-x;
	font-family: Georgia, "Times New Roman", Times, serif;
	font-size: 13px;
	color: #666666;
}

h1, h2, h3 {
	margin: 0;
	font-weight: normal;
	color: #3F586B;
}

h1 {
	font-size: 197%;
}

h2 {
	font-size: 167%;
}

h3 {
	font-size: 100%;
	font-weight: bold;
}

p, ol, ul {
	line-height: 170%;
```

```
}

p {
}

ol {
	margin-left: 0;
	padding-left: 0;
	list-style-position: inside;
}

ul {
	margin-left: 0;
	padding-left: 0;
	list-style: none;
}

ul li {
	padding-left: 15px;
	background: url(static/images/img07.gif) no-repeat 0px 7px;
}

blockquote {
	margin: 0;
	padding-left: 20px;
	font-style: italic;
}

blockquote * {
}

a {
	color: #FF5723;
}
```

```
a:hover {
	text-decoration: none;
	color: #1777B1;
}

img {
	border: none;
}

img.left {
	float: left;
	margin: 3px 15px 0 0;
}

img.right {
	float: right;
	margin: 3px 0 0 15px;
}

hr {
	display: none;
}

/* Header */

#header {
	width: 700px;
	height: 235px;
	margin: 0 auto;
	background: #A4C0C8 url(static/images/img02.jpg) no-repeat;
}

/* Logo */
```

```
#logo {
	height: 190px;
}

#logo h1, #logo h2 {
	text-align: center;
}

#logo h1 {
	padding-top: 40px;
	font-size: 350%;
}

#logo h2 {
	font-size: 150%;
}

#logo a {
	text-decoration: none;
	color: #3F586B;
}

/* Menu */

#menu {
	padding-top: 0;
	width: 798px;
}

#menu ul {
	margin: 0;
	padding: 10px 0 0 0;
	list-style: none;
```

```
        line-height: normal;
        text-align: center;
}

#menu li {
        display: inline;
        margin: 0;
        padding: 0;
}

#menu a {
        padding: 0 20px;
        text-decoration: none;
        font-size: 136%;
        font-weight: bold;
        color: #610720;
}

#menu a:hover {
        text-decoration: underline;
}

#menu .active a {
        color: #FFFFFF;
}

/* Page */

#page {
        width: 730px;
        margin: 0 auto;
        padding: 30px 0;
}
```

```
/* Content */

#content {
	float: left;
	width: 800px;
	padding-top: 8px;
}

.twocols {
}

.twocols .title {
	padding-bottom: 10px;
	border-bottom: 1px solid #97C984;
}

.twocols .col1, .twocols .col2 {
	width: 190px;
}

.twocols .col1 {
	float: left;
}

.twocols .col2 {
	float: right;
}

.twocols ul {
}

.twocols ul li {
	padding-left: 0;
}
```

```
/* Sidebar */

#sidebar {
	float: right;
	width: 260px;
}

.boxed {
	margin: 0 0 20px 0;
}

.boxed .title {
	width: 250px;
	height: 35px;
	margin: 0;
	padding: 10px 0 0 10px;
	background: #A8C3CB url(static/images/img03.jpg) no-repeat;
	font-size: 136%;
	color: #144B6B;
}

.boxed .content {
	padding: 20px;
	border: 1px solid #97C984;
	border-top: none;
}

.boxed h3 {
	margin: 0;
}

.boxed p, .boxed ul, .boxed ol {
	margin: 0;
```

```
	padding: 0;
	list-style: none;
	line-height: normal;
}

.boxed ul {
}

.boxed ul li {
	padding: 8px 0 8px 10px;
	background: url(static/images/img04.gif) no-repeat 0px 13px;
}

.boxed ul li.first {
	border: none;
}

/* Search */

#search {
}

#search form {
	margin: 0;
	padding: 0;
}

#search fieldset {
	margin: 0;
	padding: 0;
	border: none;
}

#search p {
```

```
	float: left;
	padding-top: 5px;
	font-size: 85%;
}

#searchinput {
	width: 210px;
	margin-bottom: 5px;
}

#searchsubmit {
	float: right;
}

/* Footer */

#footer {
	height: 100px;
	padding: 20px;
	background: #5F919E;
	border-top: 5px solid #4C747E;
}

#footer p {
	margin: 0;
	text-align: center;
	line-height: normal;
	font-size: 85%;
	color: #FFFFFF;
}

#footer a {
	color: #FFFFFF;
}
```

```
</style>
</head>
<body>
<div id="header">
	<div id="logo">
		<h1><a href="#">Water Quality 	</a></h1>
		<h2><a href="">Prediction</a></h2>
  </div>
	<div id="menu">
		<ul>
			<li class="active"><a href="/">Home</a></li>

	<li><a href="/user">UserLogin</a></li>
			<li><a href="/NewUser">NewUser</a></li>

