

PROJECT BASED EXPERIENTIAL LEARNING PROGRAM (NALAIYATHIRAN)

EFFICIENT WATER QUALITY ANALYSIS AND PREDICTION USING MACHINE LEARNING

A PROJECT REPORT

Submitted by

Shalima.M.S.(961419104040)

Abisha.J.(961419104003)

Godslin Sharmi(961419104024)

Sherlin Femina(961419104042)

Berlin Jose(961419104018)

TEAM ID: PNT2022TMID51670

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

**MAR EPHRAEM COLLEGE OF
ENGINEERING & TECHNOLOGY
ELAVUVILAI**



NOVEMBER 2022

TABLE OF CONTENTS

S.no	Title	Pg no
1.	Introduction	4
1.a	Project overview	4
1.b	purpose	4
2.	Literature survey	5
2.1	Existing problems	9
2.2	Reference	9
2.3	Problem statement and definition	11
3.	Ideation & proposed solution	11
3.1	Empathy Map Canvas	11
3.2	Ideation & Brainstorming	13
3.3	Proposed Solution	16
3.4	Problem Solution fit	18
4.	Requirement analysis	19
4.1	Functional requirement	19
4.2	Non-Functional requirements	20
5.	Project design	21
5.1	Data Flow Diagrams	21
5.2	Solution & Technical Architecture	22
5.3	User Stories	23
6.	Project planning & scheduling	26
6.1	Sprint Planning & Estimation	26
6.2	Sprint Delivery Schedule	27
7.	Coding & solutioning (Explain the features added	28

	in the project along with code)	
7.1	Feature 1	28
7.2	Feature 2	28
8.	Testing	29
8.1	Test Cases	29
8.	User Acceptance Testing	30
9.	Results	31
9.1.	Performance Metrics	31
10.	Advantages & disadvantages	32
11.	Conclusion	33
12.	Appendix	34
13.	Source Code	34

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 as arsenic, chromium, lead, mercury and cadmium.

LITERATURE SURVERY

[1] Shafi et al[1]estimated water quality using classical machine learning algorithms namely, Support Vector Machines (SVM), Neural Networks (NN), Deep Neural Networks (Deep NN) and K Nearest Neighbours (KNN), with the highest accuracy of 93% with Deep NN. The estimated water quality in their work is based on only three parameters: turbidity, temperature and pH, which are tested according to World Health Organization (WHO) standards (Available online at URL <https://www.who.int/airpollution/guidelines/en/>). Using only three parameters and comparing them to standardized values is quite a limitation when predicting water quality.

[2] In this research paper, the advanced artificial intelligence algorithms, namely, NARNET and LSTM were used to predict the WQI. Also, machine learning algorithms such as SVM, KNN AND Naïve Bayes were used to classify the WQI data.

Disadvantage: The efficiency and robustness is very poor in predicting the WQI.

[3] In this paper, ten indicator parameters like Solids, chloramines, sulphate, conductivity, organic carbon, trihalomethanes, turbidity, pH value, hardness and potability were used to predict the water quality. To estimate, a set of representative supervised machine learning algorithms like Decision tree, K-Nearest Neighbour were used.

[4] This paper presents a novel spatio-temporal multi-view multi-task learning framework to forecast the water quality of a station by fusing multiple sources of urban data. It also uses Spatio- temporal view alignment to work toward local information aggregation for each station and global

Prediction alignment, which incorporates the spatial correlations among stations and performs co- prediction over all stations using these correlations.

Disadvantage: Problem will arise in water quality inference in the urban water distributed systems.

[5] This study's objective is to create a water quality prediction model utilising Artificial Neural Networks (ANN) and time-series analysis to incorporate water quality parameters. Historical data on water quality are used in this study. Mean-Squared Error (MSE), Root Mean Squared Error (RMSE), and Regression Analysis are the performance evaluation metrics used to gauge how well the model is doing. ANN has received widespread recognition as a tool for classifying complicated information, including those pertaining to environmental dynamics. It can effectively explain the non-linear relationship between the intricate water quality statistics.

[6] The proposed framework has settled to an Intelligent constant IoT based Water Quality Monitoring framework which depends on Machine to Machine correspondence through AI. Turbidity and the conductivity sensors are connected. The conductivity acts as a sensor gateway. The sensor input are sent to the pi4, an edge level processor(personal computer) where in the K Means, a machine learning algorithm is used for predicting the quality of water. The predicted water quality data are stored in Cloud server for future access. The predicted data is sent to the water controller unit for further action.

References

- [1] Surface Water Pollution Detection using Internet of Things – Shafi

- [2] Water Quality Prediction Using Artificial Intelligence Algorithms – Theyazn H.H Aldhyani

- [3] Water Quality Prediction using Machine Learning – Sai Sreeja Kurra, Sambangi Geethika Naidu, Sravani Chowdala, Sree Chithra Yellanki, Dr.B.Esther Sunanda

- [4] Urban Water Quality Prediction Based on Multi-task Multi-view Learning – Ye Liu, Yu Zheng, Yuxuan Liang, Shuming Liu, David S. Rosenblum

- [5] Machine learning methods for better water quality prediction – Ali Najah Ahmeda
- [6] A Real Time Water Quality Monitoring Using Machine Learning Algorithm – S. Angel Vergina, Dr.S.Kayalvizhi ,Dr. R.M. Bhavadharini, Kalpana Devi. S

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

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

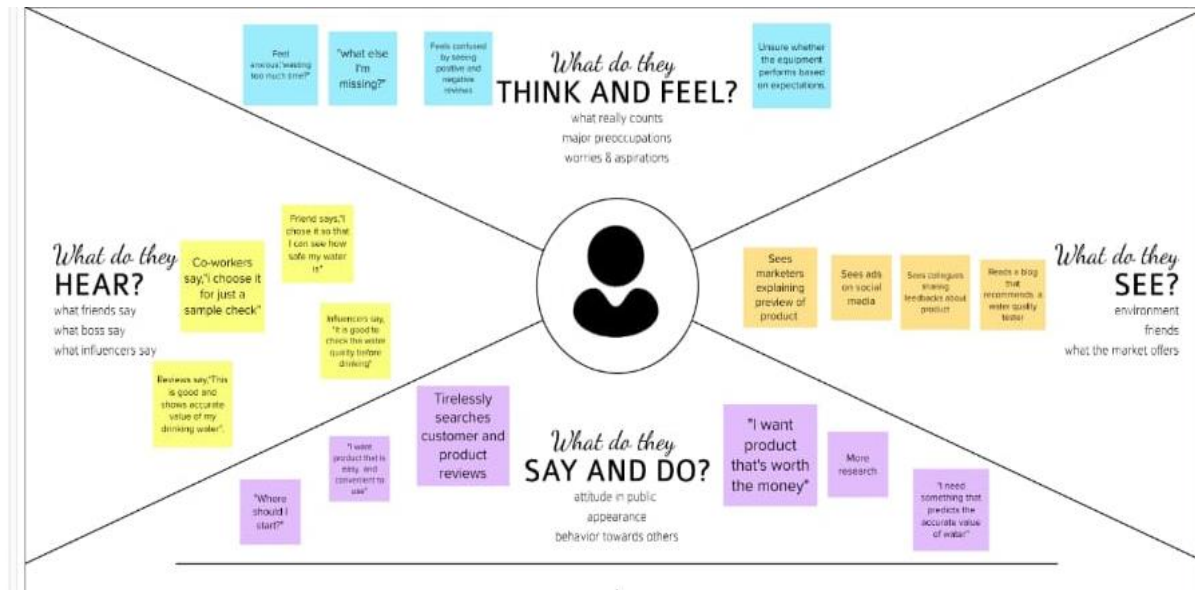
Empathy Map Canvas:

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



3.2 IDEATION AND BRAINSTORMING

2

Brainstorm
Write down any ideas that come to mind that address your problem statement.
[10 minutes](#)

Tip

When your workshop is ending, make sure you have a chance to discuss the ideas you've generated and how to move forward.

