PROJECT BASED EXPERIENTIAL LEARNING PROGRAM (NALAIYA THIRAN)

EFFICIENT WATER QUALITY ANALYSIS AND PREDICTION

A PROJECT REPORT

Submitted by

Y Dorairaj Adarsh (111719106033)
E Sri Sai Krishna (111719106035)
G Jagadeesh Varma (111719106037)
K Hariharan (111719106049)

TEAM ID: PNT2022TMID15978

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

R.M.K. ENGINEERING COLLEGE
(An Autonomous Institution)
R.S.M. Nagar, Kavaraipettai-601 206

NOVEMBER 2022

TABLE OF CONTENTS

1.INTRODUCTION

The chemical, physical, and biological properties of water can be referred to as its quality, generally in relation to how well suited it is for a certain purpose. Analysing the water's quality involves taking the necessary measurements and comparing them to the appropriate standards using established methodologies. Analysis of the quality of the water is necessary primarily for monitoring purposes. The requirements for a particular usage of the water determine all of the water quality criteria.

1.1 PROJECT OVERVIEW

Water is seen as a crucial resource that has an impact on many elements of human health and existence. People who live in metropolitan areas are often concerned about the quality of the water. The cornerstone for the prevention and management of waterborne infections is the quality of the water, which is a significant environmental influence. As a result, this project aims to develop a Machine Learning (ML) model to Predict Water Quality by taking into account all water quality standard indicators. However, this is a challenging task because the water quality varies in urban spaces non-linearly and depends on numerous factors, such as meteorology, water usage patterns, and land uses.

1.2. PURPOSE

The foundation for ensuring safe water use is water quality monitoring. Because drinking water quality is directly related to people's health, every step in the production of domestic water must be strictly inspected. Only after performing a good quality sampling survey can the water be sent to the water supply network.

Testing for water quality is a crucial component of environmental monitoring. Poor water quality has an impact on the surrounding ecology in addition to aquatic life. Each factor that affects the water quality in the environment is covered in detail in these sections.

The reason water quality is crucial is that, if appropriate standards are upheld, it ensures that end consumers will stay healthy and functional. The end users might be consumers of healthy beverages, businesses working without obstacles brought on by off-spec water, or healthy natural ecosystems prospering in the absence of pollution. Each user has a concentration limit for each contaminant above which water of lower quality will be harmful.

Human danger arises from pollutants like pesticides, heavy metals, and solvents contaminating water supplies. Blood, lung, liver, kidney, and urinary bladder cancers can all be made more likely by long-term exposure to heavy metals such arsenic, chromium, lead, mercury, and cadmium.

2. LITERATURE SURVERY

1. MACHINE LEARNING ALGORITHMS FOR EFFICIENT WATER QUALITY PREDICTION

Mourade Azrour(Université Moulay Ismail), Ghizlane Fattah(Mohammadia School of Engineers), Azidine Guezzaz(Cadi Ayyad University Morocco), Jamal Mabrouki(Mohammed V University of Rabat)

PUBLISHED : 26 AUGUST 2021

In this study, the advantages of machine learning algorithms are used to develop a model that is capable of predicting the water quality index and then the water quality class. The method we propose is based on four water parameters: temperature, pH, turbidity and coliforms. The use of the multiple regression algorithms has proven to be important and effective in predicting the water quality index. In addition, the adoption of the artificial neural network provides the most highly efficient way to classify the water quality.

2.WATER QUALITY PREDICTION BASED ON MACHINE LEARNING TECHNIQUES

Zhao Fu(Bachelor of Information Security China University of Geosciences (Wuhan))

PUBLISHED: 8 JANUARY 2020

In this dissertation, several methods have been proposed to improve the performance of ANFIS-based water quality prediction models. Stratified sampling is employed to cover different kinds of data distribution in the training and testing datasets. The wavelet denoising technique is used to remove the noise hidden in the dataset. A deep prediction performance comparison between MLR, ANN, and ANFIS model is presented after stratified sampling and wavelet denoising techniques are applied. Because water quality data can be thought as a time series dataset, a time series analysis method is integrated with the ANFIS model to improve prediction performance. Lastly, intelligence algorithms are used to optimize the parameters of membership functions in the ANFIS model to promote the prediction accuracy. Experiments based on water quality datasets collected from Las Vegas Wash since 2007 and Boulder Basin of Lake Mead, Nevada, between 2011 and 2016 are used to evaluate the proposed models. Various ANN models have been designed to predict water and wastewater discharge quality based on previous existing datasets. A

twolayer ANN model has been applied to predict the DO concentration in the Mathura River, and the experimental result showed that the ANN model worked well. An ANFIS model with eight input parameters is used to predict total phosphorus and total nitrogen, the experiment result based on 120 water samples. Time series analysis is also proposed to address dissolve oxygen prediction, and the experimental results show that the proposed analysis method can find out valuable knowledge from water quality historical timeseries data In this , MLR, ANN, ANFIS, and FTS models are integrated with statistical analysis, wavelet denoising, and intelligence algorithm to explore the prediction of water quality.

3.GROUND WATER QUALITY PREDICTION USING MACHINE LEARNING ALGORITHMS IN R

S.Vijay & 2Dr.K.Kamaraj Assistant Professor,Assistant ProfessorDepartment of Computer Science, 2Department of Computer Science Vivekanandha College of Arts and Sciences for Women, Tiruchengode, Namakkal Dt, SSM College of Arts and Science, Komarapalayam, Namakkal.

PUBLISHED : MARCH 2019

Sundarambal Palani proposed ANN models to predict water quality parameters whereas salinity, temperature, dissolved oxygen and Chl-a concentrations using continuous weekly measurements at different locations.Changjun Zhu proposed fuzzy neural network(FNN) model to evaluate and classify outer water quality in suzhou. The FNN model is reliable and effective and can deal with the problem of solitary elements which reflects the water quality at current stage.Yafra Khan has developed a water quality forecast model using the support of water quality components applying Artificial Neural Network (ANN) and time-series analysis with ANN-NAR. The performance measures such as Regression, Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) indicated the best prediction accuracy results with ANN-NAR time series algorithm.

4.WATER QUALITY ANALYSIS USING MACHINE LEARNING ALGORITHMS

Aleksei Shkurin(Bachelor's Thesis in Environmental Engineering)

PUBLISHED : 29 MARCH 2016

Random forests (RF) show the best performance and are advised for scientists andengineers working with environmental data. Artificial neural networks (ANN) are another alternative, though their performance is inferior and they are prone to overfitting. Support vector machines (SVM) are the good example for the cases where a baseline model is needed, being one of the basic algorithms. K-nearest neighbours (KNN) model was successfully used for data imputation and is also suggested for this task for other researchers. Clusterization techniques, such as kmeans clustering, may assist data scientist with possible algorithms to classify given data, for example defining good, average and bad conditions of the water based on various chemical, biological and physical parameter models generated during this research may be used by IT students for producing software meant to help environmental specialists in analysing collected water quality data.

5.GROUND WATER QUALITY PREDICTION USING MACHINE LEARNING ALGORITHMS IN R

S.Vijay(AssistantProfessor, Department of Computer Science, Vivekanandha College of Arts and Sciences for Women, Tiruchengode, Namakkal Dt) and Dr.K.Kamaraj (Assistant Professor, Department of Computer Science, SSM College of Arts and Science, Komarapalayam, Namakkal

PUBLISHED : 01 MARCH 2019

Water plays a dominant role in the growth of the country's economy and essential for all the activities. The present study deals with the physico-chemical characteristics of ground water quality in Ranipet, Arcot, Walljah pet, towns in vellore district. Such a water samples were collected from different identified bore wells for the purpose of studying the quality of groundwater . The bore wells from which the samples were collected are extensively used for drinking purpose. The water quality parameters such as TDS, EC, Chloride, Sulphate, Nitrate, Carbonate, Bicarbonate, metal ions, trace elements have been estimated. There are two major classifications like High , Low level of water contamination observed in Vellore

district. This paper focus on predicting water quality by using Machine Learning classifier algorithm C5.0, Naïve Bayes and Random forest as leaner for water quality prediction with high accuracy any efficiency.

6.WATER QUALITY FACTOR PREDICTION USING SUPERVISED MACHINE LEARNINGMACHINE LEARNING

Kathleen Joslyn(Portland State University)

PUBLISHED : 2018

The objective of this research is to explore prediction accuracy of water quality factors with techniques and algorithms in machine learning consisting of a variation of support vector machines - Support Vector Regression (SVR) and the gradient boosting algorithm Extreme Gradient Boosting (XGBoost). Both the XGBoost and SVR algorithms were used to predict nine different factors with success rates ranging from 79% to 99%. Parameters of these algorithms were also explored to test the prediction accuracy levels of individual water quality factors. These parameters included normalizing the data, filling missing data points, and training and testing on a large set of data.

7.WATER QUALITY PREDICTION AND CLASSIFICATION BASED ON PRINCIPAL COMPONENT REGRESSION AND GRADIENT BOOSTING CLASSIFIER APPROACH

Md. SaikatIslam Khan(Department of Computer Science and Engineering, Santosh, Tangail-1902, Bangladesh) , NazrulIslam(Department of Information and Communication and Technology, Santosh, Tangail-1902, Bangladesh) , SifatulIslam Department of Technology Studies, Endicott College, Woosong University,
Daejeon, South Korea , Mostofa KamalNasir(Mawlana Bhashani Science and Technology University, Santosh, Tangail 1902, Bangladesh)

PUBLISHED : 3 JUNE 2021

Estimating water quality has been one of the significant challenges faced by the world in recent decades. This paper presents a water quality prediction model utilizing the principal component regression technique. Firstly, the water quality index (WQI) is calculated using the weighted arithmetic index method. Secondly,

the principal component analysis (PCA) is applied to the dataset, and the most dominant WQI parameters have been extracted. Thirdly, to predict the WQI, different regression algorithms are used to the PCA output. Finally, the Gradient Boosting Classifier is utilized to classify the water quality status. The proposed system is experimentally evaluated on a Gulshan Lake-related dataset. The results demonstrate 95% prediction accuracy for the principal component regression method and 100% classification accuracy for the Gradient Boosting Classifier method, which show credible performance compared with the state-of-art models.

2.1 EXISITNG PROBLEM

The major issue is here. We must run lab tests on the water to test its quality, which is both costly and time-consuming. As a result, in this study, we present an alternate technique to predicting water quality based on artificial intelligence. This method makes use of a significant and easily accessible water quality index established by the WHO. The information in this research comes from the PCPB India, which has 3277 cases of the distinct sources. The WQI (Water Quality Index) is computed utilising AI methods in this study. Therefore, in the future, we can combine this with an IoT-based framework to analyse large datasets and broaden the scope of our research. Compared to any other IoT framework, it can estimate the water quality using that method quickly and correctly. To verify parameters like pH, temperature, turbidity, and other variables, that IoT framework system employs various restrictions for the sensor. Additionally, after this parameter has been read, send these measurements to the Arduino microcontroller and ZigBee device for additional prediction.

2.2 REFERENCES

1.GROUND WATER QUALITY PREDICTION USING MACHINE LEARNING ALGORITHMS IN R
S.Vijay & 2Dr.K.Kamaraj Assistant Professor,Assistant ProfessorDepartment of Computer Science, 2Department of Computer Science Vivekanandha College of Arts and Sciences for Women, Tiruchengode, Namakkal Dt, SSM College of Arts and Science, Komarapalayam, Namakkal. PUBLISHED : MARCH 2019

2.WATER QUALITY ANALYSIS USING MACHINE LEARNING ALGORITHMS
Aleksei Shkurin(Bachelor's Thesis in Environmental Engineering)
PUBLISHED : 29 MARCH 2016

3.GROUND WATER QUALITY PREDICTION USING MACHINE LEARNING ALGORITHMS IN R

S.Vijay(AssistantProfessor, Department of Computer Science, Vivekanandha College of Arts and Sciences for Women, Tiruchengode, Namakkal Dt) and Dr.K.Kamaraj (Assistant Professor, Department of Computer Science, SSM College of Arts and Science, Komarapalayam, Namakkal

PUBLISHED : 01 MARCH 2019

4. EFFICIENT PREDICTION OF WATER QUALITY INDEX (WQI) USING MACHINE LEARNING ALGORITHMS

Md. Mehedi Hassan (Computer Science and Engineering, North Western University, Khulna, Bangladesh), Md. Mahedi Hassan(Computer Science and Engineering, Bangladesh University of Business and Technology, Dhaka, Bangladesh), Laboni Akter(Biomedical Engineering, Khulna University of Engineering & Technology, Khulna, Bangladesh), Md. Mushfiqur Rahman, Sadika

Zaman, Khan Md. Hasib, Nusrat Jahan, Raisun Nasa Smrity, Jerin Farhana, M. Raihan, Swarnali Mollick

PUBLISHED : 22 November 2021

5. WATER QUALITY PREDICTION AND CLASSIFICATION BASED ON PRINCIPAL COMPONENT REGRESSION AND GRADIENT BOOSTING CLASSIFIER APPROACH

Md. SaikatIslam Khan(Department of Computer Science and Engineering, Santosh, Tangail-1902, Bangladesh) , NazrulIslam(Department of Information and Communication and Technology, Santosh, Tangail-1902, Bangladesh) , SifatulIslam Department of Technology Studies, Endicott College, Woosong University,

Daejeon, South Korea , Mostofa KamalNasir(Mawlana Bhashani Science and Technology University, Santosh, Tangail 1902, Bangladesh). PUBLISHED : 3 JUNE 2021

6. PERFORMANCE OF MACHINE LEARNING METHODS IN PREDICTING WATER QUALITY INDEX BASED ON IRREGULAR DATA SET: APPLICATION ON ILLIZI REGION (Algerian southeast)

Saber Kouadri, Ahmed Elbeltagi, Abu Reza Md. Towfiqul Islam & Samir Kateb

PUBLISHED : 06 NOVEMBER 2021

2.3 PROBLEM STATEMENT DEFINITION

To determine if the water is safe to drink is important for health, is a fundamental human right, and is a key element of successful health protection policies. On a national, regional, and local level, this is significant as a health and development concern. Investments in water supply and sanitation have been shown to produce a net economic benefit in some areas because they reduce negative health effects and medical expenses more than they cost to implement.