		</ul>
	</div>
</div>
<div id="page">
	<div id="content">
		<div style="margin-bottom: 20px;">
			<h1 class="title">VirtualEye - Life Guard For Swimming Pools To Detect Active Drowning </h1>
			<p align="justify"><strong>Swimming is one of the best exercises that helps people to reduce stress in this urban lifestyle.
			Swimming pools are found larger in number in hotels, and weekend tourist spots and barely people have them in
			their house backyard. Beginners, especially, often feel it difficult to breathe underwater which causes breathing
			trouble which in turn causes a drowning accident. Worldwide, drowning produces a higher rate of mortality
			without causing injury to children. Children under six of their age are found to be suffering the highest
```

drowning mortality rates worldwide. Such kinds of deaths account for the third cause of unplanned death globally, with about 1.2 million cases yearly. To overcome this conflict, a meticulous system is to be implemented along the swimming pools to save human life.

```
</p>
                        <h2> </h2>
                    <p> </p>
                        </blockquote>
            </div>
                    <div> </div>
            <div class="twocols"></div>
          </div>
          <!-- end content -->
          <!-- end sidebar -->
          <div style="clear: both;"> </div>
      </div>
      <!-- end page -->
      <div id="footer">
          <p id="legal"> <a href="#" title="This page validates as CSS"><abbr
      title="Cascading Style Sheets"></abbr></a></p>
          </div>
      <div align=center> <a href='#'></a></div>
      </body>
</html>
//NewUser

      <html xmlns="">
      <head>
      <meta http-equiv="content-type" content="text/html; charset=utf-8" />
      <title>Secured data communication</title>
      <meta name="keywords" content="" />
      <meta name="description" content="" />
```

```
<link href="default.css" rel="stylesheet" type="text/css" />
<style>

/*
Design by Free CSS Templates
http://www.freecsstemplates.org
Released for free under a Creative Commons Attribution 2.5 License
*/

body {
	margin: 0;
	padding: 0;
	background: #FFFFFF url(static/images/img01.gif) repeat-x;
	font-family: Georgia, "Times New Roman", Times, serif;
	font-size: 13px;
	color: #666666;
}

h1, h2, h3 {
	margin: 0;
	font-weight: normal;
	color: #3F586B;
}

h1 {
	font-size: 197%;
}

h2 {
	font-size: 167%;
}

h3 {
	font-size: 100%;
```

```
	font-weight: bold;
}

p, ol, ul {
	line-height: 170%;
}

p {
}

ol {
	margin-left: 0;
	padding-left: 0;
	list-style-position: inside;
}

ul {
	margin-left: 0;
	padding-left: 0;
	list-style: none;
}

ul li {
	padding-left: 15px;
	background: url(static/images/img07.gif) no-repeat 0px 7px;
}

blockquote {
	margin: 0;
	padding-left: 20px;
	font-style: italic;
}

blockquote * {
```

```
}

a {
	color: #FF5723;
}

a:hover {
	text-decoration: none;
	color: #1777B1;
}

img {
	border: none;
}

img.left {
	float: left;
	margin: 3px 15px 0 0;
}

img.right {
	float: right;
	margin: 3px 0 0 15px;
}

hr {
	display: none;
}

/* Header */

#header {
	width: 700px;
	height: 235px;
```

```
	margin: 0 auto;
	background: #A4C0C8 url(static/images/img02.jpg) no-repeat;
}

/* Logo */

#logo {
	height: 190px;
}

#logo h1, #logo h2 {
	text-align: center;
}

#logo h1 {
	padding-top: 40px;
	font-size: 350%;
}

#logo h2 {
	font-size: 150%;
}

#logo a {
	text-decoration: none;
	color: #3F586B;
}

/* Menu */

#menu {
	padding-top: 0;
	width: 798px;
}
```

```
#menu ul {
	margin: 0;
	padding: 10px 0 0 0;
	list-style: none;
	line-height: normal;
	text-align: center;
}

#menu li {
	display: inline;
	margin: 0;
	padding: 0;
}

#menu a {
	padding: 0 20px;
	text-decoration: none;
	font-size: 136%;
	font-weight: bold;
	color: #610720;
}

#menu a:hover {
	text-decoration: underline;
}

#menu .active a {
	color: #FFFFFF;
}

/* Page */

#page {
```

```
	width: 730px;
	margin: 0 auto;
	padding: 30px 0;
}

/* Content */

#content {
	float: left;
	width: 800px;
	padding-top: 8px;
}

.twocols {
}

.twocols .title {
	padding-bottom: 10px;
	border-bottom: 1px solid #97C984;
}

.twocols .col1, .twocols .col2 {
	width: 190px;
}

.twocols .col1 {
	float: left;
}

.twocols .col2 {
	float: right;
}

.twocols ul {
```

```
}

.twocols ul li {
	padding-left: 0;
}

/* Sidebar */

#sidebar {
	float: right;
	width: 260px;
}

.boxed {
	margin: 0 0 20px 0;
}

.boxed .title {
	width: 250px;
	height: 35px;
	margin: 0;
	padding: 10px 0 0 10px;
	background: #A8C3CB url(static/images/img03.jpg) no-repeat;
	font-size: 136%;
	color: #144B6B;
}