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Stimulus

Step-2: Brain Storm, Idea Listing, Grouping

3

Group ideas
Take turns sharing your ideas while clustering similar or related notes as you go.
In the last 10 minutes, give each cluster a sentence-like label. If a cluster is bigger than six sticky notes, try and use it to break it up into smaller sub-groups.

20 minutes

Cluster each sticky note into a cluster for each quality factor

Using sticky notes, brainstorm ideas

20%
Give sticky notes a label to make them easier to read. In the last 10 minutes, give each cluster a sentence-like label. If a cluster is bigger than six sticky notes, try and use it to break it up into smaller sub-groups.

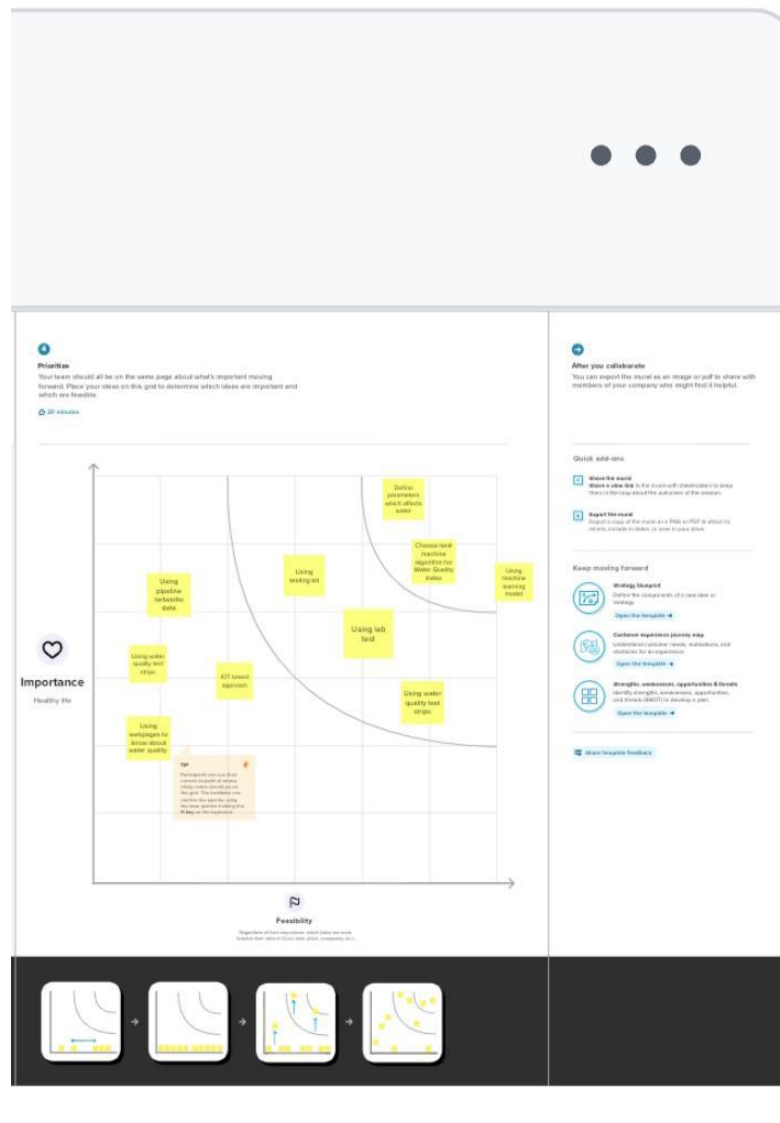
Using sticky notes, brainstorm ideas

Using sticky notes, brainstorm ideas

Using sticky notes, brainstorm ideas

Using sticky notes, brainstorm ideas

Step -3: Idea Prioritization



3.3 PROPOSED SOLUTION

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	Due to the fast growing urbanization supply of safe drinking water is a challenge for the every urban areas. They get water from sources like their own well, from municipal through pipes, pumps etc. Many of them use simple water purifier to make the water safe to drink. But it does not purifies the entire water. There will be small chemical particles the may cause diseases. So we need an water quality predicting system that will predict the correct quality of water.
2.	Idea / Solution description	To make our water safe to drink, we are developing a model that will predict the waterquality accurately using water quality standardindicators like pH, Turbidity, Alkalinity, Nitrate, Hardness etc.
3.	Novelty / Uniqueness	Predicting accurate value.Better that tester kit.
4.	Social Impact / Customer Satisfaction	Customer satisfaction is an important goal. To meet this goal ,it is necessary to use an evaluation model for measuring the customer satisfaction level. Some important criteria such as, water quality,responsibility of municipal etc. are distinguished and used in the proposed model.To integrate all of these criteria in a unit index, the Analytic Hierarchy Process (AHP) technique is used.
5.	Business Model (Revenue Model)	Water is one of the essential component of human living.Checking the water is safe to drink or not can help us to know whether we can drink or not.Avoiding drinking unsafe water can reduce water-borne diseases like typhoid,diarrhoea etc.
6.	Scalability of the Solution	We can use reverse osmosis method. ReverseOsmosis works by using a high pressure pump to increase the pressure on the salt side of the RO and force the water across the semi-permeable RO membrane, leaving almost all (around 95% to 99%) of dissolvedsalts behind in the reject stream. The amountof pressure required depends on the salt concentration of the feed water. The more concentrated the feed water, the more pressure is required to overcome the osmotic pressure.

3.4 PROBLEM SOLUTION FIT

Project Title: Efficient Water Analysis And Prediction using Machine Learning

Project Design Phase-I – Solution fit Template

Team ID: PNT2022TMI51670

Define CS, fit into CC	1. CUSTOMER SEGMENT(S) <ul style="list-style-type: none">Who need chemical free waterWho need clean, odourless water	6. CUSTOMER CONSTRAINTS <ul style="list-style-type: none">Whether the solution is in budget or not.Whether it give the expected result.	5. AVAILABLE SOLUTIONS <ul style="list-style-type: none">Have lab to test the quality of water.Pros-Give the correct resultCons-Takes more time to get the result	Explore AS, different
	2. JOBS-TO-BE-DONE / PROBLEMS <ul style="list-style-type: none">Recommend water testerRecommend to dig wellRecommend water purifier	9. PROBLEM ROOT CAUSE <ul style="list-style-type: none">Without checking the quality of water health issues may arise.Lab analysis may take long time to give the result	7. BEHAVIOUR <ul style="list-style-type: none">Finding right tester and testing the quality of waterDigging deeply to know more	

3. TRIGGERS <ul style="list-style-type: none">Seeing their children not well.Knowing new tools to solve the quality.	10. YOUR SOLUTION <ul style="list-style-type: none">Having lab test to check the safe of water takes so much time even days.Using the machine learning model, can know the result suddenly.	8. CHANNELS of BEHAVIOUR 8.1 ONLINE <ul style="list-style-type: none">Browse on internet. 8.2 OFFLINE <ul style="list-style-type: none">Communicate with local print publications
4. EMOTIONS: BEFORE / AFTER <ul style="list-style-type: none">Feel nervous not knowing the answer to a problem and after getting answer to the problem feel excited.Feel confused of selecting tools to meet the problem and after getting the correct tool feel satisfied.		

