Project Report Format

Date	**18.11.2022**
Team id	**PNT2022TMID49528**
Project name	**Efficient water quality analysis and prediction using machine learning**

1. INTRODUCTION

1.1 Project Overview

Water is the most significant resource of life, crucial for supporting the life of most existing creatures and human beings. Living organisms need water with enough quality to continue their lives. Water quality has a direct impact on public health and the environment. In this project we are going to implement a water quality prediction using machine learning techniques. In this technique, our model predicts that the water is safe to drink or not using some parameters like Ph value, conductivity, hardness, etc.

1.2 Purpose

The goal is to predict the spatio - temporal water quality in terms of the power of hydrogen (pH), value for the next day based on the historical data of water measurement indices. This model predicts water quality and is used to indicate whether or not it is suitable for drinking based on some parameters

2. LITERATURE SURVEY

2.1 Existing problem

https://www.mdpi.com/2073-4441/11/11/2210

2.2 References

1. PCRWR. National Water Quality Monitoring Program, Fifth Monitoring Report (2005–2006); Pakistan Council of Research in Water Resources Islamabad: Islamabad, Pakistan, 2007. Available online:http://www.pcrwr.gov.pk/Publications/Water%20Quality%20Reports/Water%20Quality%20Monitoring%20Report%202005-06.pdf (accessed on 23 August 2019).

2. Mehmood, S.; Ahmad, A.; Ahmed, A.; Khalid, N.; Javed, T. Drinking Water Quality in Capital City of Pakistan. Open Access Sci. Rep. 2013, 2. [CrossRef]

3. PCRWR. Water Quality of Filtration Plants, Monitoring Report; PCRWR: Islamabad, Pakistan, 2010. Available online: http://www.pcrwr.gov.pk/Publications/Water%20Quality%20Reports/FILTRTAION%20PLANTS%20REPOT-CDA.pdf (accessed on 23 August 2019).

2.3 Problem Statement Definition

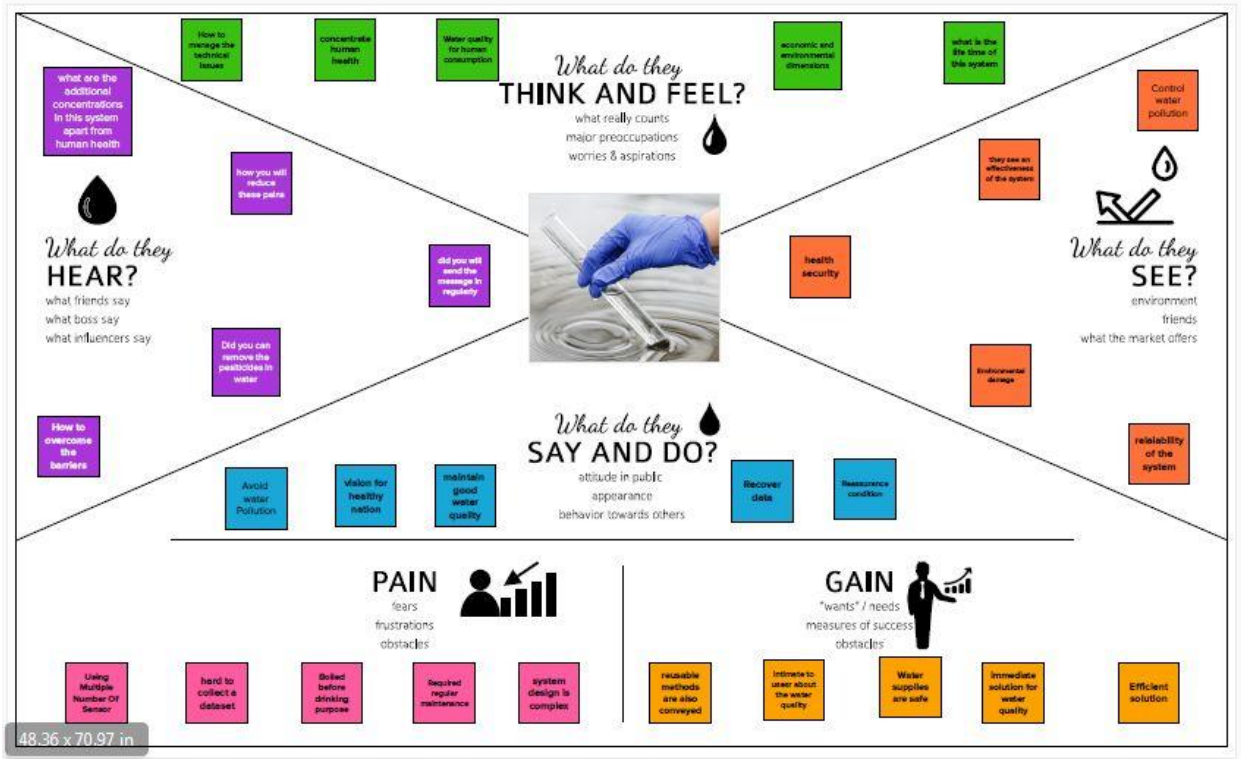
1. Deepak is a/an Officer
   Who needs To monitor the purity level of drinking water
   Because He is responsible for people welfare