.boxed .content {
	padding: 20px;
	border: 1px solid #97C984;
	border-top: none;
}

.boxed h3 {
```

```
	margin: 0;
}

.boxed p, .boxed ul, .boxed ol {
	margin: 0;
	padding: 0;
	list-style: none;
	line-height: normal;
}

.boxed ul {
}

.boxed ul li {
	padding: 8px 0 8px 10px;
	background: url(static/images/img04.gif) no-repeat 0px 13px;
}

.boxed ul li.first {
	border: none;
}

/* Search */

#search {
}

#search form {
	margin: 0;
	padding: 0;
}

#search fieldset {
	margin: 0;
```

```
	padding: 0;
	border: none;
}

#search p {
	float: left;
	padding-top: 5px;
	font-size: 85%;
}

#searchinput {
	width: 210px;
	margin-bottom: 5px;
}

#searchsubmit {
	float: right;
}

/* Footer */

#footer {
	height: 100px;
	padding: 20px;
	background: #5F919E;
	border-top: 5px solid #4C747E;
}

#footer p {
	margin: 0;
	text-align: center;
	line-height: normal;
	font-size: 85%;
	color: #FFFFFF;
```

```
}

#footer a {
	color: #FFFFFF;
}
.style5 {color: #1A3B5C}
</style>
</head>
<body>
<div id="header">
	<div id="logo">
    <h1><a href="#">Water Quality    </a></h1>
		<h2><a href="">Prediction</a></h2>
  </div>
	<div id="menu">
		<ul>
			<li><a href="/">Home</a></li>
	<li><a href="/user">UserLogin</a></li>
			<li class="active"><a href="/NewUser">NewUser</a></li>



		</ul>
	</div>
</div>
<div id="page">
	<div id="content">
		<div style="margin-bottom: 20px;">

			<p align="justify"><strong>
			  <form name="form1" method="post" action="/newuse">
           <table width="389" height="214" border="0" align="center">
             <tr>
               <td colspan="2"><div align="center" class="style5">
                 <h2 class="style5">New User </h2>
```

```
  </div></td>
 </tr>


 <tr>
  <td height="37"><h2 class="style5">Name </h2></td>
  <td><label>
   <input name="name" type="text" id="name" />
  </label></td>
 </tr>
  <tr>
  <td height="37"><h2 class="style5">Email id </h2></td>
  <td><label>
   <input name="_id" type="text" id="_id" />
  </label></td>

 </tr>
 <tr>
  <td><h2 class="style5">Password</h2></td>
  <td><label>
   <input name="psw" type="password" id="psw" />
  </label></td>
 </tr>
 <tr>
  <td> </td>
  <td>
   <input type="submit" name="Submit" value="Submit" /> <input
type="reset" name="reset" value="Reset" />
  </td>
 </tr>
</table>
        </form>
            </p>
            <p> </p>
```

```
                    </blockquote>
                </div>
                    <div> </div>
                <div class="twocols"></div>
            </div>
            <!-- end content -->
            <!-- end sidebar -->
            <div style="clear: both;"> </div>
        </div>
        <!-- end page -->
        <div id="footer">
            <p id="legal"> <a href="#" title="This page validates as CSS"><abbr
        title="Cascading Style Sheets"></abbr></a></p>
        </div>
        <div align=center> <a href='#'></a></div>
        </body>
</html>
//user

        <html xmlns="">
        <head>
        <meta http-equiv="content-type" content="text/html; charset=utf-8" />
        <title>Secured data communication</title>
        <meta name="keywords" content="" />
        <meta name="description" content="" />
        <link href="default.css" rel="stylesheet" type="text/css" />
        <style>

        /*
        Design by Free CSS Templates
        http://www.freecsstemplates.org
        Released for free under a Creative Commons Attribution 2.5 License
        */
```