3.IDEATION AND PROPOSED SOLUTION

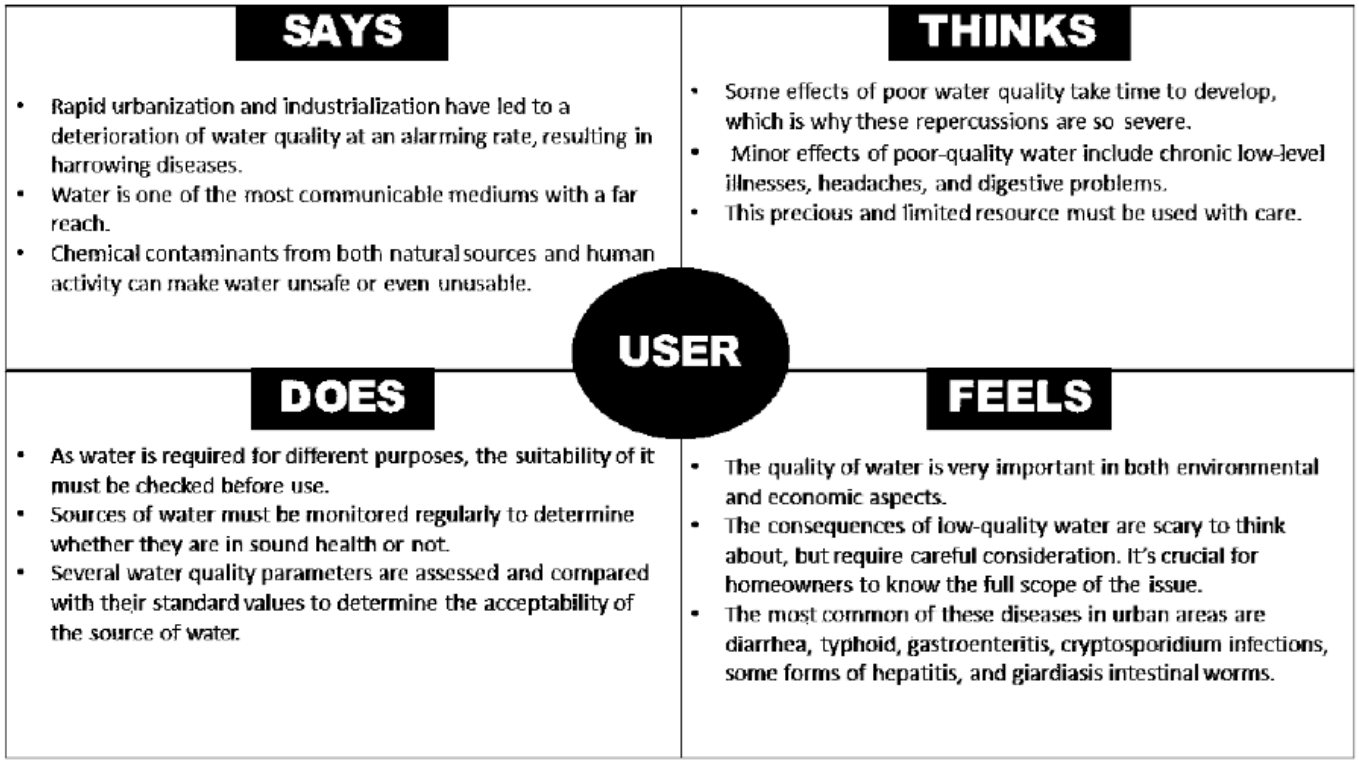
3.1 EMPATHY MAP CANVAS

Empathy Map Canvas:

An empathy map is a straightforward, easy-to-understand picture that gathers information about a user's behaviours and attitudes.

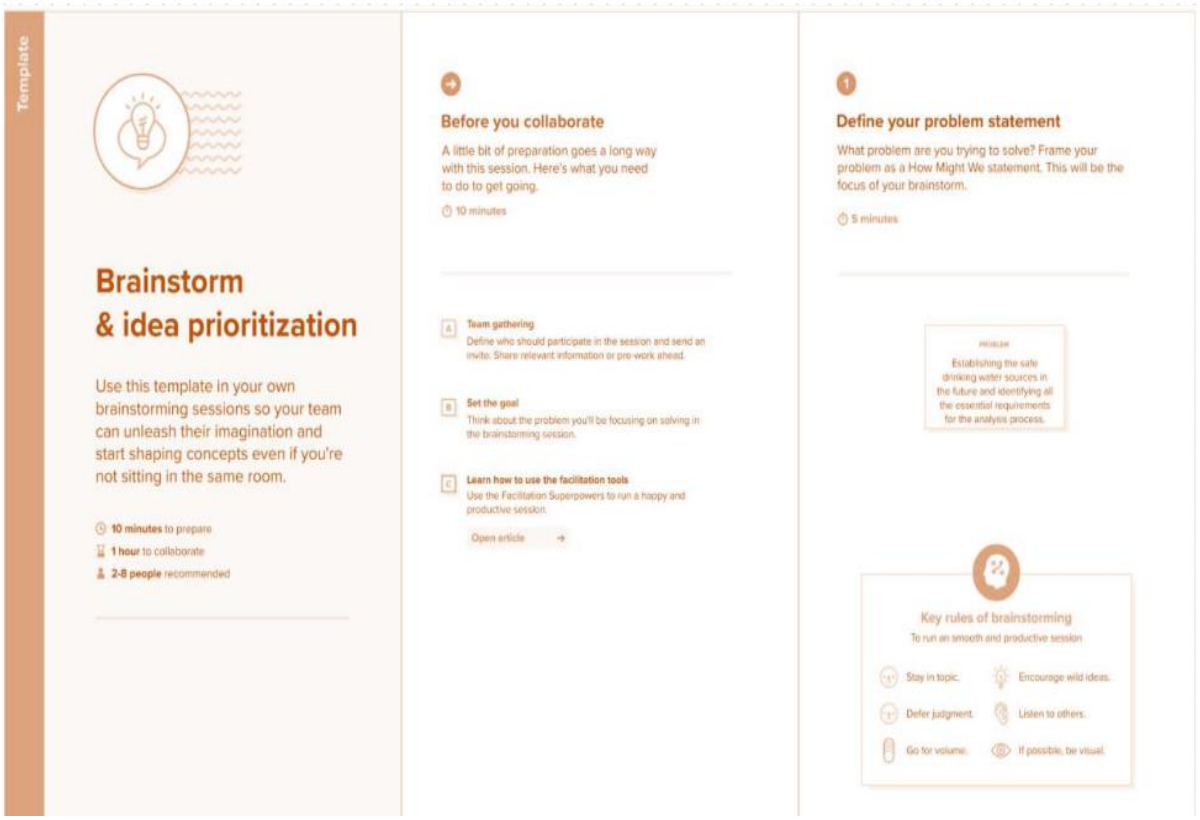
It is a valuable tool for assisting teams in better understanding their users.
Understanding the real problem and the person experiencing it is necessary for developing an effective solution. The map-making activity lets participants analyse things from the user's point of view, as well as his or her goals and obstacles.

EMPATHY MAP


SAYS
Rapid urbanization and industrialization have led to a deterioration of water quality at an alarming rate, resulting in harrowing diseases.
Water is one of the most communicable mediums with a far reach.
Chemical contaminants from both natural sources and human activity can make water unsafe or even unusable.
THINKS
Some effects of poor water quality take time to develop, which is why these repercussions are so severe.
Minor effects of poor-quality water include chronic low-level illnesses, headaches, and digestive problems.
This precious and limited resource must be used with care.
USER
DOES
As water is required for different purposes, the suitability of it must be checked before use.
Sources of water must be monitored regularly to determine whether they are in sound health or not.
Several water quality parameters are assessed and compared with their standard values to determine the acceptability of the source of water.
FEELS
The quality of water is very important in both environmental and economic aspects.
The consequences of low-quality water are scary to think about, but require careful consideration. It's crucial for homeowners to know the full scope of the issue.
The most common of these diseases in urban areas are diarrhea, typhoid, gastroenteritis, cryptosporidium infections, some forms of hepatitis, and giardiasis intestinal worms.