4.REQUIREMENT ANALYSIS

4.1 FUNCTIONAL REQUIREMENT

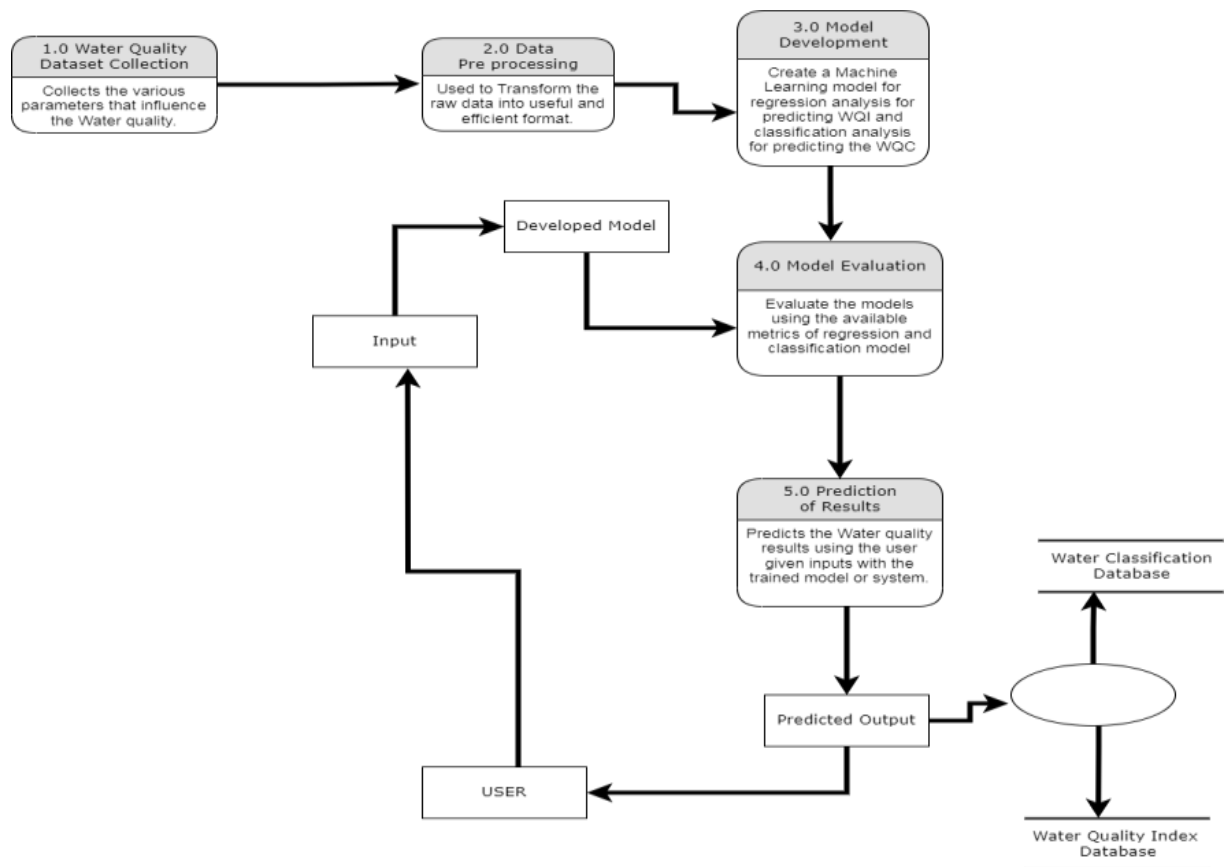
Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Form Registration through Gmail Registration through LinkedIn
FR-2	User Confirmation	Confirmation via Email Confirmation via OTP
FR-3	User Authentication	Authentication with OTP via SMS/email Authentication via social networks
FR-4	User Authorization	Role-Based Access Controls(RBAC) OpenID Authorization
FR-5	External Interfaces	Interaction logic between user and software Software interfaces like frontend, backend, etc.Buttons, functions on the model
FR-6	Reporting	SMS notification for reports/alerts Email notification for reports/alerts

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.

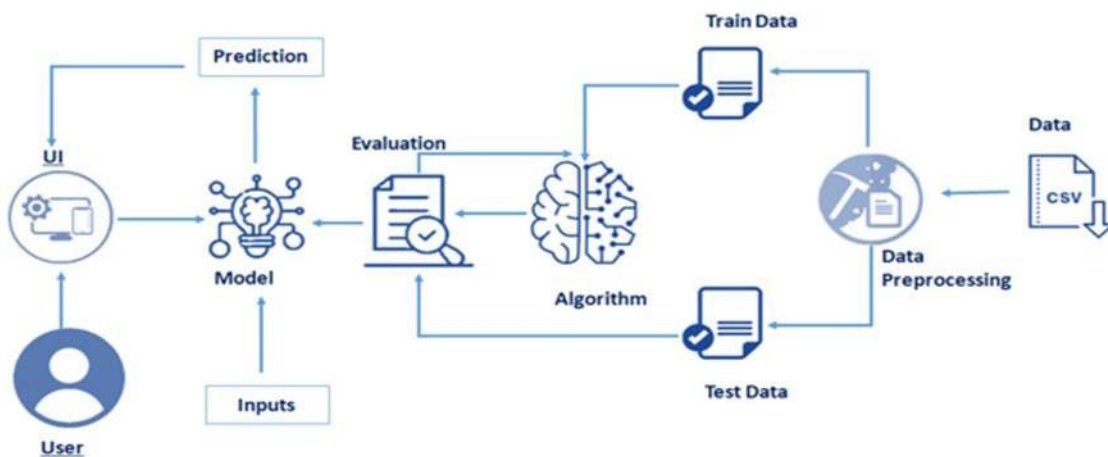


DFD LEVEL 1 DIAGRAM

5.2 SOLUTION AND TECHNICAL 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 (Epic)	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 Gmail	I can register & access the dashboard with Gmail login	Medium	Sprint-1
	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 log into the dashboard by entering username & password	I can access the website	High	Sprint-1
Customer (Webuser)	Analysis of water quality	USN-7	As a user, I can access the water quality prediction section.	I can get the water quality	High	Sprint-1
Customer Care Executive	Customer queries	USN-8	As a customer care Executive, I can check the customer queries they posted in the website.	I can improve the customer satisfaction.	High	Sprint-1
Administrator	Maintaining website	USN-9	As an administrator, I can maintain website and enhance the online presence.	I can improve the website's appearance & usability.	High	Sprint-2
		USN-10	As an administrator, I can maintain issues in analysing values.	I can improve the accuracy of predicting values.	High	Sprint-2
		USN-11	As an administrator, I can update the website content	I can ensure the content is in harmony with the customer's overall objectives	Medium	Sprint-2
		USN-12	As an administrator, I can improve the website.	I can enhance user experience	High	Sprint-2