```
body {
	margin: 0;
	padding: 0;
	background: #FFFFFF url(static/images/img01.gif) repeat-x;
	font-family: Georgia, "Times New Roman", Times, serif;
	font-size: 13px;
	color: #666666;
}

h1, h2, h3 {
	margin: 0;
	font-weight: normal;
	color: #3F586B;
}

h1 {
	font-size: 197%;
}

h2 {
	font-size: 167%;
}

h3 {
	font-size: 100%;
	font-weight: bold;
}

p, ol, ul {
	line-height: 170%;
}

p {
}
```

```
ol {
	margin-left: 0;
	padding-left: 0;
	list-style-position: inside;
}

ul {
	margin-left: 0;
	padding-left: 0;
	list-style: none;
}

ul li {
	padding-left: 15px;
	background: url(static/images/img07.gif) no-repeat 0px 7px;
}

blockquote {
	margin: 0;
	padding-left: 20px;
	font-style: italic;
}

blockquote * {
}

a {
	color: #FF5723;
}

a:hover {
	text-decoration: none;
	color: #1777B1;
```

```
}

img {
	border: none;
}

img.left {
	float: left;
	margin: 3px 15px 0 0;
}

img.right {
	float: right;
	margin: 3px 0 0 15px;
}

hr {
	display: none;
}

/* Header */

#header {
	width: 700px;
	height: 235px;
	margin: 0 auto;
	background: #A4C0C8 url(static/images/img02.jpg) no-repeat;
}

/* Logo */

#logo {
	height: 190px;
}
```

```
#logo h1, #logo h2 {
	text-align: center;
}

#logo h1 {
	padding-top: 40px;
	font-size: 350%;
}

#logo h2 {
	font-size: 150%;
}

#logo a {
	text-decoration: none;
	color: #3F586B;
}

/* Menu */

#menu {
	padding-top: 0;
	width: 798px;
}

#menu ul {
	margin: 0;
	padding: 10px 0 0 0;
	list-style: none;
	line-height: normal;
	text-align: center;
}
```

```
#menu li {
	display: inline;
	margin: 0;
	padding: 0;
}

#menu a {
	padding: 0 20px;
	text-decoration: none;
	font-size: 136%;
	font-weight: bold;
	color: #610720;
}

#menu a:hover {
	text-decoration: underline;
}

#menu .active a {
	color: #FFFFFF;
}

/* Page */

#page {
	width: 730px;
	margin: 0 auto;
	padding: 30px 0;
}

/* Content */

#content {
	float: left;
```

```
	width: 800px;
	padding-top: 8px;
}

.twocols {
}

.twocols .title {
	padding-bottom: 10px;
	border-bottom: 1px solid #97C984;
}

.twocols .col1, .twocols .col2 {
	width: 190px;
}

.twocols .col1 {
	float: left;
}

.twocols .col2 {
	float: right;
}

.twocols ul {
}

.twocols ul li {
	padding-left: 0;
}

/* Sidebar */

#sidebar {
```

```
        float: right;
        width: 260px;
}

.boxed {
        margin: 0 0 20px 0;
}

.boxed .title {
        width: 250px;
        height: 35px;
        margin: 0;
        padding: 10px 0 0 10px;
        background: #A8C3CB url(static/images/img03.jpg) no-repeat;
        font-size: 136%;
        color: #144B6B;
}

.boxed .content {
        padding: 20px;
        border: 1px solid #97C984;
        border-top: none;
}

.boxed h3 {
        margin: 0;
}

.boxed p, .boxed ul, .boxed ol {
        margin: 0;
        padding: 0;
        list-style: none;
        line-height: normal;
}
```

```
.boxed ul {
}

.boxed ul li {
	padding: 8px 0 8px 10px;
	background: url(static/images/img04.gif) no-repeat 0px 13px;
}

.boxed ul li.first {
	border: none;
}

/* Search */

#search {
}

#search form {
	margin: 0;
	padding: 0;
}

#search fieldset {
	margin: 0;
	padding: 0;
	border: none;
}

#search p {
	float: left;
	padding-top: 5px;
	font-size: 85%;
}
```

```
#searchinput {
	width: 210px;
	margin-bottom: 5px;
}

#searchsubmit {
	float: right;
}

/* Footer */

#footer {
	height: 100px;
	padding: 20px;
	background: #5F919E;
	border-top: 5px solid #4C747E;
}

#footer p {
	margin: 0;
	text-align: center;
	line-height: normal;
	font-size: 85%;
	color: #FFFFFF;
}

#footer a {
	color: #FFFFFF;
}
.style5 {color: #1A3B5C}
</style>
</head>
<body>
```

```
<div id="header">
	<div id="logo">
			<h1><a href="#">Water Quality    </a></h1>
		<h2><a href="">Prediction</a></h2>
  </div>
	<div id="menu">
		<ul>
			<li><a href="/">Home</a></li>
			 <li class="active"><a href="/user">UserLogin</a></li>
				<li ><a href="/NewUser">NewUser</a></li>

		</ul>
	</div>
</div>
<div id="page">
	<div id="content">
		<div style="margin-bottom: 20px;">

			<p align="justify"><strong>
			 <form name="form1" method="post" action="/userlog">
	  <table width="389" height="214" border="0" align="center">
	    <tr>
	      <td colspan="2"><div align="center" class="style5">
	        <h2 class="style5">UserLogin </h2>
	      </div></td>
	    </tr>
	    <tr>
	      <td height="37"><h2 class="style5">EmailId  </h2></td>
	      <td><label>
	        <input name="_id" type="text" id="_id" />
	      </label></td>
	    </tr>
	    <tr>
	      <td><h2 class="style5">Password</h2></td>
```