2. Banu is a/an counselor of the area
   Who needs To know about the alternate methods to using a impurity water
   Because she will not know about these alternative methods

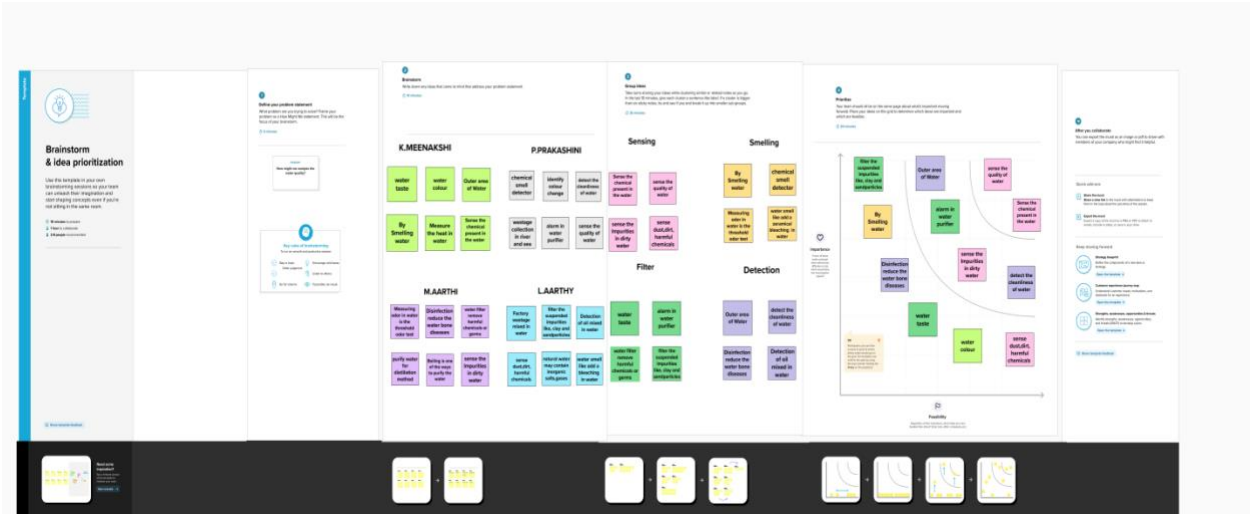
3. Priya is a/an one of person in this person
   Who needs To know about the alternate methods to using a impurity water
   Because she will not know about these alternative methods