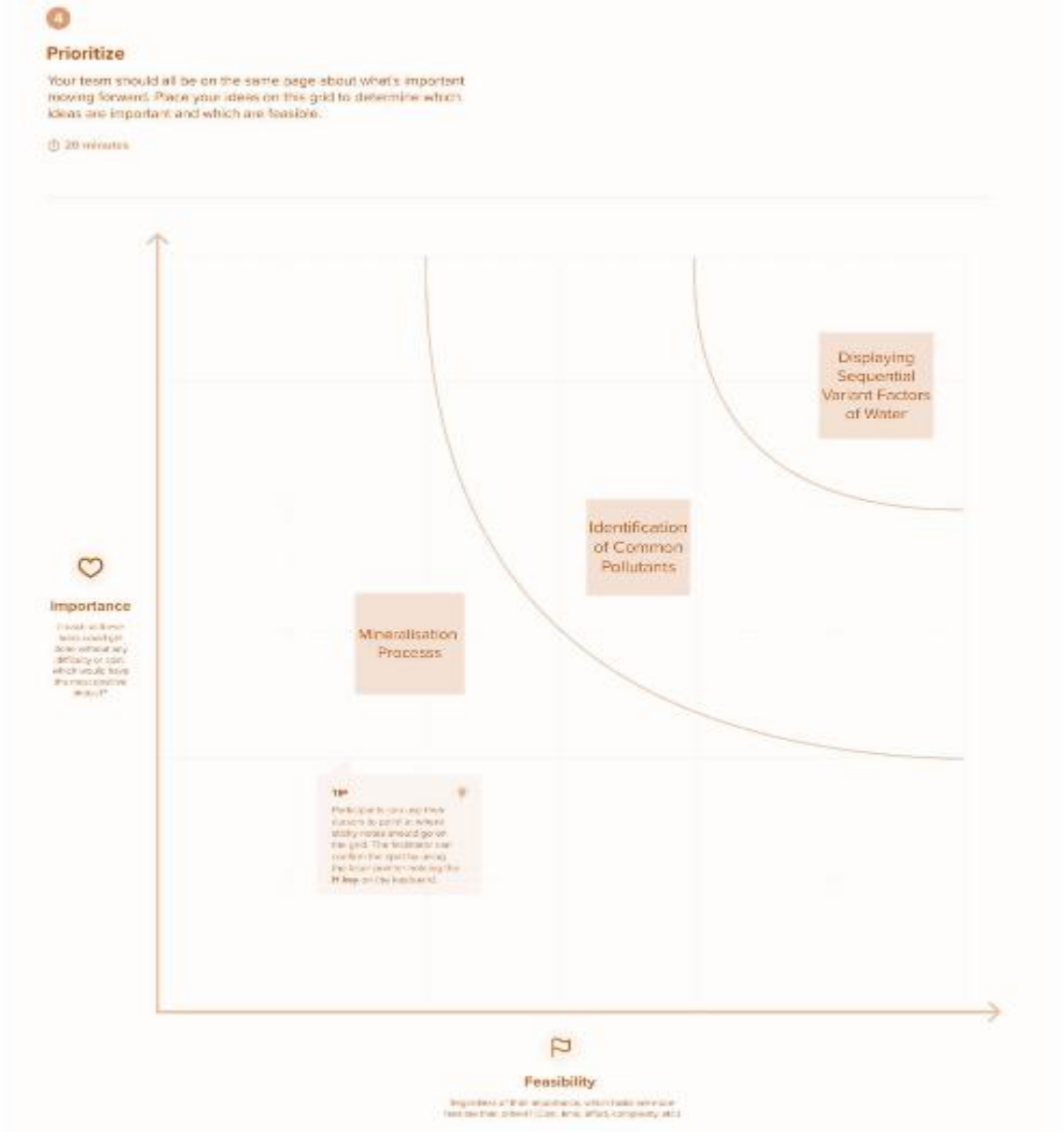
3.2 IDEATION AND BRAINSTORMING

Step-1: Team Gathering, Collaboration and Select the Problem Statement



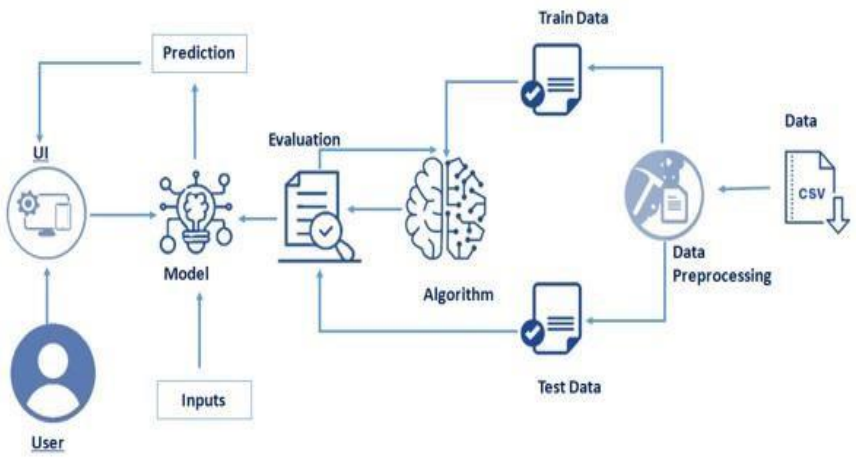
Step-2: Brain Storm, Idea Listing, Grouping

Step -3: Idea Prioritization

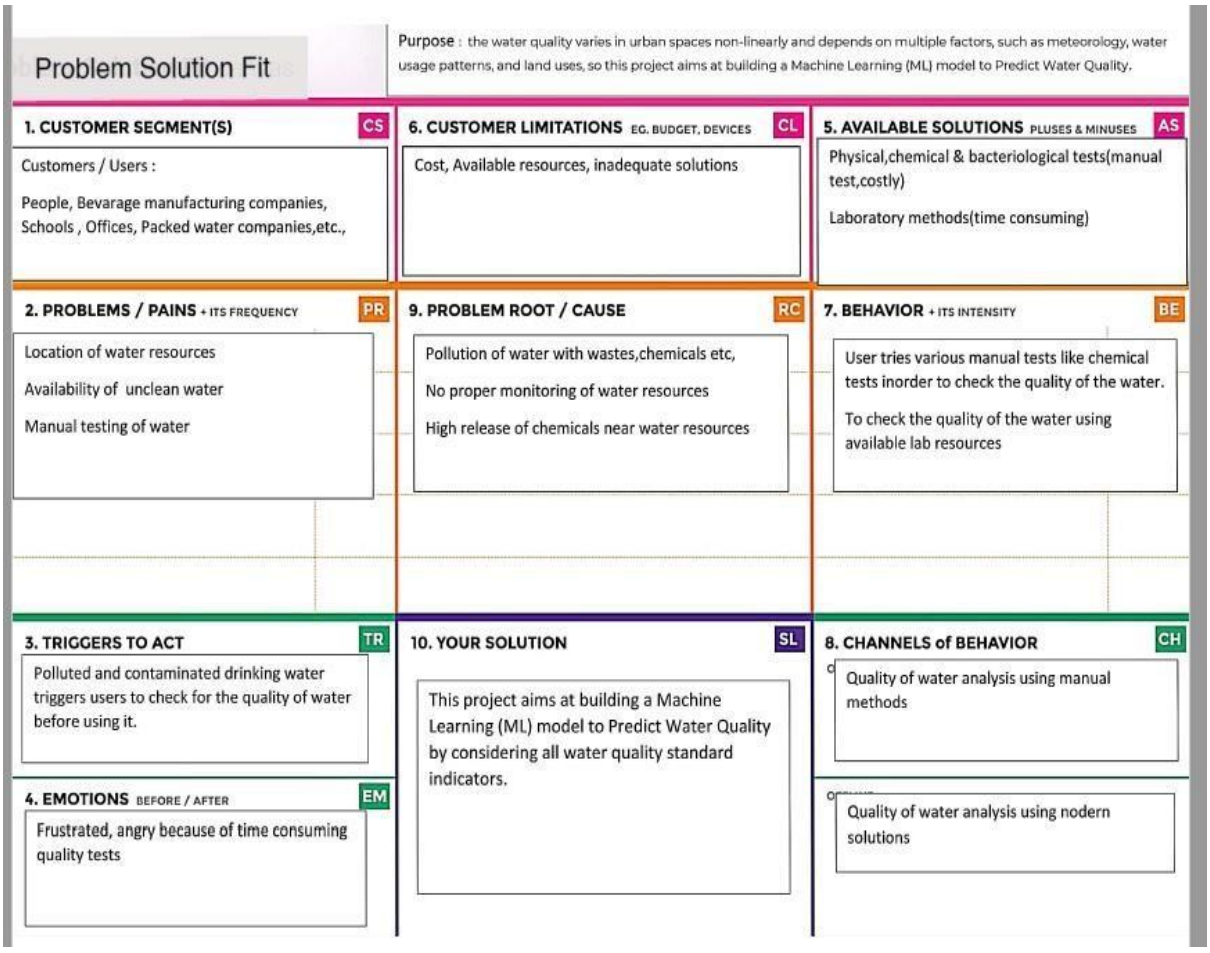


3.3 PROPOSED SOLUTION

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	Efficient Water Quality Analysis & Prediction Using Machine Learning Water is considered as a vital resource that affects various aspects of human health and lives. The quality of water is a major concern for people living in urban areas. The quality of water serves as a powerful environmental determinant and a foundation for the prevention and control of waterborne diseases. However predicting the urban water quality is a challenging task since the water quality varies in urban spaces nonlinearly and depends on multiple factors, such as meteorology, water usage patterns, and land uses, so this project aims at building a Machine Learning (ML) model to Predict Water Quality by considering all water quality standard indicators.
2.	Idea / Solution description	You will need to train the datasets to run smoothly and see an incremental improvement in the prediction rate. Different measures are used to assess the accuracy of the applied machine learning algorithms.
3.	Novelty / Uniqueness	Since the results from various algorithms are compared we can identify which method gives the highest accuracy.
4.	Social Impact / Customer Satisfaction	Assessing customer satisfaction is critical for companies looking to gain competitive advantage in the market. Therefore, customer satisfaction is tied to the standards that apply to different areas of activity, including water quality analysis. The present work aims to assess the satisfaction of the customers based on parameters that influence the management requirements and the technical requirements of the ISO/IEC 17025 standard.

5.	Business Model (Revenue Model)	
6.	Scalability of the Solution	Reliable water quality prediction can improve environmental flow monitoring and the sustainability of the stream ecosystem.

3.3 PROBLEM SOLUTION FIT



4. REQUIREMENT ANALYSIS

4.1 FUNCTIONAL REQUIREMENT

Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	Install Google colab and download libraries.	Install google collab and Import all the required libraries which are used to train the model or visualise the data.
FR-2	Dataset	Initial process 1.Download and import the dataset 2.Read the dataset
FR-3	WQI(water quality index)	The outcome to be found from the dataset, 1.calculate the quality index for each column. 2. Calculate the avg of WQI.
FR-4	Application Building	Use flask architecture which is used to create a user interface . 1.It accepts the individual inputs**(year,D.O, P.H,** **B.O.D, C.O, N,A, T.C)** and inturn produce the WQI as output
FR-5	Interface sensor	Confirmation via email and it is predicted by water level sensor

4.2 NON-FUNCTIONAL REQUIREMENT

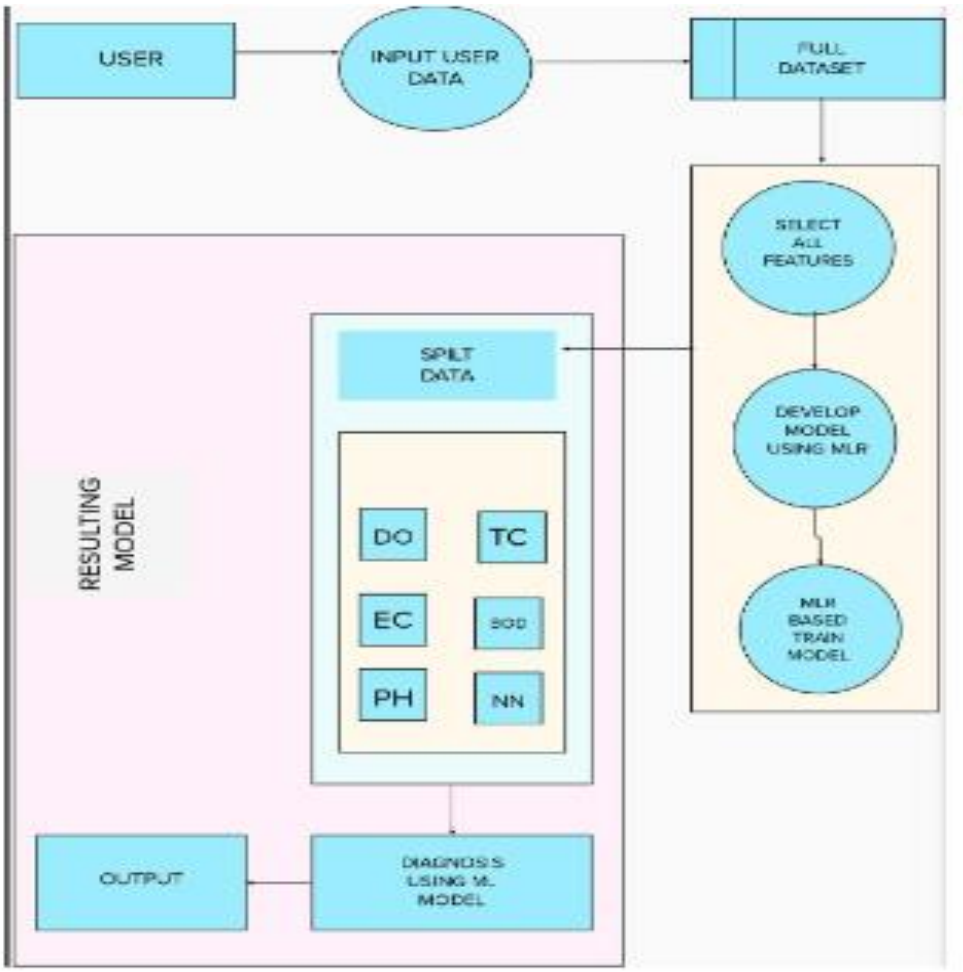
Following are the non-functional requirements of the proposed solution.

NFR-1	**Usability**	The aim of this model is to predict the WQI(water quality index) based on some factors like(PH, B.O.D, Conductivity etc..). WQI helps in determining overall water quality status. Accurate water quality prediction is the basis of water environment management and is of great significance for water environment protection.
NFR-2	**Security**	It provides secured feel for the people while drinking water
NFR-3	**Reliability**	This project helps in protecting people health and the environment.
NFR-4	**Performance**	PH, Turbidity, temperature etc are calculated by sensors and recorded; the data is pre processed and WQI is calculated
NFR-5	**Availability**	By developing and deploying the software we can analyse the drinking water
NFR-6	**Scalability**	The project helps in providing a purified water and pollution free water

5.PROJECT DESIGN

5.1 DATA FLOW DIAGRAM

A Data Flow Diagram (DFD) is a classic visual depiction of a system's information flows. A tidy and clear DFD may graphically display the appropriate quantity of system need. It demonstrates how data enters and exits the system, what changes the data, and where data is stored.