```
                    <td><label>
                      <input name="psw" type="password" id="psw" />
                    </label></td>
                  </tr>
                  <tr>
                    <td> </td>
                    <td>
                      <input type="submit" name="Submit" value="Submit" />  <input
        type="reset" name="reset" value="Reset" />
                    </td>
                  </tr>
                </table>
                        </form>
                            </p>
                            <p> </p>
                            </blockquote>
              </div>
                    <div> </div>
              <div class="twocols"></div>
            </div>
            <!-- end content -->
            <!-- end sidebar -->
            <div style="clear: both;"> </div>
        </div>
        <!-- end page -->
        <div id="footer">
            <p id="legal"> <a href="#" title="This page validates as CSS"><abbr
        title="Cascading Style Sheets"></abbr></a></p>
        </div>
        <div align=center> <a href='#'></a></div>
        </body>
</html>
```

//userhome

```
<html xmlns="">
<head>
<meta http-equiv="content-type" content="text/html; charset=utf-8" />
<title>Secured data communication</title>
<meta name="keywords" content="" />
<meta name="description" content="" />
<link href="default.css" rel="stylesheet" type="text/css" />
<style>

/*
Design by Free CSS Templates
http://www.freecsstemplates.org
Released for free under a Creative Commons Attribution 2.5 License
*/

body {
	margin: 0;
	padding: 0;
	background: #FFFFFF url(static/images/img01.gif) repeat-x;
	font-family: Georgia, "Times New Roman", Times, serif;
	font-size: 13px;
	color: #666666;
}

h1, h2, h3 {
	margin: 0;
	font-weight: normal;
	color: #3F586B;
}

h1 {
	font-size: 197%;
}
```

```
h2 {
	font-size: 167%;
}

h3 {
	font-size: 100%;
	font-weight: bold;
}

p, ol, ul {
	line-height: 170%;
}

p {
}

ol {
	margin-left: 0;
	padding-left: 0;
	list-style-position: inside;
}

ul {
	margin-left: 0;
	padding-left: 0;
	list-style: none;
}

ul li {
	padding-left: 15px;
	background: url(static/images/img07.gif) no-repeat 0px 7px;
}
```

```
blockquote {
	margin: 0;
	padding-left: 20px;
	font-style: italic;
}

blockquote * {
}

a {
	color: #FF5723;
}

a:hover {
	text-decoration: none;
	color: #1777B1;
}

img {
	border: none;
}

img.left {
	float: left;
	margin: 3px 15px 0 0;
}

img.right {
	float: right;
	margin: 3px 0 0 15px;
}

hr {
	display: none;
```

```
}

/* Header */

#header {
	width: 700px;
	height: 235px;
	margin: 0 auto;
	background: #A4C0C8 url(static/images/img02.jpg) no-repeat;
}

/* Logo */

#logo {
	height: 190px;
}

#logo h1, #logo h2 {
	text-align: center;
}

#logo h1 {
	padding-top: 40px;
	font-size: 350%;
}

#logo h2 {
	font-size: 150%;
}

#logo a {
	text-decoration: none;
	color: #3F586B;
}
```

```
/* Menu */

#menu {
	padding-top: 0;
	width: 798px;
}

#menu ul {
	margin: 0;
	padding: 10px 0 0 0;
	list-style: none;
	line-height: normal;
	text-align: center;
}

#menu li {
	display: inline;
	margin: 0;
	padding: 0;
}

#menu a {
	padding: 0 20px;
	text-decoration: none;
	font-size: 136%;
	font-weight: bold;
	color: #610720;
}

#menu a:hover {
	text-decoration: underline;
}
```

```
#menu .active a {
	color: #FFFFFF;
}

/* Page */

#page {
	width: 730px;
	margin: 0 auto;
	padding: 30px 0;
}

/* Content */

#content {
	float: left;
	width: 800px;
	padding-top: 8px;
}

.twocols {
}

.twocols .title {
	padding-bottom: 10px;
	border-bottom: 1px solid #97C984;
}

.twocols .col1, .twocols .col2 {
	width: 190px;
}

.twocols .col1 {
	float: left;
```

```
}

.twocols .col2 {
	float: right;
}

.twocols ul {
}

.twocols ul li {
	padding-left: 0;
}

/* Sidebar */

#sidebar {
	float: right;
	width: 260px;
}

.boxed {
	margin: 0 0 20px 0;
}

.boxed .title {
	width: 250px;
	height: 35px;
	margin: 0;
	padding: 10px 0 0 10px;
	background: #A8C3CB url(static/images/img03.jpg) no-repeat;
	font-size: 136%;
	color: #144B6B;
}
```

```
.boxed .content {
	padding: 20px;
	border: 1px solid #97C984;
	border-top: none;
}