6 PROJECT PLANNING AND SCDULING

6.1SPRINT PLANNING AND ESTIMATION

Sprint	Function al Require ment (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Registrati on	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	2	High	Shalima M S
Sprint-1		USN-2	As a user, I will receive confirmation email oncel have registered for the application	1	High	Berlin Jose J
Sprint-2		USN-3	As a user, I can register for the application through Facebook	2	Low	Sherlin Femina B
Sprint-1		USN-4	As a user, I can register for the application through Gmail	2	Medium	Godslin Sharmi R
Sprint-1	Login	USN-5	As a user, I can log into the application by entering email & password	1	High	Abisha J
Sprint-2	Dashboard	USN-6	As a user, I can log into the dashboard by entering username & password	2	High	Shalima M S
Sprint-3	Analysis of water quality	USN-7	As a user, I can access the water quality prediction section	3	High	Berlin Jose J
Sprint-3	Customer queries	USN-8	As a customer care Executive, I can check the customer queries they posted in the website.	2	High	Godslin Sharmi R
Sprint-4	Maintai ning website	USN-9	As an administrator, I can maintain website and enhance the online presence.	1	Medium	Sherlin Femina B

	Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
	Sprint-4		USN-10	As an administrator, I can maintain issues in analysing values.	1	Medium	Shalima M S
	Sprint-3		USN-11	As an administrator, I can update the website content	2	High	Abisha J
	Sprint-4		USN-12	As an administrator, I can improve the website	2	High	Berlin JoseJ

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	6 Days	24 Oct 2022	29 Oct 2022	20	29 Oct 2022
Sprint-2	20	6 Days	31 Oct 2022	05 Nov 2022	20	05 Nov 2022
Sprint-3	20	6 Days	07 Nov 2022	12 Nov 2022	20	12 Nov 2022
Sprint-4	20	6 Days	14 Nov 2022	19 Nov 2022	20	19 Nov 2022

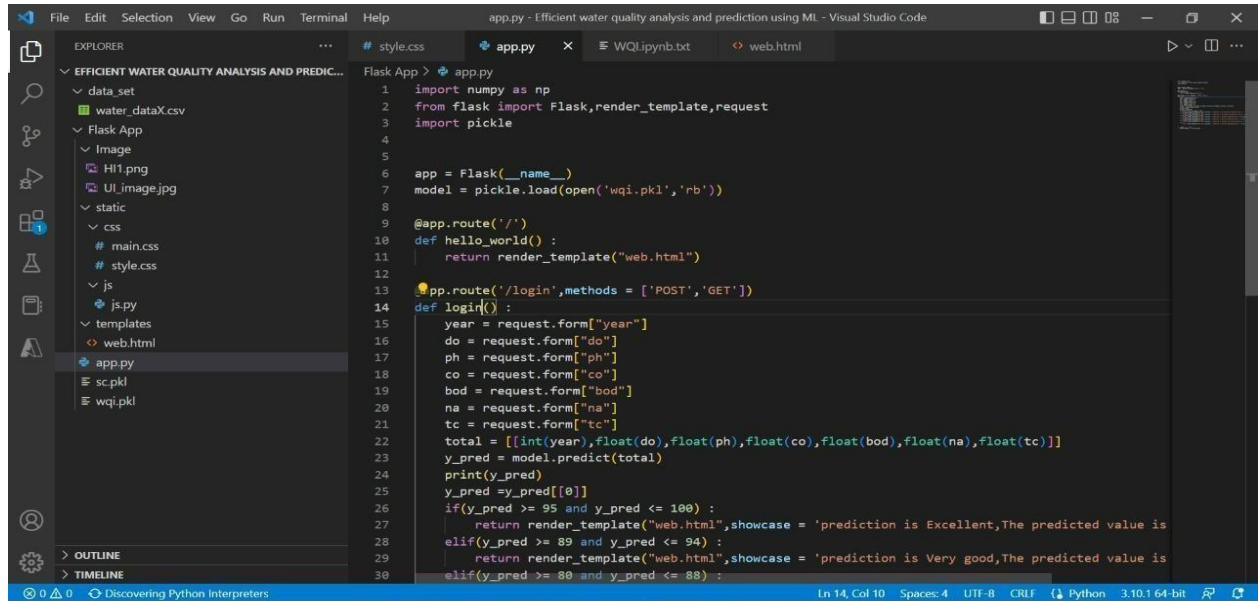
Burndown Chart:



7 CODING AND SOLUTIONS

7.2 FEATURE 1

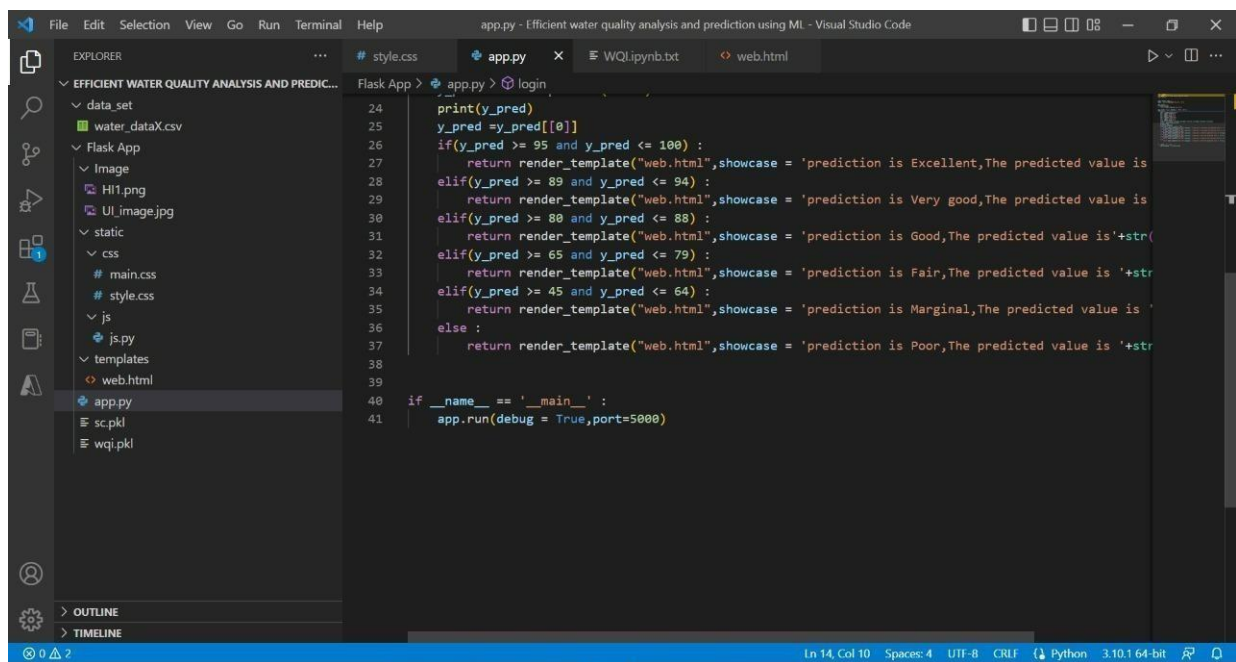
Flask is a web framework. This means flask provides you with tools, libraries and technologies that allow you to build a web application. This web application can be some web pages, a blog, a wiki or go as big as a web-based calendar application or a commercial website.



```
File Edit Selection View Go Run Terminal Help
app.py - 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
        H11.png
        UI_image.jpg
      static
        css
          main.css
          style.css
        js
          js.py
      templates
        web.html
      app.py
      sc.pkl
      wqi.pkl

Flask App > app.py
1 import numpy as np
2 from flask import Flask,render_template,request
3 import pickle
4
5
6 app = Flask(__name__)
7 model = pickle.load(open('wqi.pkl','rb'))
8
9
10 @app.route('/')
11 def hello_world():
12     return render_template("web.html")
13
14 @app.route('/login',methods = ['POST','GET'])
15 def login():
16     year = request.form["year"]
17     do = request.form["do"]
18     ph = request.form["ph"]
19     co = request.form["co"]
20     bod = request.form["bod"]
21     na = request.form["na"]
22     tc = request.form["tc"]
23     total = [[int(year),float(do),float(ph),float(co),float(bod),float(na),float(tc)]]
24     y_pred = model.predict(total)
25     print(y_pred)
26     y_pred = y_pred[0]
27     if(y_pred >= 95 and y_pred <= 100):
28         return render_template("web.html",showcase = 'prediction is Excellent,The predicted value is
29     elif(y_pred >= 89 and y_pred <= 94):
30         return render_template("web.html",showcase = 'prediction is Very good,The predicted value is
31     elif(y_pred >= 80 and y_pred <= 88):
32         return render_template("web.html",showcase = 'prediction is Good,The predicted value is'+str(
33     elif(y_pred >= 65 and y_pred <= 79):
34         return render_template("web.html",showcase = 'prediction is Fair,The predicted value is '+str
35     elif(y_pred >= 45 and y_pred <= 64):
36         return render_template("web.html",showcase = 'prediction is Marginal,The predicted value is '
37     else:
38         return render_template("web.html",showcase = 'prediction is Poor,The predicted value is '+str
39
40 if __name__ == '__main__':
41     app.run(debug = True,port=5000)
```



```
File Edit Selection View Go Run Terminal Help
app.py - 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
        H11.png
        UI_image.jpg
      static
        css
          main.css
          style.css
        js
          js.py
      templates
        web.html
      app.py
      sc.pkl
      wqi.pkl

Flask App > app.py > login
24 print(y_pred)
25 y_pred = y_pred[0]
26 if(y_pred >= 95 and y_pred <= 100):
27     return render_template("web.html",showcase = 'prediction is Excellent,The predicted value is
28 elif(y_pred >= 89 and y_pred <= 94):
29     return render_template("web.html",showcase = 'prediction is Very good,The predicted value is
30 elif(y_pred >= 80 and y_pred <= 88):
31     return render_template("web.html",showcase = 'prediction is Good,The predicted value is'+str(
32 elif(y_pred >= 65 and y_pred <= 79):
33     return render_template("web.html",showcase = 'prediction is Fair,The predicted value is '+str
34 elif(y_pred >= 45 and y_pred <= 64):
35     return render_template("web.html",showcase = 'prediction is Marginal,The predicted value is '
36 else:
37     return render_template("web.html",showcase = 'prediction is Poor,The predicted value is '+str
38
39
40 if __name__ == '__main__':
41     app.run(debug = True,port=5000)
```

8 TESTING

8.2 TEST CASE

The screenshot shows a web browser window with the address bar displaying 'localhost:8888/view/Downloads/nalaiyathiran/nalaiyathiran/predict.html'. The page title is 'Water Quality Predictor'. The form contains the following inputs:

Input Field	Value
Year	2014
Temperature (°C)	6.7
PH	7.5
Dissolved Oxygen (mg/L)	203
Hardness (mg/L)	0
Chloride (mg/L)	0.1
Sulfate (mg/L)	27

Below the inputs is a 'Submit' button. The output area, highlighted with a yellow border, displays the result: 'Prediction is Good, the prediction value is 80.308'.

8.3 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

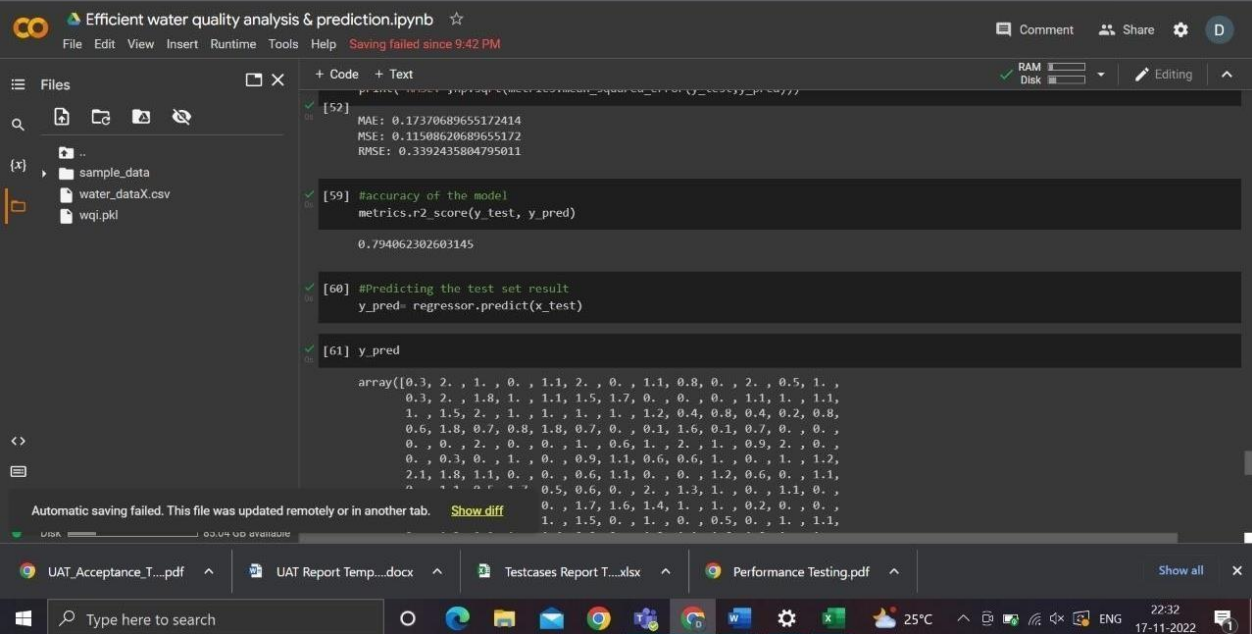
3. Test Case Analysis

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 predict output for valid inputs	5	0	0	5
Verify user is able to get output with Invalid input	5	0	0	5
Verify user is able to get output without any input	5	0	0	5