3. IDEATION & PROPOSED SOLUTION

3.1 Empathy Map Canvas



3.2 Ideation & Brainstorming

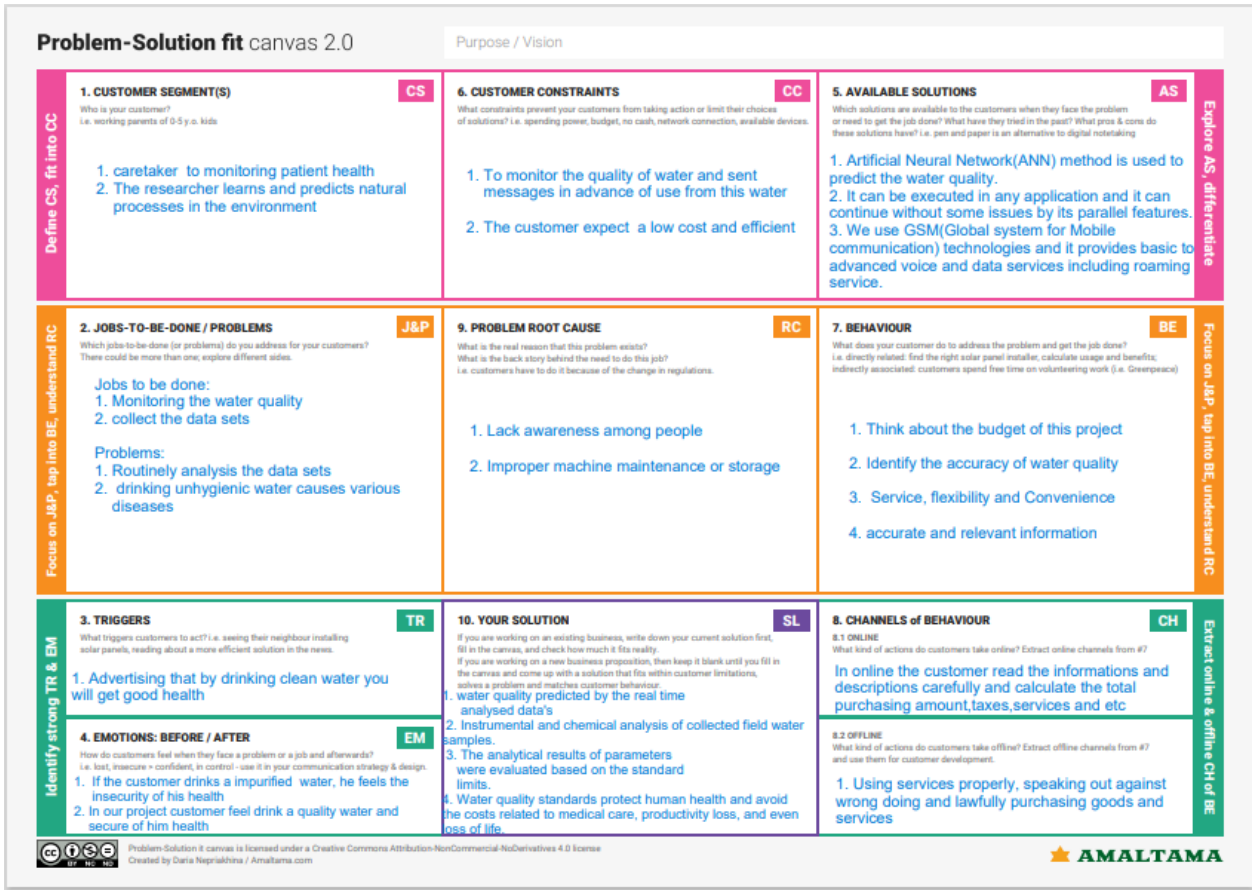


3.3 Proposed Solution

S.NO	Parameter	Description
1.	Problem Statement (Problem to be solved)	➢ In a current situation the human beings health in a bad condition . It is the water that is used daily that causes people to formed a diseases.
2.	Idea / Solution description	➢ Collecting data sets ,Analysing a water quality by comparing the current data set values with previous collected data set values and convey message to the user about the water quality and reusable methods also.
3.	Novelty / Uniqueness	➢ The information is sent to the user very quickly without any delay. ➢ Recycling methods are also reported along with water quality.
4.	Social Impact / Customer satisfaction	➢ Clean water consumption leads to healthy life. ➢ Learn about ways to reuse water without wasting it. ➢ Medical care, loss of productivity and even death can be avoided..
5.	Business Model (Revenue Model)	➢ Through advertisement can sell my project the private organization and public sectors can earn more. ➢ In an organization (hospital, school, college etc..) we explain our process to them and do water quality analysis and earn income in our business. ➢ Water quality is an important factor in economic development.

6.	Scalability of the solution	➢ Obtain quantitative information on the physical, chemical, and biological characteristics of water.