.boxed h3 {
	margin: 0;
}

.boxed p, .boxed ul, .boxed ol {
	margin: 0;
	padding: 0;
	list-style: none;
	line-height: normal;
}

.boxed ul {
}

.boxed ul li {
	padding: 8px 0 8px 10px;
	background: url(static/images/img04.gif) no-repeat 0px 13px;
}

.boxed ul li.first {
	border: none;
}

/* Search */

#search {
}
```

```
#search form {
	margin: 0;
	padding: 0;
}

#search fieldset {
	margin: 0;
	padding: 0;
	border: none;
}

#search p {
	float: left;
	padding-top: 5px;
	font-size: 85%;
}

#searchinput {
	width: 210px;
	margin-bottom: 5px;
}

#searchsubmit {
	float: right;
}

/* Footer */

#footer {
	height: 100px;
	padding: 20px;
	background: #5F919E;
	border-top: 5px solid #4C747E;
}
```

```
#footer p {
	margin: 0;
	text-align: center;
	line-height: normal;
	font-size: 85%;
	color: #FFFFFF;
}

#footer a {
	color: #FFFFFF;
}
.style5 {color: #1A3B5C}
.style6 {font-size: 16px}
.style8 {color: #3d5b99}
</style>
</head>
<body>
<div id="header">
	<div id="logo">
			<h1><a href="#">Water Quality    </a></h1>
		<h2><a href="">Prediction</a></h2>
  </div>
	<div id="menu">
		<ul>
			<li><a href="/userhome">Home</a></li>


			<li><a href="/">Logout</a></li>
		</ul>
	</div>
</div>
<div id="page">
	<div id="content">
```

```
                <div style="margin-bottom: 20px;">

                        <p align="justify"><strong>
                          <form action="/predict" method="post"
enctype="multipart/form-data" name="form1">
            <table  border="0" align="center">
              <tr>
                <td colspan="2"><div align="center" class="style5">
                   <h2 class="style5"><span class="style8">Enter Water  Info
</span></h2>
                </div></td>
              </tr>


              <tr>
                <td height="37"><h2 class="style5 style6">year </h2></td>
                <td><label>
                   <input name="year" type="text" id="year" />
                </label></td>
              </tr>

                <tr>
                <td height="37"><h2 class="style5 style6">D.O. (mg/l) </h2></td>
                <td><label>
                   <input name="do" type="text" id="do" />
                </label></td>
              </tr>

                 <tr>
                <td height="37"><h2 class="style5 style6">PH </h2></td>
                <td><label>
                   <input name="ph" type="text" id="ph" />
                </label></td>
              </tr>
```

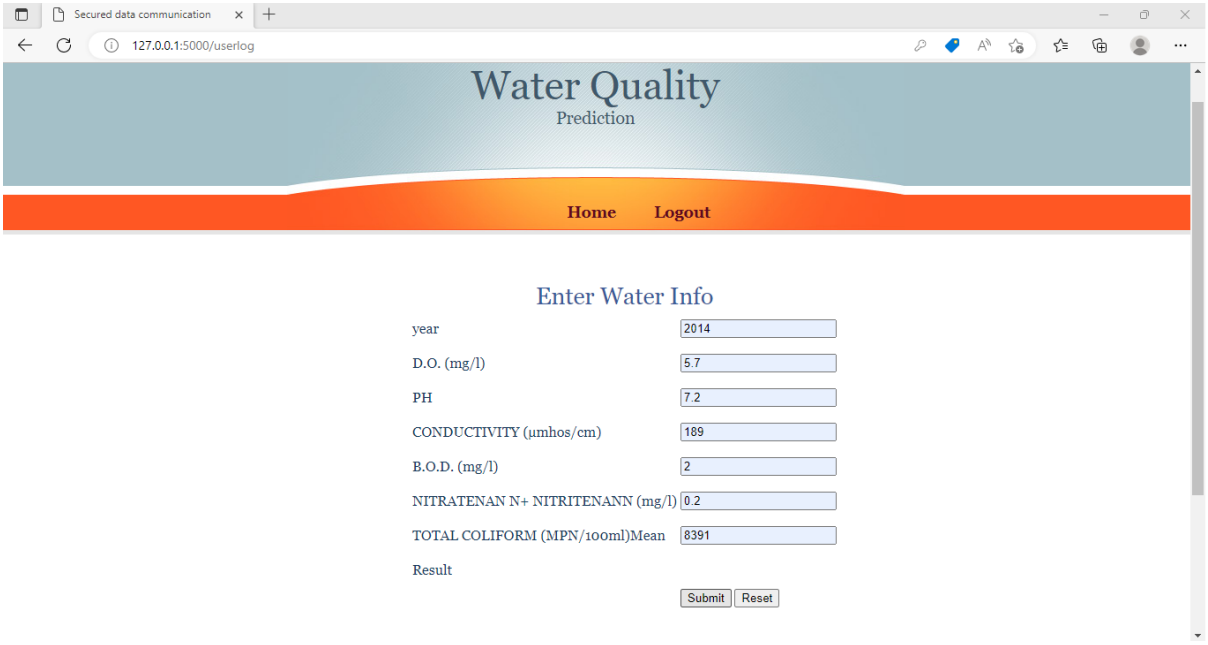
```
  <tr>
  <td height="37"><h2 class="style5 style6">CONDUCTIVITY
(µmhos/cm) </h2></td>
  <td><label>
    <input name="co" type="text" id="co" />
  </label></td>
 </tr>