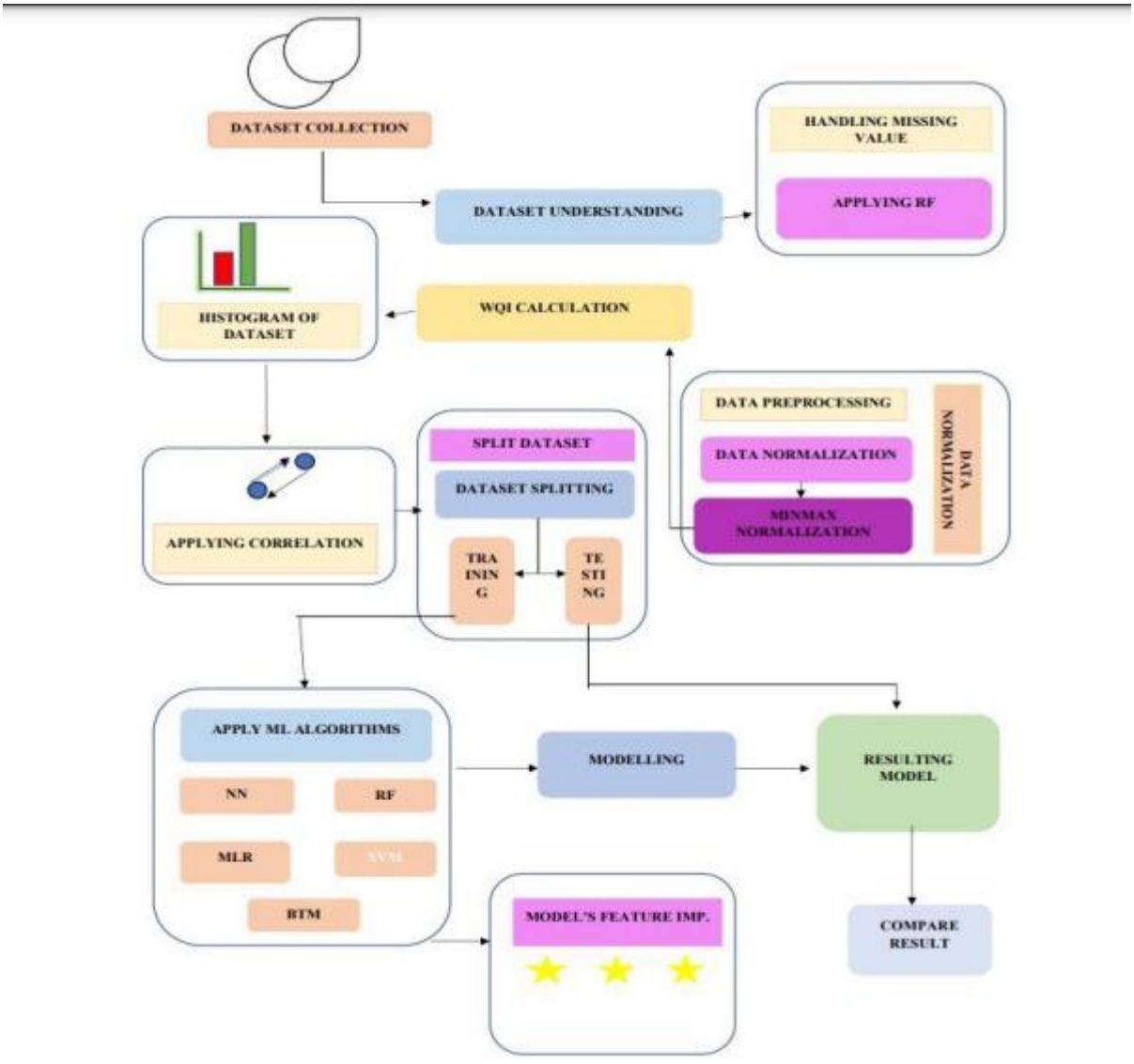
5.2 SOLUTION AND TECHNINCAL ARCHITECTURE

Solution architecture is a multi-step approach that bridges the gap between business challenges and technological solutions. Its objectives are as follows: • Find the best technological solution to existing business difficulties.

• Explain to project stakeholders the structure, features, behaviour, and other elements of the programme.

• Specify features, phases of development, and solution requirements.

• Provide requirements for defining, managing, and delivering the solution.



5.3 USER STORIES

User Type	Functional Requirement	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Customer (Mobile user)	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can access my account / dashboard.	High	Sprint-1
		USN-2	As a user, I will receive confirmation email once I have registered for the application.	I can receive confirmation email & click confirm.	High	Sprint-1
		USN-3	As a user, I can register for the application through Facebook.	I can register & access the dashboard with Facebook Login.	Low	Sprint-2
		USN-4	As a user, I can register for the application through		Medium	Sprint-1

			Gmail			
	Login	USN-5	As a user, I can log into the application by entering email & password		High	Sprint-1
	Dashboard	USN-6	As a user, I can check my login details and work details		High	Sprint-1
Customer (Web user)	Web Access	USN-7	As a user, I can enter the values about the water.	I can access the webpage through internet.	High	Sprint-1
		USN-8	As a user, I can submit the values into the webpage.	I can click the submit button.	High	Sprint-2
		USN-9	As a user, I expect correct coefficient of water.		Medium	Sprint-3
	Data preprocessing	USN-10	As a user, I can see the loading information.		Medium	Sprint-3
	User Input Evaluation	USN-11	I can see the evaluation		High	Sprint-4

			quickly.			
	Prediction	USN-12	As a user, I can see the result of the water efficient.	The results are visible on webpage.	High	Sprint-4
Customer Care Executive	Solving Customer issues.	USN-13	As a customer care executive, I solve the customer issues in using the application and webpage.	It results in user interaction.	Medium	Sprint-5
Administrator		USN-14	I can manage the application.		Medium	Sprint-5

6. PROJECT PLANNING AND SCDULING

6.1 SPRINT PLANNING AND ESTIMATION

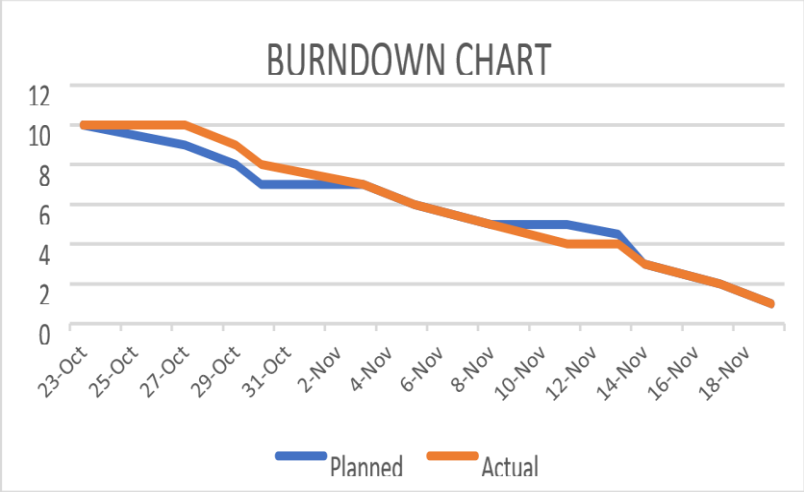
Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint 1	Dataset collection	USN-1	Collect the required data for the water quality prediction	3	High	Y.Dorairaj Adarsh E.Sri Sai Krishna G. Jagadeesh varma K.Hariharan
Sprint 1	Data preprocessing	USN-2	Perform data cleaning to optimize the dataset	7	Medium	Y.Dorairaj Adarsh E.Sri Sai Krishna G. Jagadeesh varma K.Hariharan
Sprint 2	Training & Building Model , Model evaluation	USN-3	Build the model using regression algorithms to classify the data	10	High	Y.Dorairaj Adarsh E.Sri Sai Krishna G. Jagadeesh varma K.Hariharan
Sprint 3	Application Building	USN-5	Build the html and python code. Run flak app.	7	High	Y.Dorairaj Adarsh E.Sri Sai Krishna G. Jagadeesh varma K.Hariharan
		USN-6	Run Flask app.	3	Medium	Y.Dorairaj Adarsh E.Sri Sai Krishna G. Jagadeesh varma K.Hariharan

Sprint 4	Implementation of the Application	USN-7	Deploy the model on IBM cloud	10	Medium	Y.Dorairaj Adarsh E.Sri Sai Krishna G. Jagadeesh varma K.Hariharan

6.2 Sprint Delivery Schedule

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	20	4 Days	25 Oct 2022	28 Oct 2022	20	01 Nov 2022
Sprint-2	20	4 Days	29 Oct 2022	1 Nov 2022	20	04 Nov 2022
Sprint-3	20	4 Days	02 Nov 2022	5 Nov 2022	20	11 Nov 2022
Sprint-4	20	4 Days	12 Nov 2022	15 Nov 2022	20	19 Nov 2022

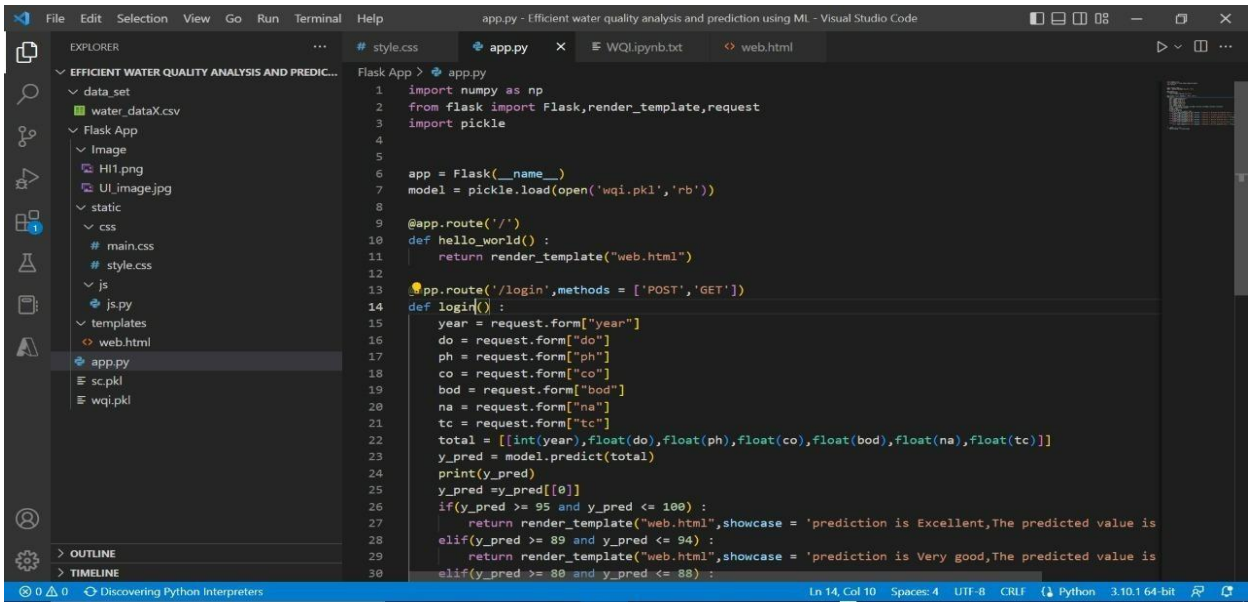
Burndown Chart:

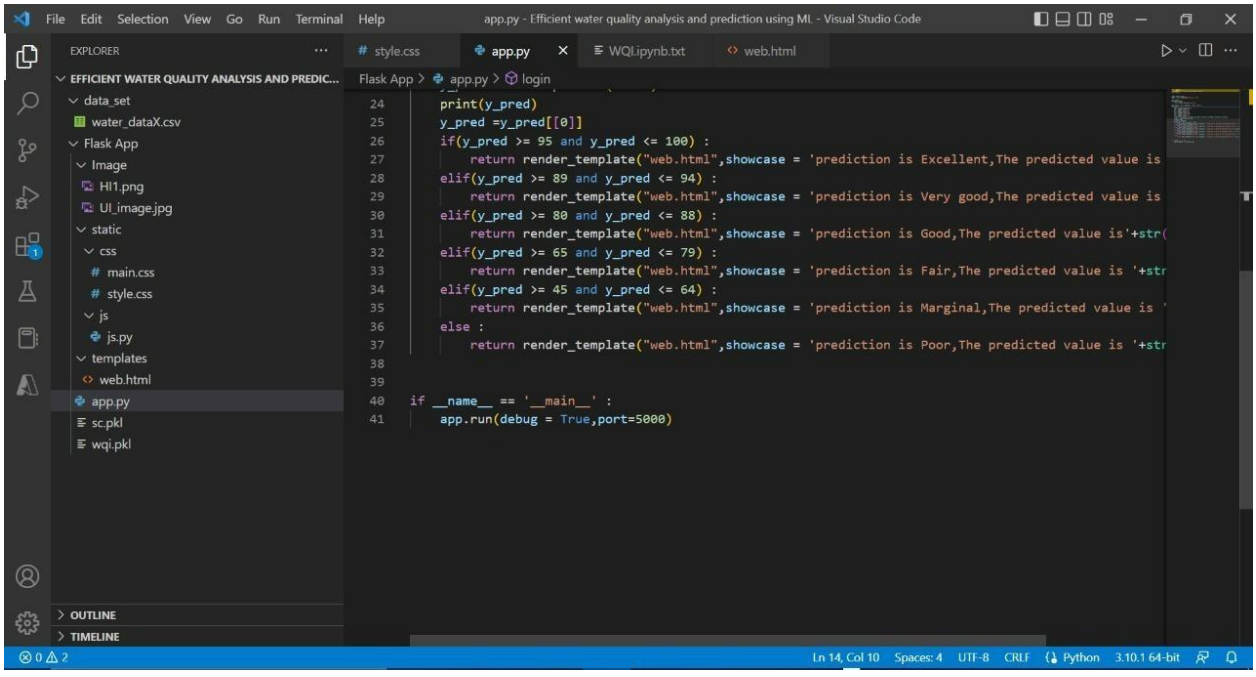


7. CODING AND SOLUTIONS

7.1 FEATURE 1

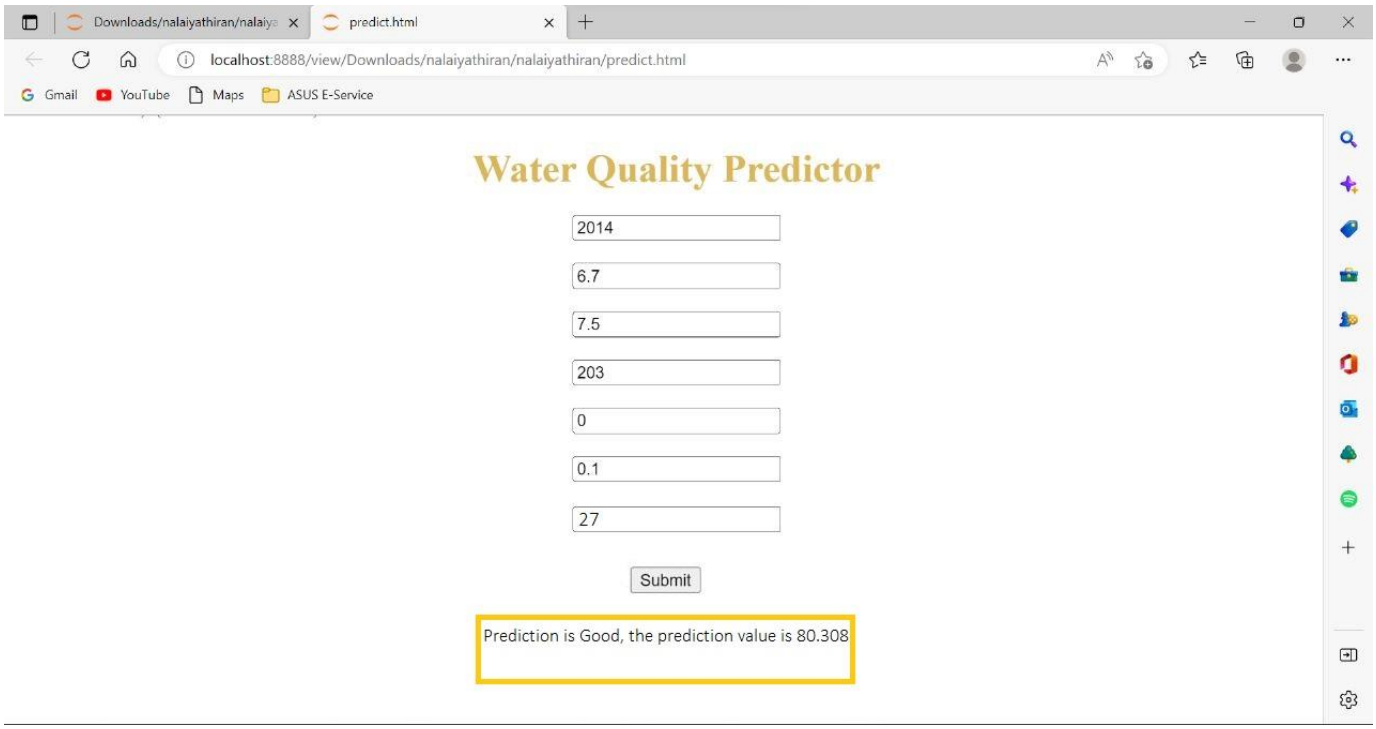
Flask is a web framework. This means flask provides you with tools, libraries and technologies that allow you to build a web application. This web application can be some web pages, a blog, a wiki or go as big as a web-based calendar application or a commercial website.





8.TESTING

8.1 TEST CASE



8.2 USER ACCEPTANCE TESTING

1. Purpose of Document

The purpose of this document is to briefly explain the test coverage and open issues of the Efficient water quality analysis and prediction using machine learning project at the time of the release to User Acceptance Testing (UAT).

2. Defect Analysis

This report shows the number of resolved or closed bugs at each severity level, and how they were resolved.

Resolution	Severity 1	Severity 2	Severity 3	Subtotal
By Design	2	4	2	8
Duplicate	0	0	0	0
Fixed	2	2	0	4
Not Reproduced	0	0	0	0
Skipped	0	0	1	1
Won't Fix	0	0	0	0
Totals	4	6	1	13

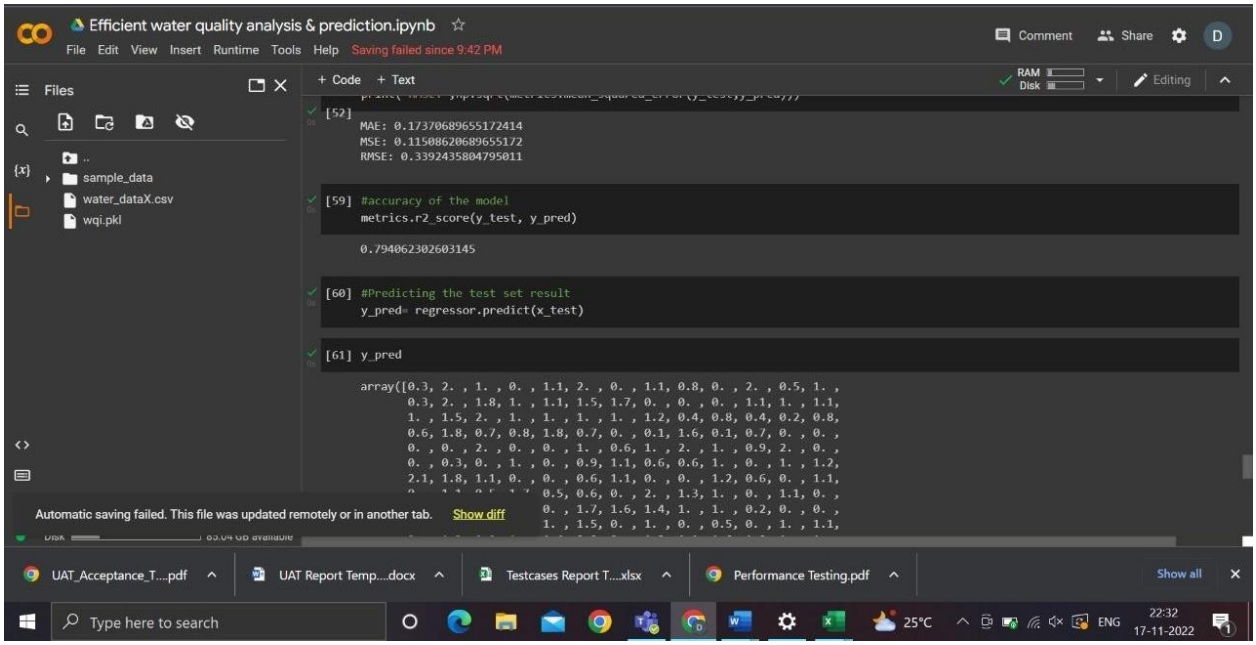
3. Test Case Analysis

This report shows the number of test cases that have passed, failed, and untested

Verify user is able to input valid inputs	5	0	0	5
Verify user is able to predit output for valid inputs	5	0	0	5
Verify user is able to get ouput with InValid input	5	0	0	5
Verify user is able to get ouput without any input	5	0	0	5

9.RESULT

9.1 PERFORMANCE METRICS



Testing : 25%
Training : 75%

10. ADVANTAGES

There are several reasons why you should do frequent water quality monitoring, whether for groundwater, surface water, or open water. If you want to provide a firm basis for a more comprehensive water management strategy, investing in water quality testing should be your first step. This testing will also help you to comply with tight permit restrictions and Australian legislation.
You may find out where your water might need some assistance by assessing the health of your water. In the end, you will be able to save money over time by identifying the cause of the pollution or by continuing to be proactive with your monitoring. Each body of water has distinct features that can only be determined by testing, so making assumptions and purchasing things based on a hunch or a general trend is not a good idea.
Another significant benefit of water quality testing is measuring the quantity of dissolved oxygen in your water since, normally, the less oxygen present, the hotter the water is, and the more hazardous the environment is for aquatic life.

DISADVANTAGES

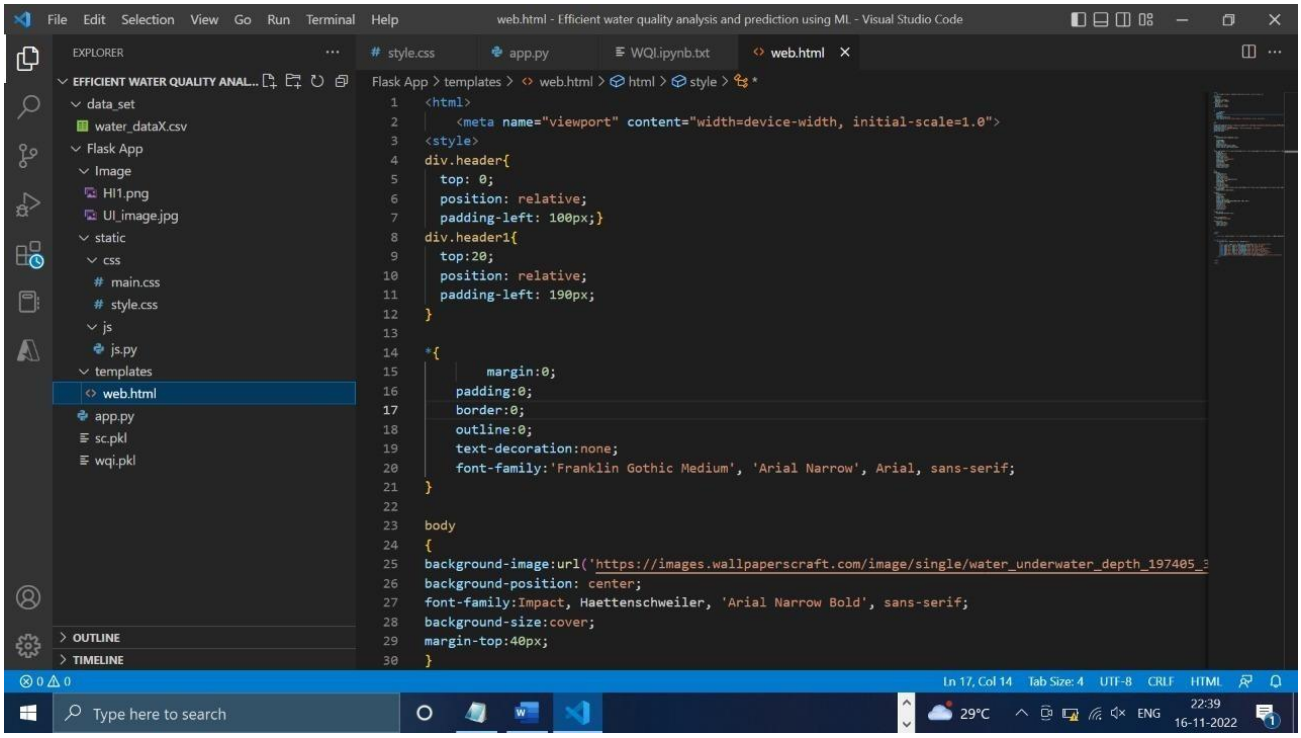
- Challenging to maintain over time and with big data set.
- Data submission requires manual intervention, and some setting is necessary expensive, and typically only practical with Exchange Network funding Requires technical know-how and a network server.
- Human effort is needed to submit data cannot communicate with the Exchange Network by responding to data requests from other nodes. Requires technical know-how and a network server.