3.4 Problem Solution fit



4. REQUIREMENT ANALYSIS

4.1 Functional requirement

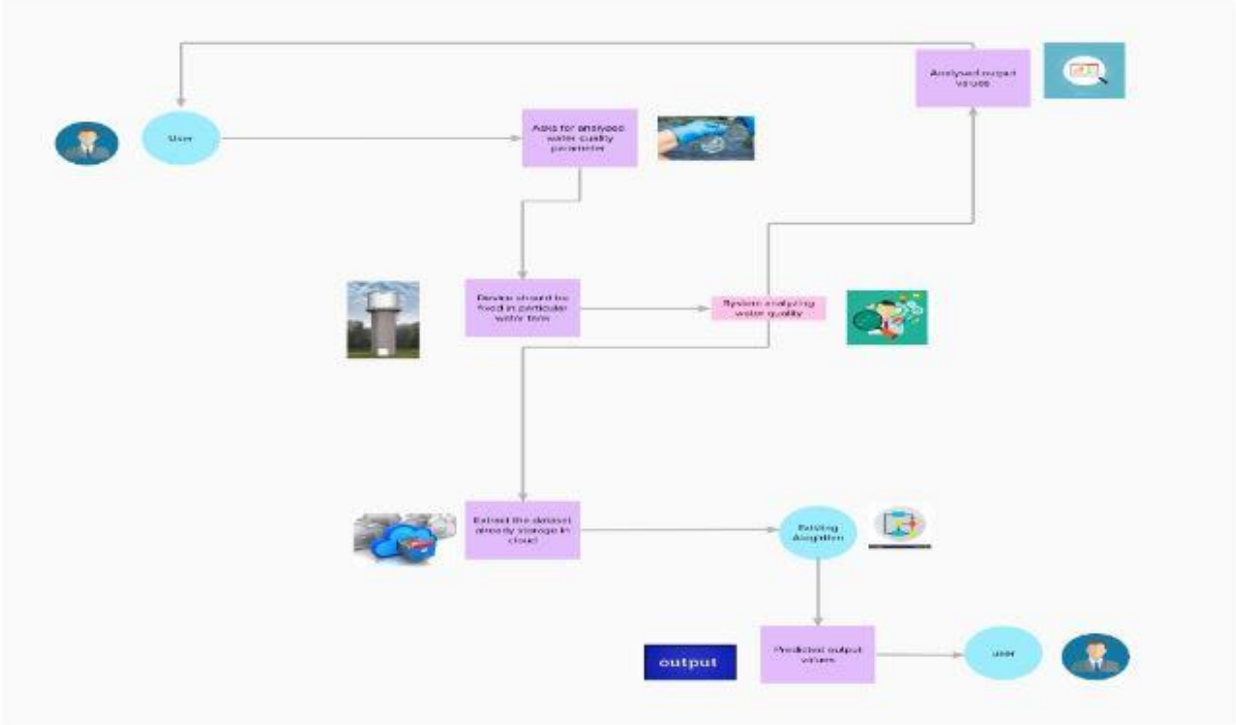
FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	To every family we will provide a form when a new user account is created
FR-2	User Confirmation	The system send an approval message after the user account is activated
FR-3	Authorization level	We provide secure water quality monitoring system approved by TNPCB(Tamilnadu pollution control board)
FR-4	Transaction processing system	analyse ,send a message about real time water quality and reusable method via online,
FR-5	Reporting	1.Analysis the real time water quality and send the message to the users. 2.The real time water quality report is collected and the dataset is using to predict the water in upcoming days
FR-6	business rules	1.Any one of the family member fill the appropriate form and provide the current usable mobile number 2.After receiving the verification message user send the confirmation message 3.We will providing our service continuously 4. If any problem occurs register the complaint in our website ,we will provide a immediate solution

4.2 Non-Functional requirements

FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	Allows users to identify specific missing data elements available in the water quality portal data.
NFR-2	**Security**	To ensure that the access of safe drinking water for all people in a country
NFR-3	**Reliability**	Above 90% of the operations that are completed correctly.
NFR-4	**Performance**	System effectively compare the incoming water quality parameters with the required dataset
NFR-5	**Availability**	This system is available for every family or any part of the area people.
NFR-6	**Scalability**	High mineral levels are found in water as well as Water Quality Index (WQI) and Water Quality Classification (WQC) are accurately predicted.