  <tr>
  <td height="37"><h2 class="style5 style6">B.O.D. (mg/l) </h2></td>
  <td><label>
    <input name="bod" type="text" id="bod" />
  </label></td>
 </tr>

  <tr>
  <td height="37"><h2 class="style5 style6">NITRATENAN N+
NITRITENANN (mg/l) </h2></td>
  <td><label>
    <input name="na" type="text" id="na" />
  </label></td>
 </tr>

  <tr>
  <td height="37"><h2 class="style5 style6">TOTAL COLIFORM
(MPN/100ml)Mean </h2></td>
  <td><label>
    <input name="tc" type="text" id="tc" />
  </label></td>
 </tr>
 <tr>
  <td height="37"><h2 class="style5 style6"> Result </h2></td>
  <td>
```

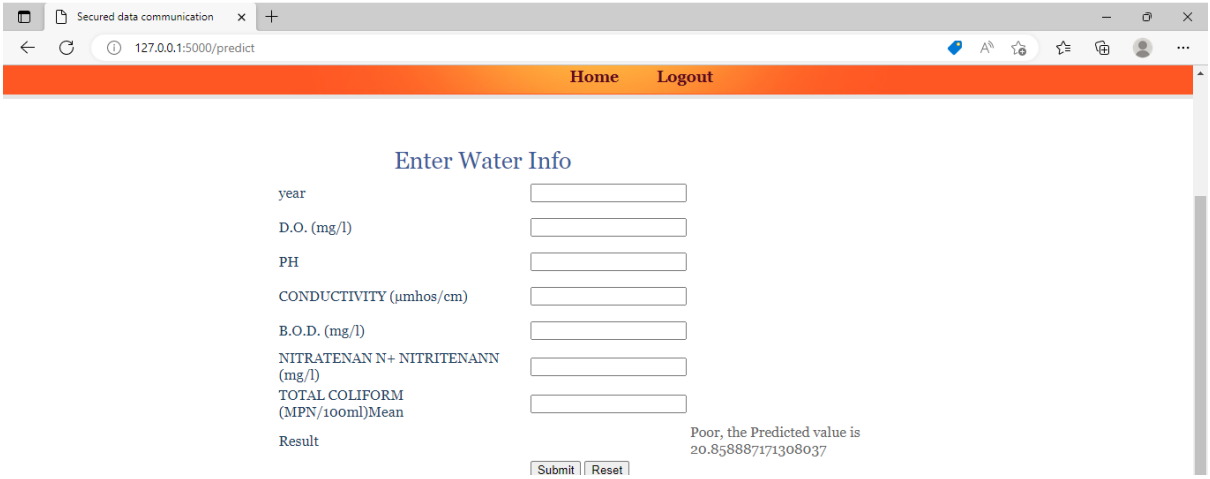
```
                    <td>{{prediction}}</td>
                      </td>
                    </tr>
                    <tr>
                      <td> </td>
                      <td><label>
                        <input type="submit" name="Submit" value="Submit" />
                        <input type="reset" name="Submit2" value="Reset">
                      <a href="user_reg.jsp"></a></label></td>
                    </tr>
                  </table>
                            </form>
                                </p>
                                <p> </p>