9.RESULT

9.1 PERFORMANCE METRICS



The screenshot displays a Jupyter Notebook titled "Efficient water quality analysis & prediction.ipynb". The interface includes a file explorer on the left showing a directory with files like "sample_data", "water_dataX.csv", and "wqi.pkl". The main area shows a code cell with the following output:

```
[52] MAE: 0.17370689655172414  
MSE: 0.11508620689655172  
RMSE: 0.3392435804795011
```

```
[59] #accuracy of the model  
metrics.r2_score(y_test, y_pred)  
  
0.794062302603145
```

```
[60] #Predicting the test set result  
y_pred= regressor.predict(x_test)
```

```
[61] y_pred  
  
array([0.3, 2., 1., 0., 1.1, 2., 0., 1.1, 0.8, 0., 2., 0.5, 1.,  
0.3, 2., 1.8, 1., 1.1, 1.5, 1.7, 0., 0., 0., 1.1, 1., 1.1,  
1., 1.5, 2., 1., 1., 1., 1., 1.2, 0.4, 0.8, 0.4, 0.2, 0.8,  
0.6, 1.8, 0.7, 0.8, 1.8, 0.7, 0., 0.1, 1.6, 0.1, 0.7, 0., 0.,  
0., 0., 2., 0., 0., 1., 0.6, 1., 2., 1., 0.9, 2., 0.,  
0., 0.3, 0., 1., 0., 0.9, 1.1, 0.6, 0.6, 1., 0., 1., 1.2,  
2.1, 1.8, 1.1, 0., 0., 0.6, 1.1, 0., 0., 1.2, 0.6, 0., 1.1,  
0., 0., 0., 0., 0.5, 0.6, 0., 2., 1.3, 1., 0., 1.1, 0.,  
0., 1.7, 1.6, 1.4, 1., 1., 0.2, 0., 0.,  
1., 1.5, 0., 1., 0., 0.5, 0., 1., 1.1,
```

Below the code cell, a message states: "Automatic saving failed. This file was updated remotely or in another tab. [Show diff](#)". The bottom of the image shows a Windows taskbar with various application icons and a system tray displaying the date and time as 22:32 on 17-11-2022.

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

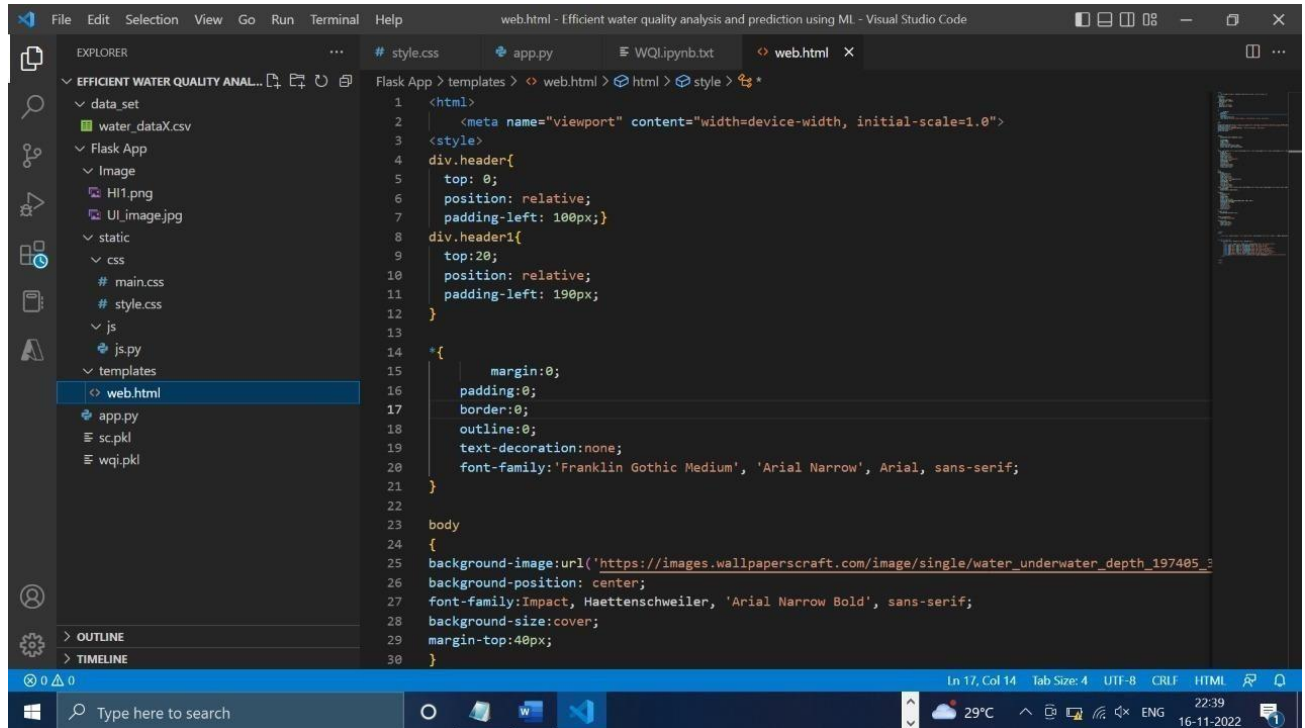
- Challenging to maintain over time and with big data set.
- Data submission requires manual intervention, and some setting is necessary expensive, and typically only practical with Exchange Network funding Requires technical know-how and a network server.
- Human effort is needed to submit data cannot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

11. CONCLUSION

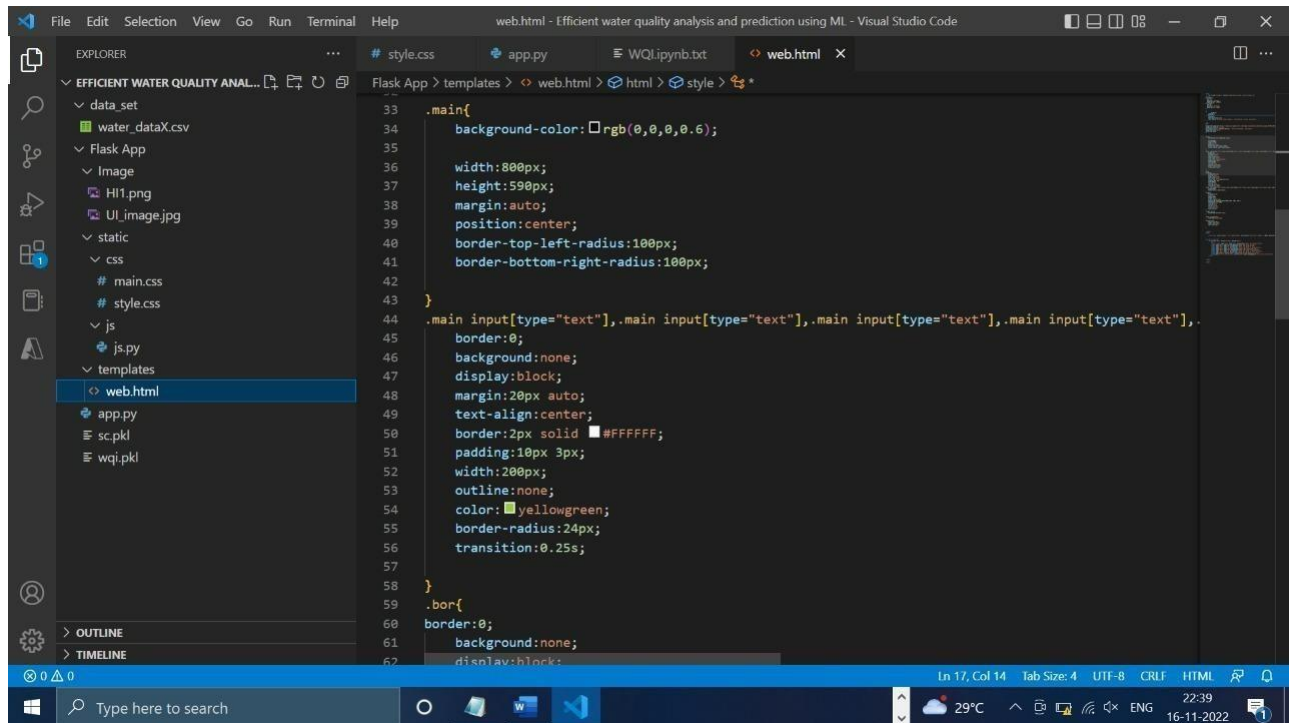
Water is one of the most vital resources for life, and WQI measures its quality. In the past, determining the quality of the water necessitated an expensive and time-consuming lab examination. This work investigated a different machine learning approach for forecasting water quality using just a few 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