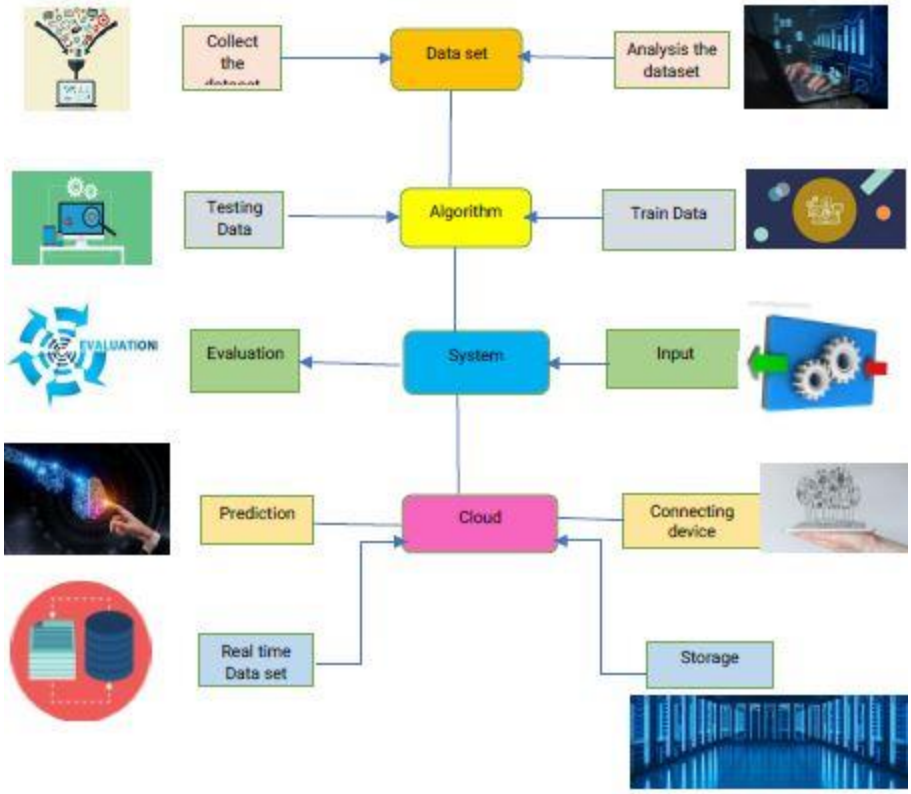
5. PROJECT DESIGN

5.1 Data Flow Diagrams



5.2 Solution & Technical Architecture

SOLUTION ARCHITECTURE FIT



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 through website	I can register and access the account with website	High	Sprint-1
		USN-4	As a user, I can register for the application through Gmail	I can register and access the gmail	Medium	Sprint-1
	Login	USN-5	As a user, I can log into the application by entering email & password	I can successfully login into application	High	Sprint-1
	Dashboard	USN-6	As a user,I can access the dashboard	I can referred dashboard for certainty	Medium	Sprint-1
Customer (Ordinary people,Industry)	Analysis the water quality	USN-7	As a user,I can access the water quality analysis in all over india	I can predict the water quality earlier	High	Sprint-1
Customer Care Executive	Customer queries	USN-8	As a user ,I can register the complaint in website	I can get immediate solution	High	Sprint-1
Administrator	Getting value	USN-9	when there is a issuses in getting analysed value	through administrator getting predicted value	Low	Sprint-2

6. PROJECT PLANNING & SCHEDULING

6.1 Sprint Planning & Estimation

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Analysing the water quality	USN-1	The man who wants to save our health to do the necessary steps	2	High	L.Aarthy P.Prakashini
Sprint-2	Prevention of drinking polluted water	USN-2	The officer worried about the people health because unpurified water causes many health issues	2	High	P.Prakashini M.Aarthi
Sprint-3	Detect the water quality	USN-3	The socialist can collect the various water parameters to detect the water quality at various environment	2	High	M.Aarthi K.Meenakshi
Sprint-4	Using hardware kit,cloud and various sensors	USN-4	The government take more steps to implementing this hardware setup	2	Medium	K.Meenakshi M.Aarthi
Sprint-5	Training and testing the water quality detection	USN-4	The programmer build a model for water quality detection by train the dataset	2	High	P.Prakashini L.Aarthy
Sprint-6	Notification	USN-6	The model can detect the water quality this detected values is notify to the users	2	High	L.Aarthy K.Meenakshi

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	4Days	24 Oct 2022	27 Oct 2022	20	27 Oct 2022
Sprint-2	20	4 Days	28 Oct 2022	01 Nov 2022	20	01 Nov 2022
Sprint-3	20	4 Days	02 Nov 2022	06 Nov 2022	20	06 Nov 2022
Sprint-4	20	4 Days	07 Nov 2022	10 Nov 2022	20	10 Nov 2022
Sprint-5	20	4 Days	11 Nov 2022	15 Nov 2022	20	15 Nov 2022
Sprint-6	20	4 Days	16 Nov 2022	19 Nov 2022	20	19 Nov 2022