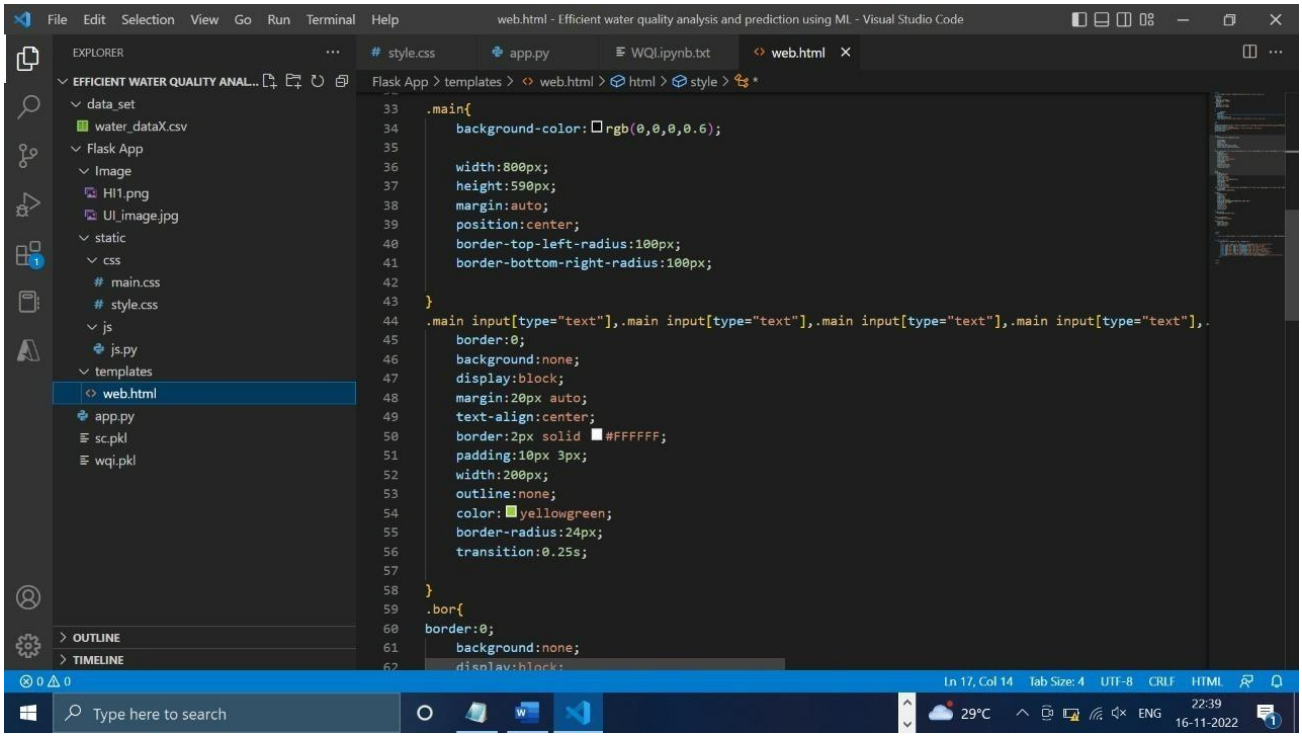
11. CONCLUSION

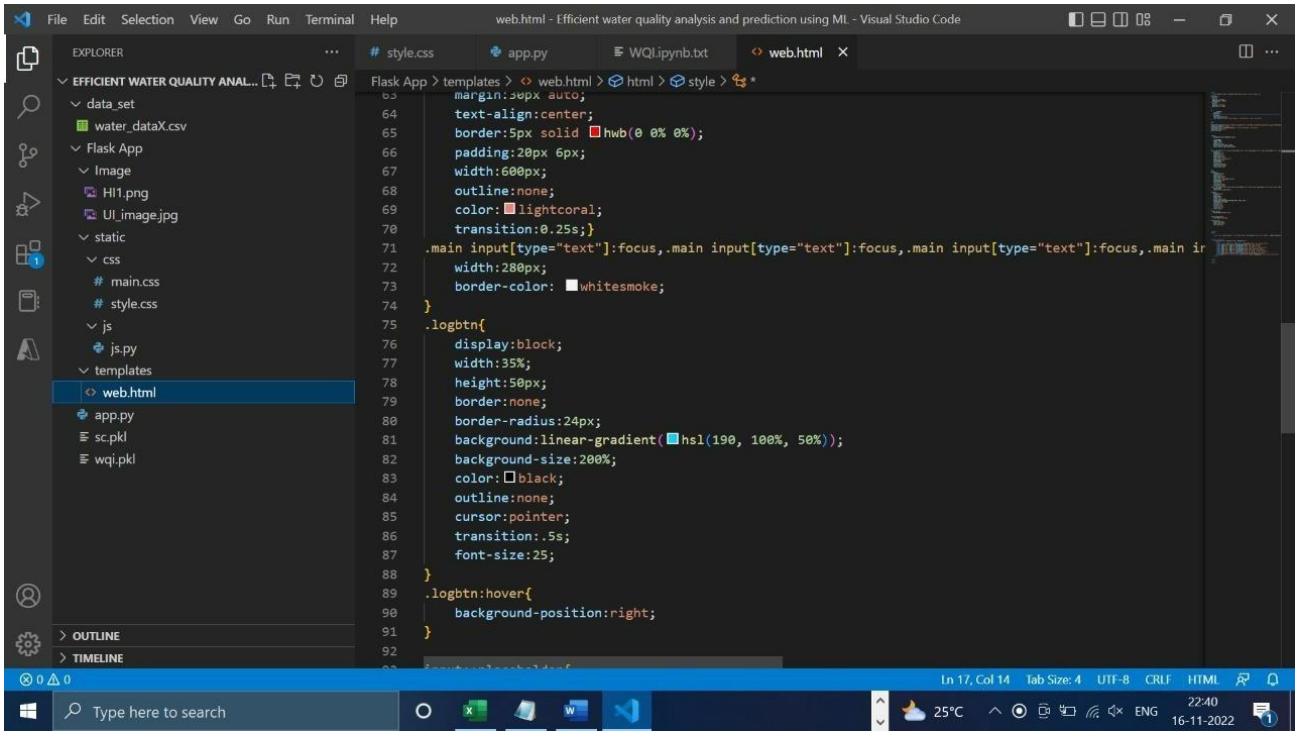
Water is one of the most vital resources for life, and WQI measures its quality. In the past, determining the quality of the water necessitated an expensive and time-consuming lab examination. This work investigated a different machine learning approach for forecasting water quality using just a few basic water quality variables. A group of representative supervised machine learning methods were utilised to estimate. It would identify water of poor quality before it was made available for consumption and alert the necessary authorities. By reducing the number of people who consume water of poor quality, the danger of illnesses like typhoid and diarrhoea should decrease. The use 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.