                                </blockquote>
                </div>
                        <div> </div>
                  <div class="twocols"></div>
                </div>
                <!-- end content -->
                <!-- end sidebar -->
                <div style="clear: both;"> </div>
        </div>
        <!-- end page -->
        <div id="footer">
                <p id="legal"> <a href="#" title="This page validates as CSS"><abbr
        title="Cascading Style Sheets"></abbr></a></p>
        </div>
        <div align=center> <a href='#'></a></div>
        </body>
</html>
```

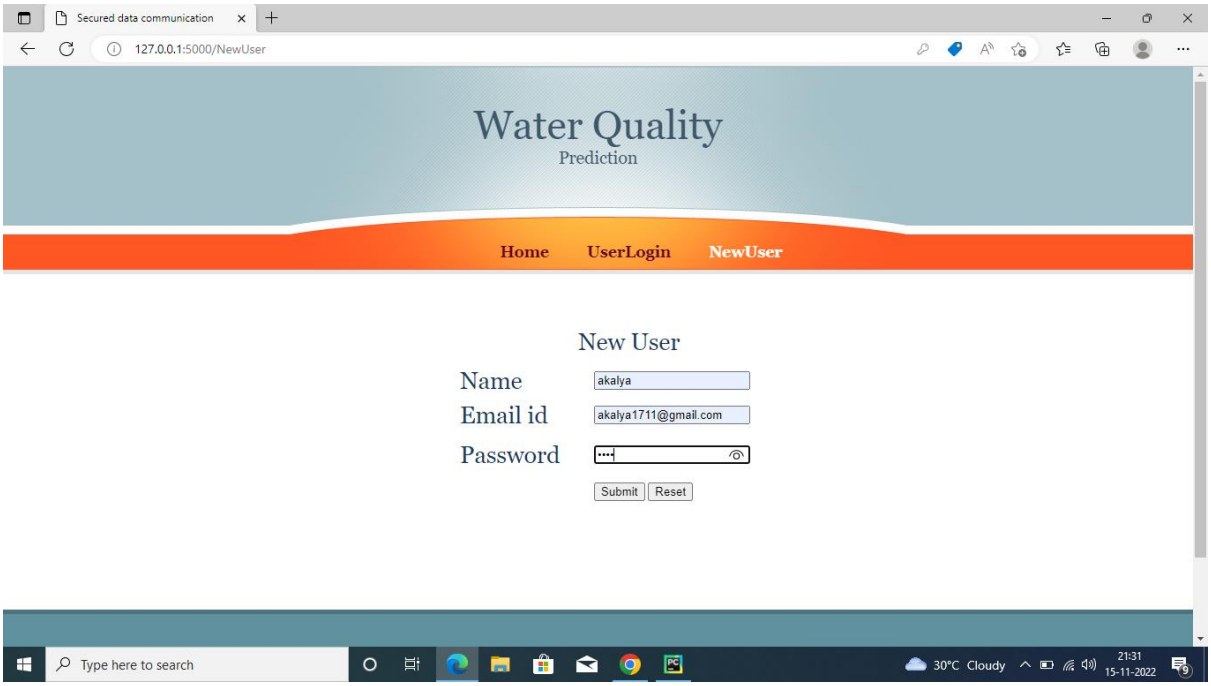
APPENDIX-2

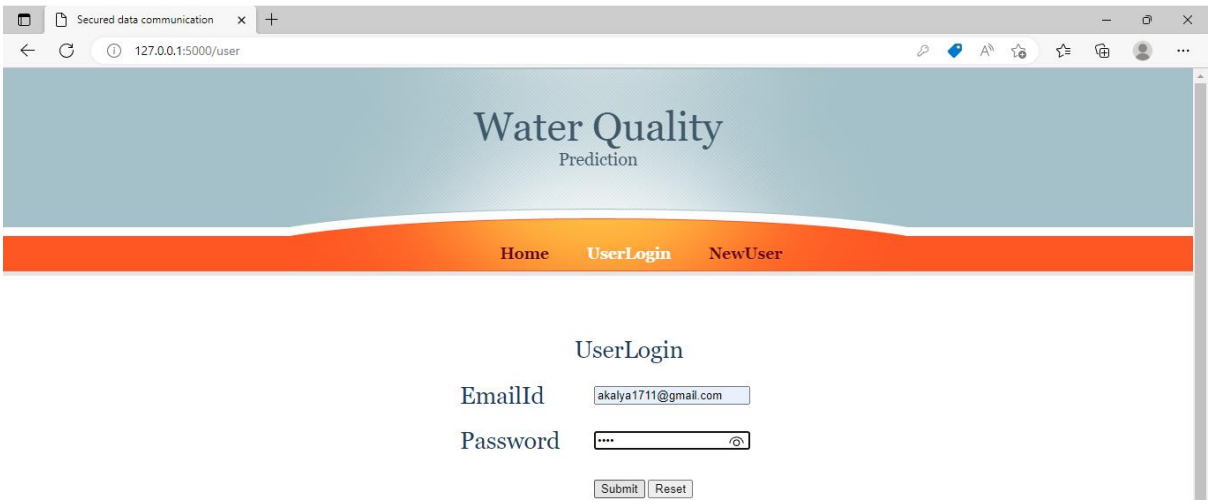
OUTPUT:

SCREENSHOTS:






Secured data communication
127.0.0.1:5000/NewUser
Water Quality
Prediction
Home
UserLogin
NewUser
New User
Name
akalya
Email id
akalya1711@gmail.com
Password
Submit
Reset
Type here to search
30°C Cloudy
21:31
15-11-2022


Secured data communication
127.0.0.1:5000/user
Water Quality
Prediction
Home
UserLogin
NewUser
UserLogin
EmailId
akalya1711@gmail.com
Password
Submit
Reset

GITHUB & PROJECT DEMO LINK:

https://github.com/IBM-EPBL/IBM-Project-4883-1658741757