6.3 Reports from JIRA

7. CODING & SOLUTIONING

7.1 Feature 1

	STATION CODE	LOCATIONS	STATE	Temp	D.O. (mg/l)	PH	CONDUCTIVITY (µmhos/cm)	B.O.D. (mg/l)	NITRATENAN N+ NITRITENANN (mg/l)	FECAL COLIFORM (MPN/100ml)	TOTAL COLIFORM (MPN/100ml)Mean	year
0	1393	DAMANGANGA AT D/S OF MADHUBAN, DAMAN	DAMAN & DIU	30.6	6.7	7.5	203	NAN	0.1	11	27	2014
1	1399	ZUARI AT D/S OF PT. WHERE KUMBARJRIA CANAL JOI...	GOA	29.8	5.7	7.2	189	2	0.2	4953	8391	2014
2	1475	ZUARI AT PANCHAWADI	GOA	29.5	6.3	6.9	179	1.7	0.1	3243	5330	2014
3	3181	RIVER ZUARI AT BORIM BRIDGE	GOA	29.7	5.8	6.9	64	3.8	0.5	5382	8443	2014
4	3182	RIVER ZUARI AT MARCAIM JETTY	GOA	29.5	5.8	7.3	83	1.9	0.4	3428	5500	2014
...	...	...	...	...	...	...	...	...	...	...	...	...
1986	1330	TAMBIRAPARANI AT ARUMUGANERI, TAMILNADU	NAN	NAN	7.9	738	7.2	2.7	0.518	0.518	202	2003
1987	1450	PALAR AT VANIYAMBADI WATER SUPPLY HEAD WORK, T...	NAN	29	7.5	585	6.3	2.6	0.155	0.155	315	2003
1988	1403	GUMTI AT U/S SOUTH TRIPURA,TRIPURA	NAN	28	7.6	98	6.2	1.2	NAN	NAN	570	2003
1989	1404	GUMTI AT D/S SOUTH TRIPURA, TRIPURA	NAN	28	7.7	91	6.5	1.3	NAN	NAN	562	2003
1990	1726	CHANDRAPUR, AGARTALA D/S OF HAORA RIVER, TRIPURA	NAN	29	7.6	110	5.7	1.1	NAN	NAN	546	2003

7.2 Feature 2

```
<!DOCTYPE html>
<html lang="en">

<head>
   <meta charset="UTF-8">
   <meta http-equiv="X-UA-Compatible" content="IE=edge">
   <meta name="viewport" content="width=device-width, initial-scale=1.0">
   <title>DEPLOYMENT</title>
   <link rel="stylesheet" href="../static/css/style.css">

</head>

<body>
   <header>
      <nav>
         <div class="row">
            <div class="row1">
               <img src="../static/logo.jpg" alt="logo">
            </div>
            <div class="row2">
               <h1>Water Quality Prediction</h1>
            </div>
         </div>
      </nav>
   </header>
   <main>
      <div class="column">
         <form action="/login" method="post">
```

```
            <label for=""></label>
            <input type="text" name="year" id="" placeholder="Enter Year">
            <label for=""></label>
            <input type="text" name="do" id="" placeholder="Enter D.O">
            <label for=""></label>
            <input type="text" name="ph" id="" placeholder="Enter PH">
            <label for=""></label>
            <input type="text" name="co" id="" placeholder="Enter Conductivity">
            <label for=""></label>
            <input type="text" name="bod" id="" placeholder="Enter B.O.D">
            <label for=""></label>
            <input type="text" name="na" id="" placeholder="Enter Nitratenen">
            <label for=""></label>
            <input type="text" name="tc" id="" placeholder="Enter Total Coliform">
            <label for=""></label>
            <div class="last">
               <input type="submit" value="Predict">
            </div>
            <div class="bor">
               {{showcase}}
            </div>

         </form>
      </div>
   </main>
  </div>
 </body>
</html>
```

8. TESTING

8.1 Test Cases

Test case ID	Feature Type	Component	Test Scenario	Pre-Requisite	Steps To Execute	Test Data	Expected R
LoginPage_TC_OO1	Functional	Home Page	Location Test Case	Indicate the near by location	By using GPS	http://127.0.0.1:5000/	Display the current l
LoginPage_TC_OO2	UI	Home Page	Industrial Test Case	Know the record of particular industry.	By using Advertisement	http://127.0.0.1:5000/	Detect the toxic cher
LoginPage_TC_OO3	Functional	Home page	Environment Test Case	Maintain the good environment	Provide more information about the surrounding of the river.	Username: chalam@gmail.com password: Testing123	Detect the nature of surrounded by the ri
LoginPage_TC_OO4	Functional	Login page	pH Test Cases	Accurate quality of water	By using pH sensor.	Username: chalam@gmail password: Testing123	Detect the water qua
LoginPage_TC_OO4	Functional	Login page	Purity and Dirty Test Cases	Good water as well as bad water	By comparing the pH level for good water and bad water.	Username: chalam@gmail.com password: Testing123678686786876876	Identify which kind c be drinked
LoginPage_TC_OO5	Functional	Login page	Agriculture Test Cases	Evergreen process	By using the different kind of strategy	Username: chalam password: Testing123678686786876876	Better growth in agri

8.2 User Acceptance Testing

- **Purpose of Document**

  The purpose of this document is to briefly explain the test coverage and open issues of the [ProductName] project at the time of the release to User Acceptance Testing (UAT).