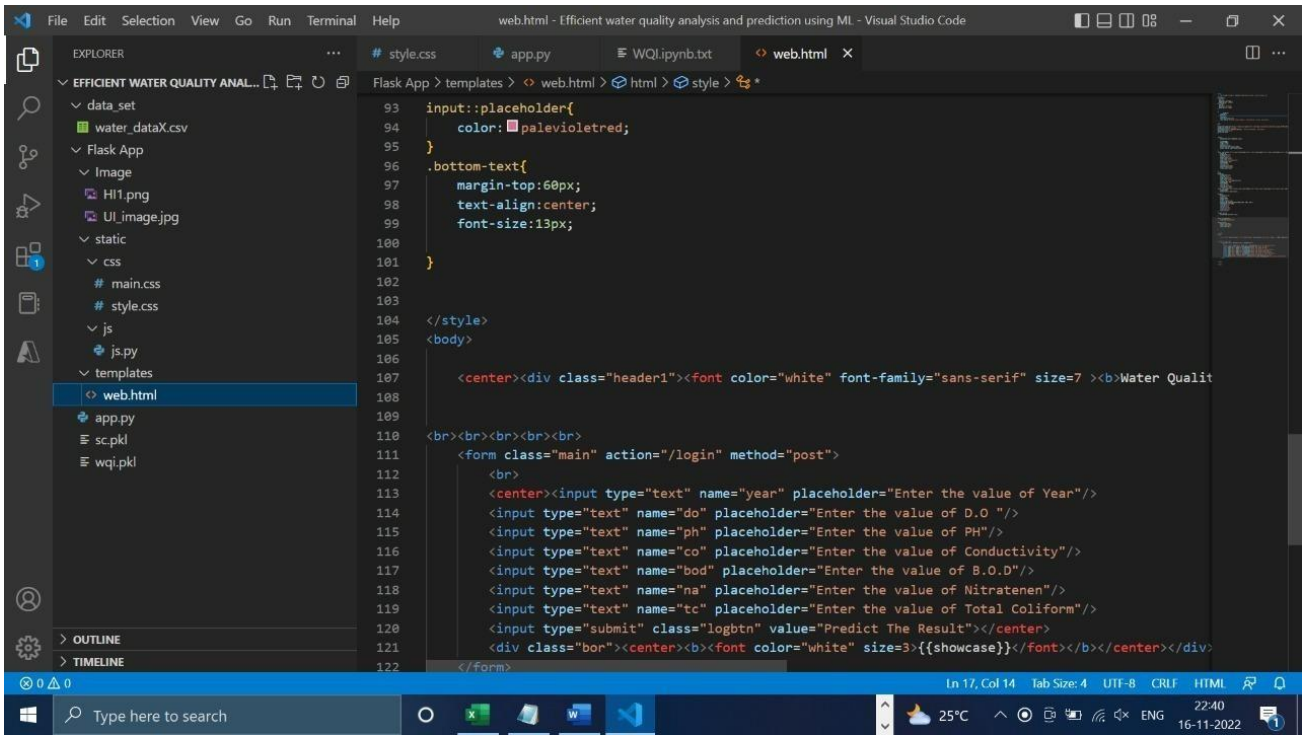
12. APPENDIX

13.SOURCE CODE



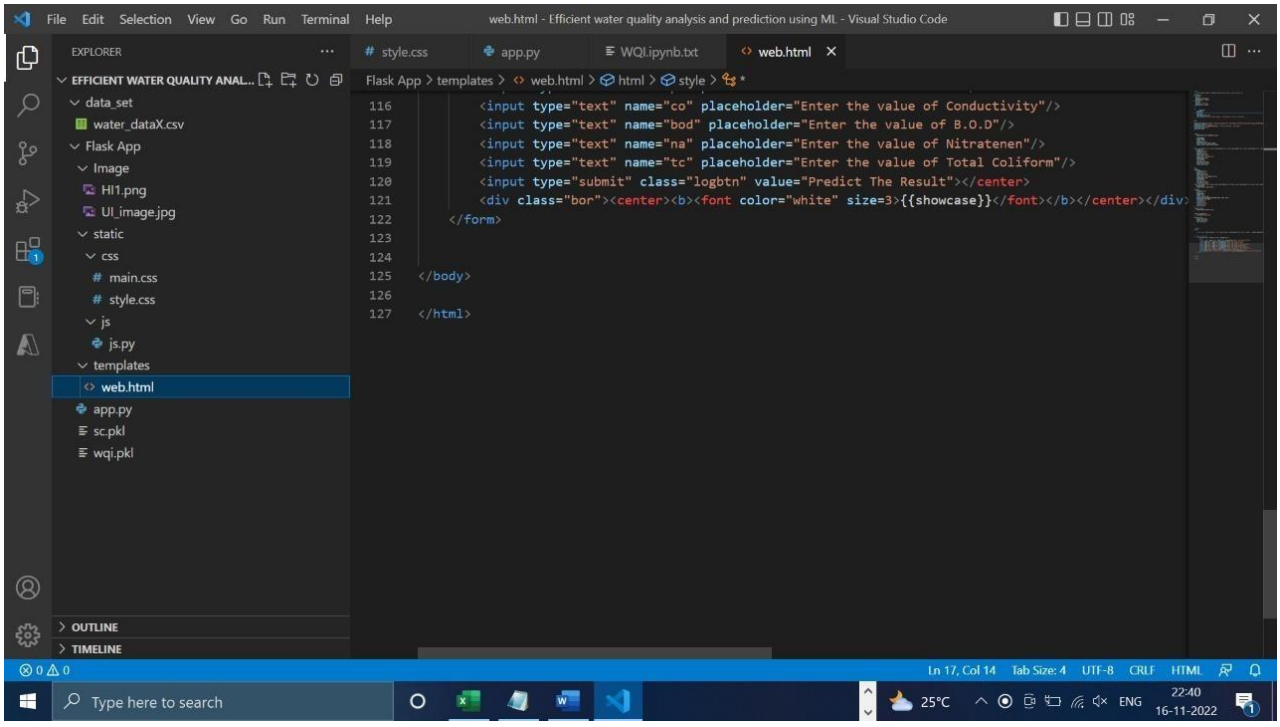


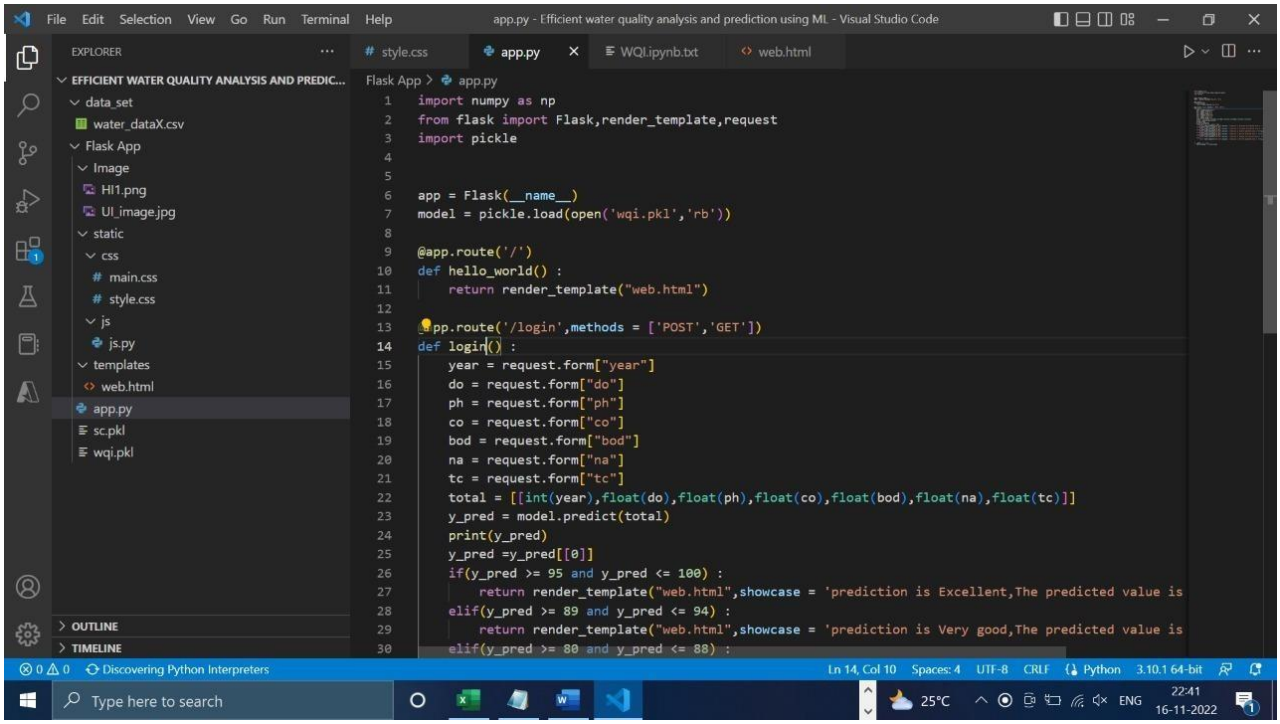

web.html - Efficient water quality analysis and prediction using ML - Visual Studio Code
EFFICIENT WATER QUALITY ANAL...
data_set
water_dataX.csv
Flask App
Image
HI1.png
UI_image.jpg
static
css
main.css
style.css
js
js.py
templates
web.html
app.py
sc.pkl
wqi.pkl
style.css
app.py
WQI.ipynb.txt
web.html
Flask App > templates > web.html > html > style > *
margin:30px auto;
text-align:center;
border:5px solid hwb(0 0% 0%);
padding:20px 6px;
width:600px;
outline:none;
color:lightcoral;
transition:0.25s;}
.main input[type="text"]:focus,.main input[type="text"]:focus,.main input[type="text"]:focus,.main ir
width:280px;
border-color: whitesmoke;
}
.logbtn{
display:block;
width:35%;
height:50px;
border:none;
border-radius:24px;
background:linear-gradient(hsl(190, 100%, 50%));
background-size:200%;
color:black;
outline:none;
cursor:pointer;
transition:.5s;
font-size:25;
}
.logbtn:hover{
background-position:right;
}
OUTLINE
TIMELINE
Ln 17, Col 14 Tab Size: 4 UTF-8 CRLF HTML
Type here to search
25°C
ENG
22:40
16-11-2022

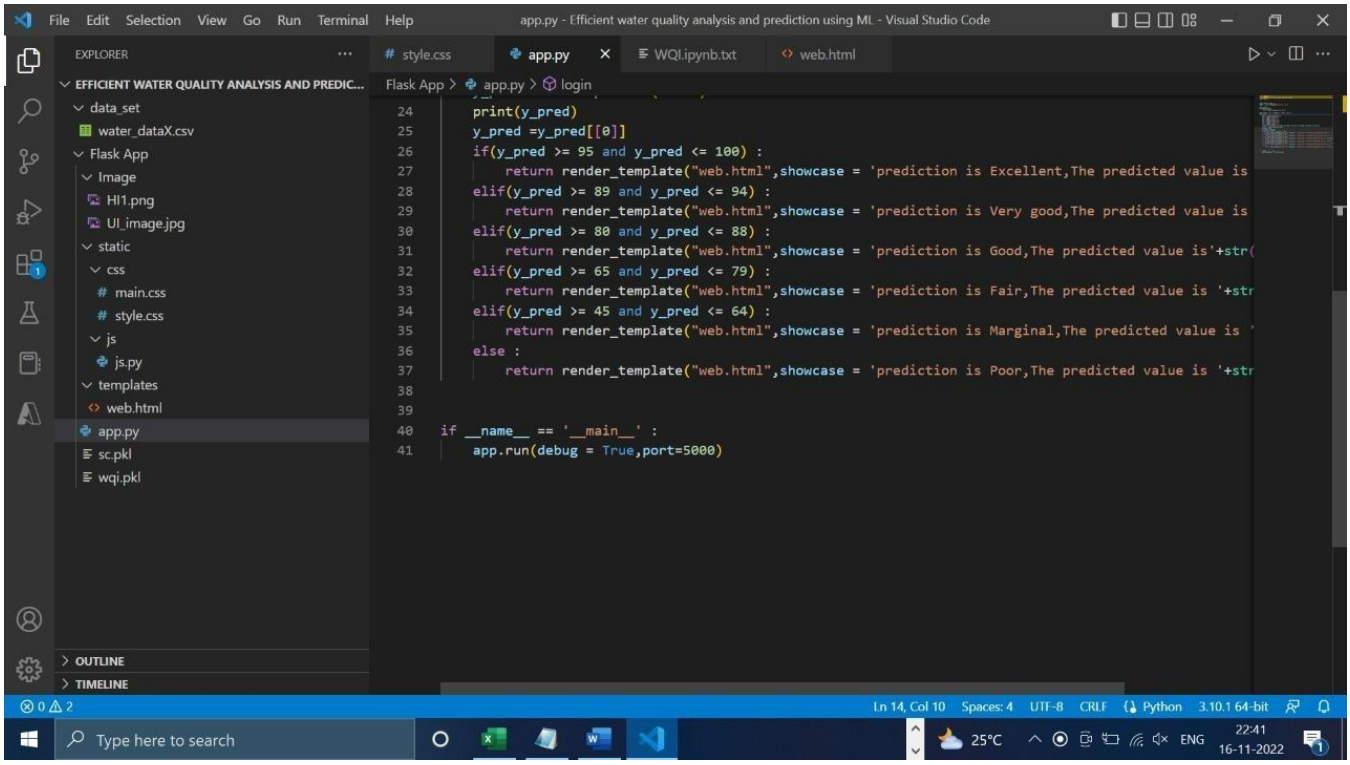

web.html - Efficient water quality analysis and prediction using ML - Visual Studio Code
EFFICIENT WATER QUALITY ANAL...
data_set
water_dataX.csv
Flask App
Image
HI1.png
UI_image.jpg
static
css
main.css
style.css
js
js.py
templates
web.html
app.py
sc.pkl
wqi.pkl
style.css
app.py
WQI.ipynb.txt
web.html
Flask App > templates > web.html > html > style > *
input::placeholder{
color:palevioletred;
}
.bottom-text{
margin-top:60px;
text-align:center;
font-size:13px;
}
</style>
<body>
Water Qualit

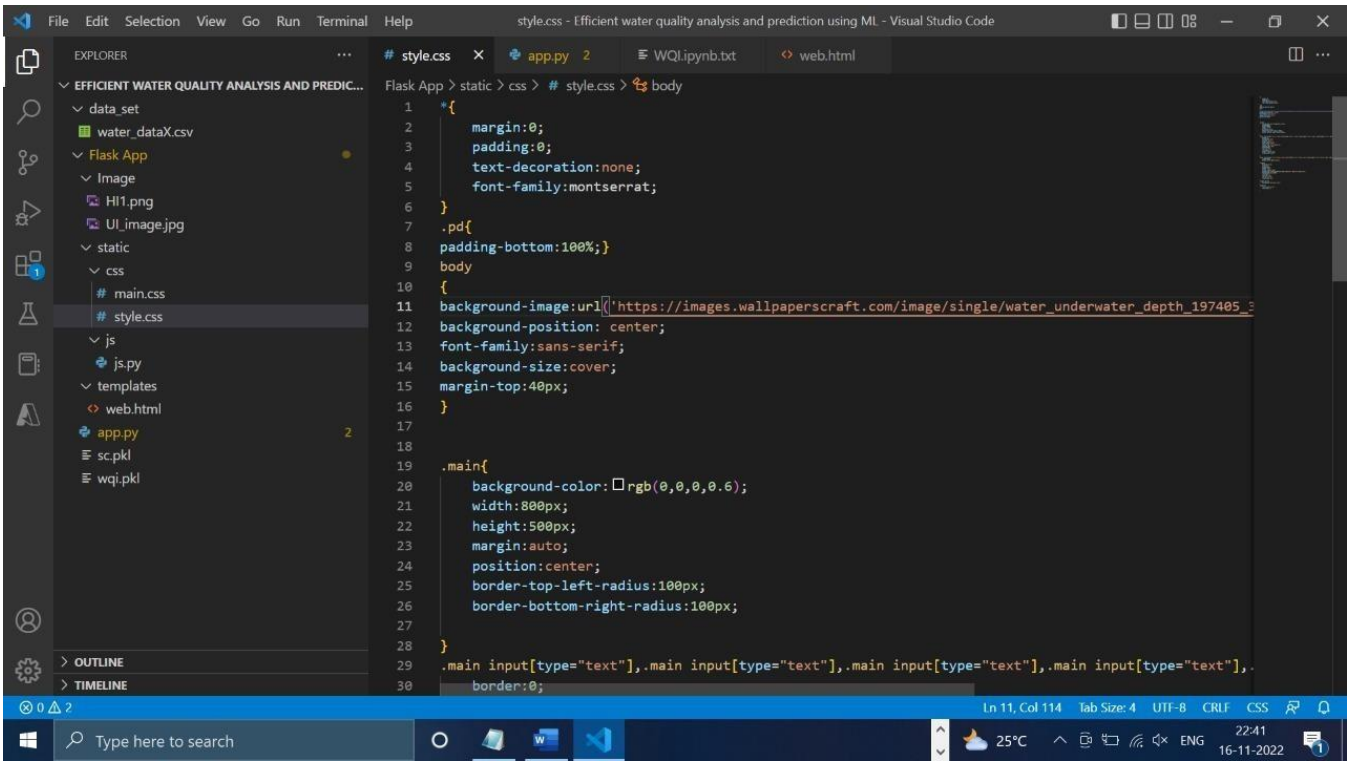
<form class="main" action="/login" method="post">

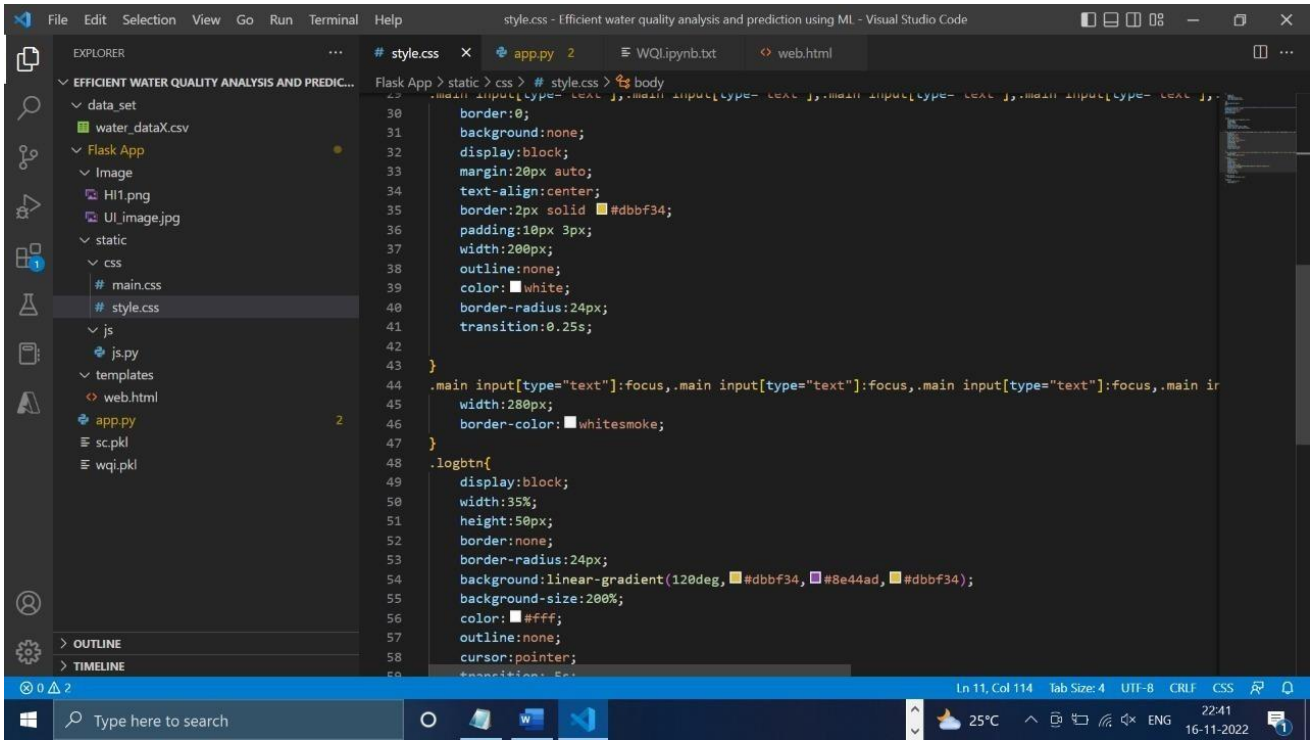
<input type="text" name="year" placeholder="Enter the value of Year"/>
<input type="text" name="do" placeholder="Enter the value of D.O "/>
<input type="text" name="ph" placeholder="Enter the value of PH"/>
<input type="text" name="co" placeholder="Enter the value of Conductivity"/>
<input type="text" name="bod" placeholder="Enter the value of B.O.D"/>
<input type="text" name="na" placeholder="Enter the value of Nitratenen"/>
<input type="text" name="tc" placeholder="Enter the value of Total Coliform"/>
<input type="submit" class="logbtn" value="Predict The Result">
{{showcase}}
</form>
OUTLINE
TIMELINE
Ln 17, Col 14 Tab Size: 4 UTF-8 CRLF HTML
Type here to search
25°C
ENG
22:40
16-11-2022