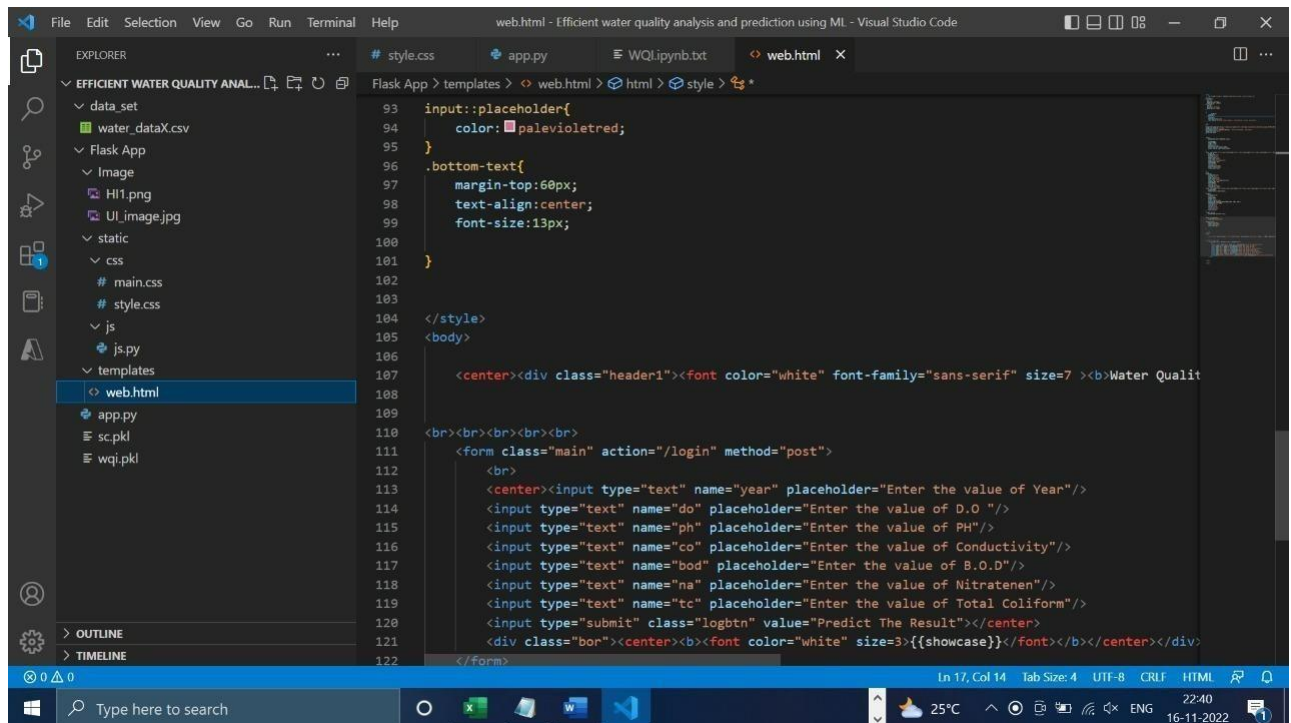
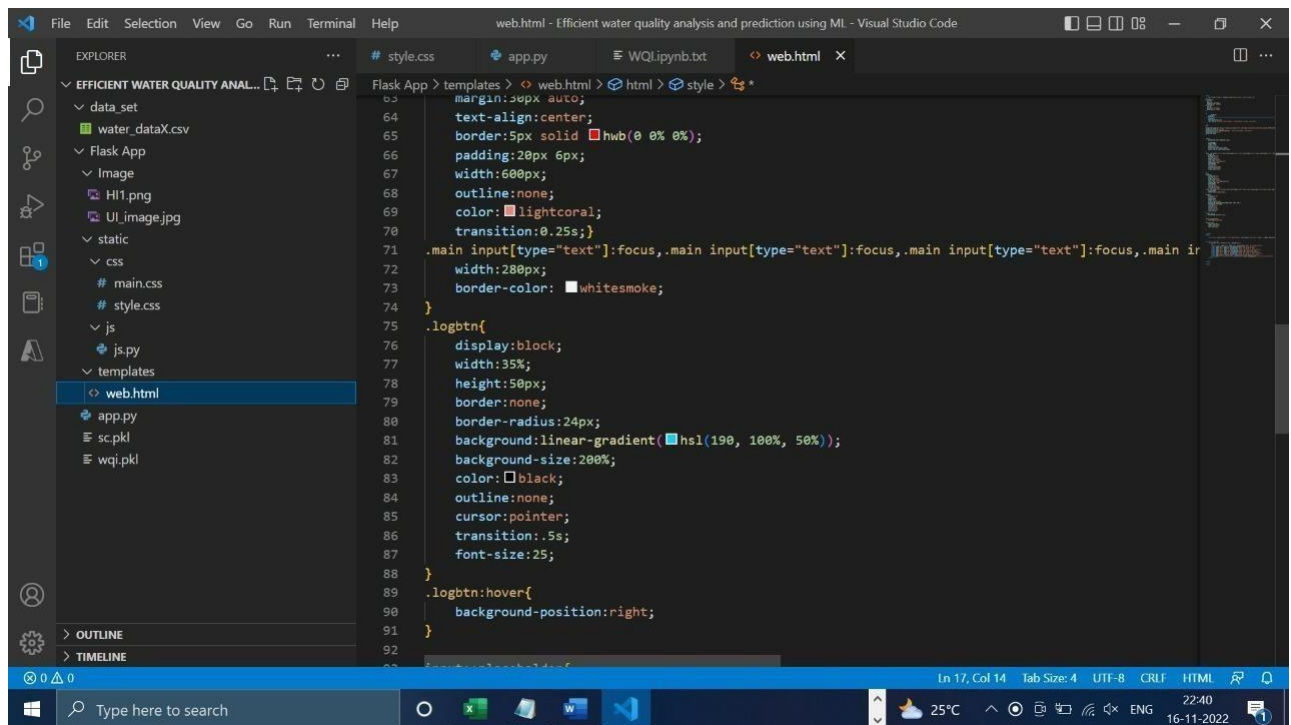
13.SOURCE CODE



```
1 <html>
2   <meta name="viewport" content="width=device-width, initial-scale=1.0">
3   <style>
4     div.header{
5       top: 0;
6       position: relative;
7       padding-left: 100px;}
8     div.header1{
9       top:20;
10      position: relative;
11      padding-left: 190px;
12    }
13  }
14  *{
15    margin:0;
16    padding:0;
17    border:0;
18    outline:0;
19    text-decoration:none;
20    font-family:'Franklin Gothic Medium', 'Arial Narrow', Arial, sans-serif;
21  }
22
23  body
24  {
25    background-image:url('https://images.wallpaperscraft.com/image/single/water_underwater_depth_197405_3');
26    background-position: center;
27    font-family:Impact, Haettenschweiler, 'Arial Narrow Bold', sans-serif;
28    background-size:cover;
29    margin-top:40px;
30  }
```



```
33 .main{
34   background-color:rgb(0,0,0,0.6);
35
36   width:800px;
37   height:590px;
38   margin:auto;
39   position:center;
40   border-top-left-radius:100px;
41   border-bottom-right-radius:100px;
42 }
43
44 .main input[type="text"],.main input[type="text"],.main input[type="text"],.main input[type="text"],.
45   border:0;
46   background:none;
47   display:block;
48   margin:20px auto;
49   text-align:center;
50   border:2px solid #FFFFFF;
51   padding:10px 3px;
52   width:200px;
53   outline:none;
54   color:yellowgreen;
55   border-radius:24px;
56   transition:0.25s;
57 }
58
59 .bor{
60   border:0;
61   background:none;
62   display:block;
```