- **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	Severity 4	Subtotal
By Design	15	4	3	4	27
Duplicate	1	0	2	0	3
External	1	2	0	1	4
Fixed	13	3	6	23	45
Not Reproduced	0	1	0	0	1
Skipped	0	1	0	1	2
Won't Fix	0	6	1	1	8
Totals	30	17	12	30	90

- **Test Case Analysis**

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

Section	Total Cases	Not Tested	Fail	Pass
Print Engine	7	0	0	7
Client Application	39	0	0	39
Security	4	0	0	4
Outsource Shipping	3	0	0	3
Exception Reporting	8	0	0	8
Final Report Output	5	0	0	5
Version Control	3	0	0	3

9. RESULTS

9.1 Performance Metrics

Model Performance Testing:

Project team shall fill the following information in model performance testing template.

S.No	Parameter	Values	Screenshot

1.	Metrics	**Model Evaluation:**
		MAE :
		1.0140200501253205
		MSE :
		5.786707157894741
		RMSE :
		2.405557556554143
		R2 score :
		0.9684566685516488
			Model Evaluation
			`[ ] from sklearn import metrics`
			`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)))`
			MAE: 1.0140200501253205
			MSE: 5.786707157894741
			RMSE: 2.405557556554143
			`[ ] metrics.r2_score(y_test, y_pred)`
			0.9684566685516488

2.	Tune the Model	Validation Method :
		Testing Accuracy
		0.9684566685516488
			Model Evaluation
			`[ ] from sklearn import metrics`
			`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_`
			MAE: 1.0140200501253205
			MSE: 5.786707157894741
			RMSE: 2.405557556554143
			`[ ] metrics.r2_score(y_test, y_pred)`
			0.9684566685516488

10. ADVANTAGES

1. In this technique, our model predicts that the water is safe to drink or not using some parameters like Ph value, conductivity, hardness, etc. Access to safe drinking-water is essential to health, a basic human right and a component of effective policy for health protection.

2. During the last years, water quality has been threatened by various pollutants. Therefore, modeling and predicting water quality have become very important in controlling water pollution.

11. DISADVANTAGES

1. The output of an algorithm after it has been trained on a historical dataset and applied to new data when forecasting the likelihood of a particular outcome.