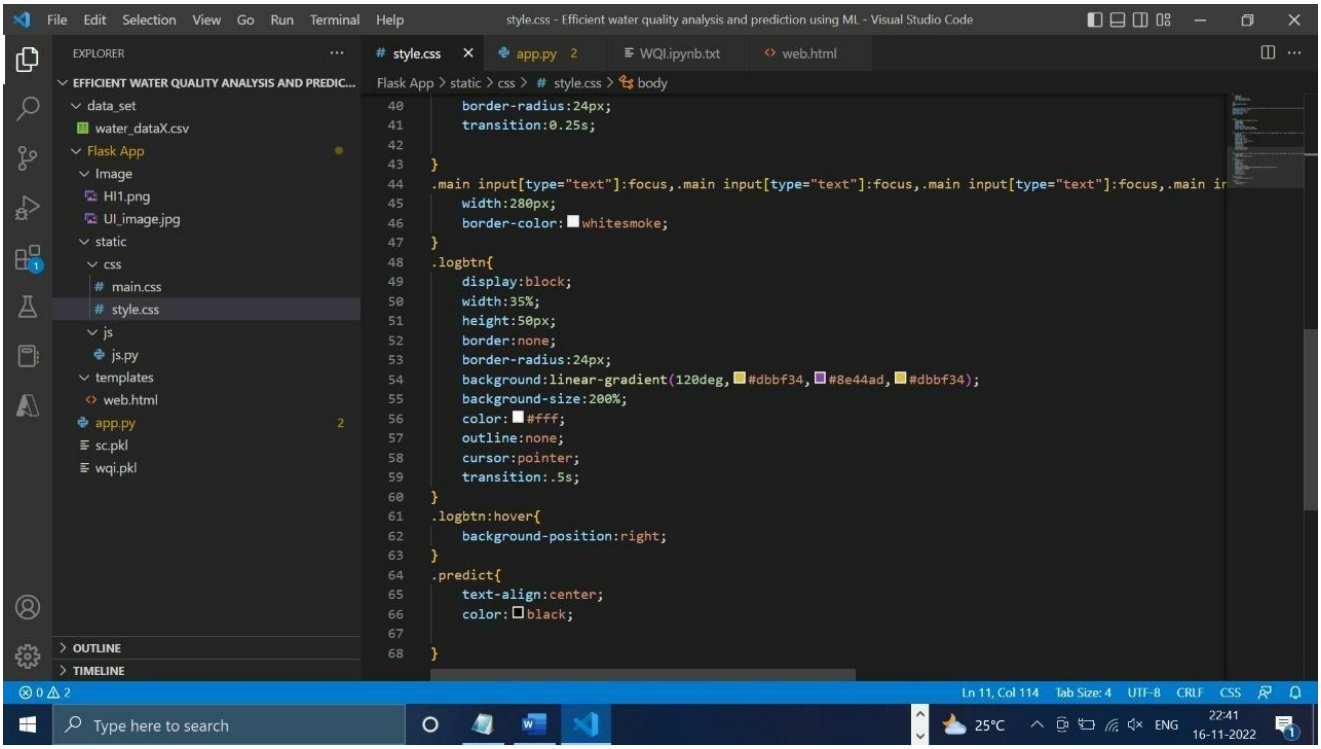

web.html - Efficient water quality analysis and prediction using ML - Visual Studio Code
EFFICIENT WATER QUALITY ANAL...
data_set
water_dataX.csv
Flask App
Image
HI1.png
UI_image.jpg
static
css
main.css
style.css
js
js.py
templates
web.html
app.py
sc.pkl
wqi.pkl
style.css
app.py
WQI.ipynb.txt
web.html
Flask App > templates > web.html > html > style > *
<input type="text" name="co" placeholder="Enter the value of Conductivity"/>
<input type="text" name="bod" placeholder="Enter the value of B.O.D"/>
<input type="text" name="na" placeholder="Enter the value of Nitratenen"/>
<input type="text" name="tc" placeholder="Enter the value of Total Coliform"/>
<input type="submit" class="logbtn" value="Predict The Result">
{{showcase}}
</form>
</body>
</html>
OUTLINE
TIMELINE
Ln 17, Col 14 Tab Size: 4 UTF-8 CRLF HTML
Type here to search
25°C ENG 22:40 16-11-2022


app.py - Efficient water quality analysis and prediction using ML - Visual Studio Code
EFFICIENT WATER QUALITY ANALYSIS AND PREDIC...
data_set
water_dataX.csv
Flask App
Image
HI1.png
UI_image.jpg
static
css
main.css
style.css
js
js.py
templates
web.html
app.py
sc.pkl
wqi.pkl
style.css
app.py
WQI.ipynb.txt
web.html
Flask App > app.py
import numpy as np
from flask import Flask,render_template,request
import pickle
app = Flask(__name__)
model = pickle.load(open('wqi.pkl','rb'))
@app.route('/')
def hello_world() :
return render_template("web.html")
@app.route('/login',methods = ['POST','GET'])
def login() :
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"]
total = [[int(year),float(do),float(ph),float(co),float(bod),float(na),float(tc)]]
y_pred = model.predict(total)
print(y_pred)
y_pred =y_pred[[0]]
if(y_pred >= 95 and y_pred <= 100) :
return render_template("web.html",showcase = 'prediction is Excellent,The predicted value is
elif(y_pred >= 89 and y_pred <= 94) :
return render_template("web.html",showcase = 'prediction is Very good,The predicted value is
elif(y_pred >= 80 and y_pred <= 88) :
OUTLINE
TIMELINE
Discovering Python Interpreters
Ln 14, Col 10 Spaces: 4 UTF-8 CRLF Python 3.10.1 64-bit
Type here to search
25°C ENG 22:41 16-11-2022


app.py - Efficient water quality analysis and prediction using ML - Visual Studio Code
EFFICIENT WATER QUALITY ANALYSIS AND PREDIC...
data_set
water_dataX.csv
Flask App
Image
HI1.png
UI_image.jpg
static
css
main.css
style.css
js
js.py
templates
web.html
app.py
sc.pkl
wqi.pkl
print(y_pred)
y_pred =y_pred[[0]]
if(y_pred >= 95 and y_pred <= 100) :
return render_template("web.html",showcase = 'prediction is Excellent,The predicted value is
elif(y_pred >= 89 and y_pred <= 94) :
return render_template("web.html",showcase = 'prediction is Very good,The predicted value is
elif(y_pred >= 80 and y_pred <= 88) :
return render_template("web.html",showcase = 'prediction is Good,The predicted value is'+str(
elif(y_pred >= 65 and y_pred <= 79) :
return render_template("web.html",showcase = 'prediction is Fair,The predicted value is '+str
elif(y_pred >= 45 and y_pred <= 64) :
return render_template("web.html",showcase = 'prediction is Marginal,The predicted value is
else :
return render_template("web.html",showcase = 'prediction is Poor,The predicted value is '+str
if __name__ == '__main__' :
app.run(debug = True,port=5000)


style.css - Efficient water quality analysis and prediction using ML - Visual Studio Code
*{
margin:0;
padding:0;
text-decoration:none;
font-family:montserrat;
}
.pd{
padding-bottom:100%;}
body
{
background-image:url('https://images.wallpaperscraft.com/image/single/water_underwater_depth_197405_
background-position: center;
font-family:sans-serif;
background-size:cover;
margin-top:40px;
}
.main{
background-color:rgb(0,0,0,0.6);
width:800px;
height:500px;
margin:auto;
position:center;
border-top-left-radius:100px;
border-bottom-right-radius:100px;
}
.main input[type="text"],.main input[type="text"],.main input[type="text"],.main input[type="text"],
border:0;


File Edit Selection View Go Run Terminal Help
style.css - Efficient water quality analysis and prediction using ML - Visual Studio Code
EXPLORER
EFFICIENT WATER QUALITY ANALYSIS AND PREDIC...
data_set
water_dataX.csv
Flask App
Image
HI1.png
UI_image.jpg
static
css
main.css
style.css
js
js.py
templates
web.html
app.py 2
sc.pkl
wqi.pkl
OUTLINE
TIMELINE
style.css
app.py 2
WQI.ipynb.txt
web.html
Flask App > static > css > style.css > body
30 border:0;
31 background:none;
32 display:block;
33 margin:20px auto;
34 text-align:center;
35 border:2px solid #dbbf34;
36 padding:10px 3px;
37 width:200px;
38 outline:none;
39 color:white;
40 border-radius:24px;
41 transition:0.25s;
42
43 }
44 .main input[type="text"]:focus,.main input[type="text"]:focus,.main input[type="text"]:focus,.main ir
45 width:280px;
46 border-color:whitesmoke;
47 }
48 .logbtn{
49 display:block;
50 width:35%;
51 height:50px;
52 border:none;
53 border-radius:24px;
54 background:linear-gradient(120deg,#dbbf34,#8e44ad,#dbbf34);
55 background-size:200%;
56 color:#fff;
57 outline:none;
58 cursor:pointer;
Ln 11, Col 114 Tab Size: 4 UTF-8 CRLF CSS
Type here to search
25°C ENG 22:41 16-11-2022


File Edit Selection View Go Run Terminal Help
style.css - Efficient water quality analysis and prediction using ML - Visual Studio Code
EXPLORER
EFFICIENT WATER QUALITY ANALYSIS AND PREDIC...
data_set
water_dataX.csv
Flask App
Image
HI1.png
UI_image.jpg
static
css
main.css
style.css
js
js.py
templates
web.html
app.py 2
sc.pkl
wqi.pkl
OUTLINE
TIMELINE
style.css
app.py 2
WQI.ipynb.txt
web.html
Flask App > static > css > style.css > body
40 border-radius:24px;
41 transition:0.25s;
42
43 }
44 .main input[type="text"]:focus,.main input[type="text"]:focus,.main input[type="text"]:focus,.main ir
45 width:280px;
46 border-color:whitesmoke;
47 }
48 .logbtn{
49 display:block;
50 width:35%;
51 height:50px;
52 border:none;
53 border-radius:24px;
54 background:linear-gradient(120deg,#dbbf34,#8e44ad,#dbbf34);
55 background-size:200%;
56 color:#fff;
57 outline:none;
58 cursor:pointer;
59 transition:.5s;
60 }
61 .logbtn:hover{
62 background-position:right;
63 }
64 .predict{
65 text-align:center;
66 color:black;
67
68 }
Ln 11, Col 114 Tab Size: 4 UTF-8 CRLF CSS
Type here to search
25°C ENG 22:41 16-11-2022

REQUIREMENT.TXT

Flask == 2.2.2

numpy == 1.23.4

pandas == 1.5.1

scikit-learn == 1.1.3

matplotlib == 3.6.2

seaborn == 0.12.1

flask-cors == 3.0.10

GitHub Link:
https://github.com/IBM-EPBL/IBM-Project-30687-1660154042

Demo Video Link:
https://drive.google.com/file/d/1vTB835Q13MMSPcpeMjrj10FDFu9YjUk3/view?usp=sharing