The screenshot shows the Visual Studio Code editor with the 'web.html' file open. The Explorer sidebar on the left shows the project structure: 'EFFICIENT WATER QUALITY ANALYSIS AND PREDICTION' with subfolders 'data_set', 'Flask App', 'Image', 'static', 'css', 'js', and 'templates'. The 'web.html' file is selected under 'templates'. The main editor displays the HTML code for the web form, which includes input fields for 'co' (Conductivity), 'bod' (B.O.D), 'na' (Nitratene), and 'tc' (Total Coliform), and a 'Predict The Result' button. The status bar at the bottom indicates 'Ln 17, Col 14'.

```
116 <input type="text" name="co" placeholder="Enter the value of Conductivity"/>
117 <input type="text" name="bod" placeholder="Enter the value of B.O.D"/>
118 <input type="text" name="na" placeholder="Enter the value of Nitratene"/>
119 <input type="text" name="tc" placeholder="Enter the value of Total Coliform"/>
120 <input type="submit" class="logbtn" value="Predict The Result"/></center>
121 <div class="bor"><center><b><font color="white" size=3>{{showcase}}</font></b></center></div>
122 </form>
123
124
125 </body>
126
127 </html>
```

The screenshot shows the Visual Studio Code editor with the 'app.py' file open. The Explorer sidebar on the left shows the project structure, with 'app.py' selected under 'Flask App'. The main editor displays the Python code for the Flask application, which includes imports for 'numpy', 'Flask', 'render_template', 'request', and 'pickle'. The code defines a Flask app, loads a pickle model, and implements routes for 'hello_world' and 'login'. The 'login' route processes form data and uses the model to predict the result, which is then rendered in the 'web.html' template. The status bar at the bottom indicates 'Ln 14, Col 10'.

```
1 import numpy as np
2 from flask import Flask,render_template,request
3 import pickle
4
5
6 app = Flask(__name__)
7 model = pickle.load(open('wqi.pkl','rb'))
8
9 @app.route('/')
10 def hello_world():
11     return render_template("web.html")
12
13
14 @app.route('/login',methods = ['POST','GET'])
15 def login():
16     year = request.form["year"]
17     do = request.form["do"]
18     ph = request.form["ph"]
19     co = request.form["co"]
20     bod = request.form["bod"]
21     na = request.form["na"]
22     tc = request.form["tc"]
23     total = [[int(year),float(do),float(ph),float(co),float(bod),float(na),float(tc)]]
24     y_pred = model.predict(total)
25     print(y_pred)
26     y_pred = y_pred[0]
27     if(y_pred >= 95 and y_pred <= 100):
28         return render_template("web.html",showcase = 'prediction is Excellent,The predicted value is
29     elif(y_pred >= 89 and y_pred <= 94):
30         return render_template("web.html",showcase = 'prediction is Very good,The predicted value is
31     elif(y_pred >= 80 and y_pred <= 88):
```



```
File Edit Selection View Go Run Terminal Help app.py - 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
    sc.pkl
    wqi.pkl

# style.css
app.py x WQIipynb.txt web.html
Flask App > app.py > login
24 print(y_pred)
25 y_pred = y_pred[0]
26 if(y_pred >= 95 and y_pred <= 100) :
27     return render_template("web.html",showcase = 'prediction is Excellent,The predicted value is
28 elif(y_pred >= 89 and y_pred <= 94) :
29     return render_template("web.html",showcase = 'prediction is Very good,The predicted value is
30 elif(y_pred >= 80 and y_pred <= 88) :
31     return render_template("web.html",showcase = 'prediction is Good,The predicted value is'+str(
32 elif(y_pred >= 65 and y_pred <= 79) :
33     return render_template("web.html",showcase = 'prediction is Fair,The predicted value is '+str
34 elif(y_pred >= 45 and y_pred <= 64) :
35     return render_template("web.html",showcase = 'prediction is Marginal,The predicted value is '
36 else :
37     return render_template("web.html",showcase = 'prediction is Poor,The predicted value is '+str
38
39
40 if __name__ == '__main__' :
41     app.run(debug = True,port=5000)

Ln 14, Col 10 Spaces: 4 UTF-8 CRLF Python 3.10.1 64-bit
```

```
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
    sc.pkl
    wqi.pkl

# style.css
app.py 2 WQIipynb.txt web.html
Flask App > static > css > # style.css > body
1 *{
2     margin:0;
3     padding:0;
4     text-decoration:none;
5     font-family:montserrat;
6 }
7 .pd{
8     padding-bottom:100%;}
9 body
10 {
11     background-image:url('https://images.wallpaperscraft.com/image/single/water_underwater_depth_197405_3
12     background-position: center;
13     font-family:sans-serif;
14     background-size:cover;
15     margin-top:40px;
16 }
17
18
19 .main{
20     background-color:rgb(0,0,0,0.6);
21     width:800px;
22     height:500px;
23     margin:auto;
24     position:center;
25     border-top-left-radius:100px;
26     border-bottom-right-radius:100px;
27 }
28
29 .main input[type="text"],.main input[type="text"],.main input[type="text"],.main input[type="text"],.
30     border:0;
```

```
# style.css
29 body
30 {
31     border:0;
32     background:none;
33     display:block;
34     margin:20px auto;
35     text-align:center;
36     border:2px solid #d9d9d9;
37     padding:10px 3px;
38     width:200px;
39     outline:none;
40     color:white;
41     border-radius:24px;
42     transition:0.25s;
43 }
44 .main input[type="text"]:focus,.main input[type="text"]:focus,.main input[type="text"]:focus,.main input[type="text"]:focus
45 {
46     width:280px;
47     border-color:whitesmoke;
48 }
49 .logbtn{
50     display:block;
51     width:35%;
52     height:50px;
53     border:none;
54     border-radius:24px;
55     background:linear-gradient(120deg,#d9d9d9,#8e44ad,#d9d9d9);
56     background-size:200%;
57     color:#fff;
58     outline:none;
59     cursor:pointer;
60     transition:0.5s;
61 }
```

```
# style.css
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 input[type="text"]:focus
45 {
46     width:280px;
47     border-color:whitesmoke;
48 }
49 .logbtn{
50     display:block;
51     width:35%;
52     height:50px;
53     border:none;
54     border-radius:24px;
55     background:linear-gradient(120deg,#d9d9d9,#8e44ad,#d9d9d9);
56     background-size:200%;
57     color:#fff;
58     outline:none;
59     cursor:pointer;
60     transition:0.5s;
61 }
62 .logbtn:hover{
63     background-position:right;
64 }
65 .predict{
66     text-align:center;
67     color:black;
68 }
```


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