2. Dataset collection is difficult, because more number of dataset is needed for training the model

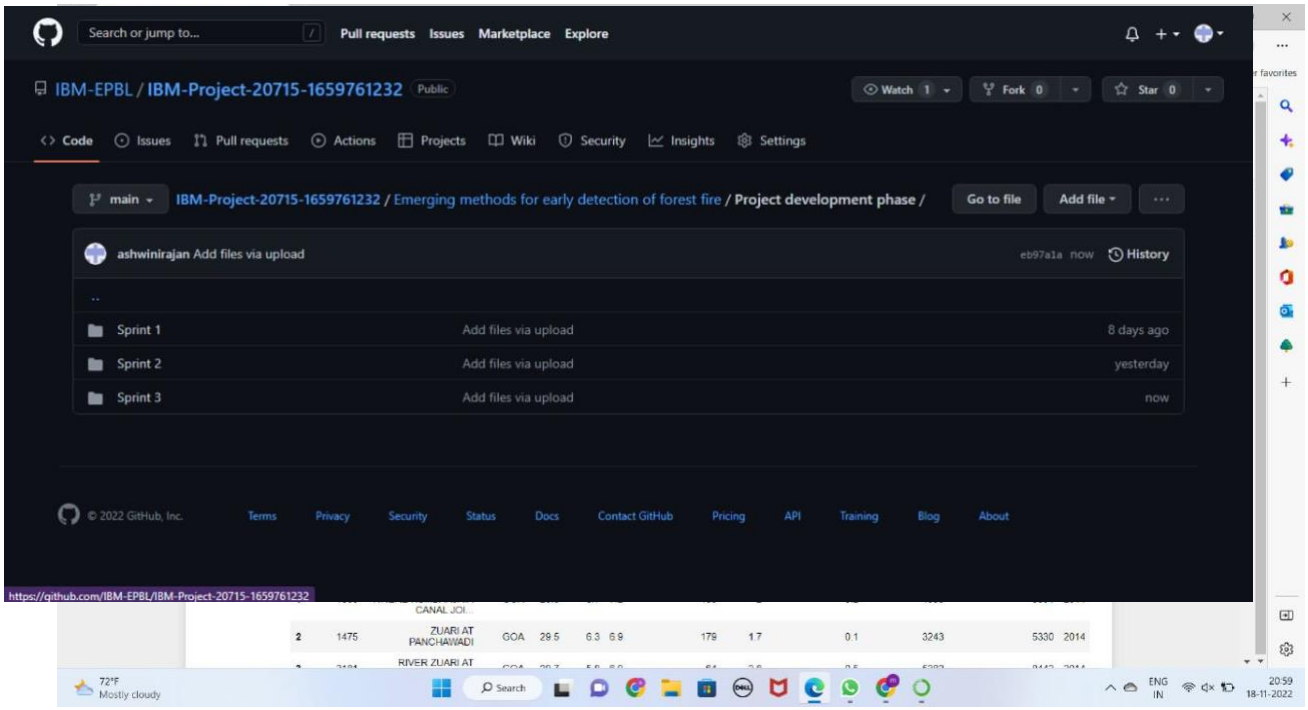
12. CONCLUSION

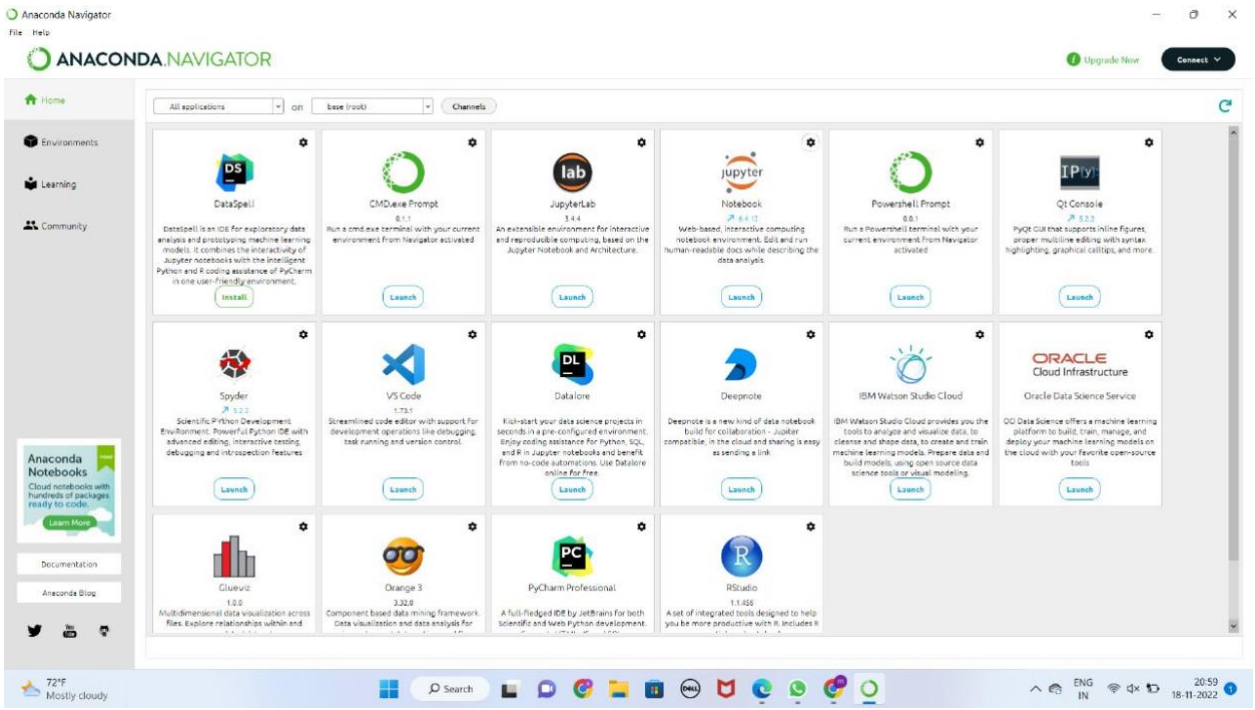
13. FUTURE SCOPE

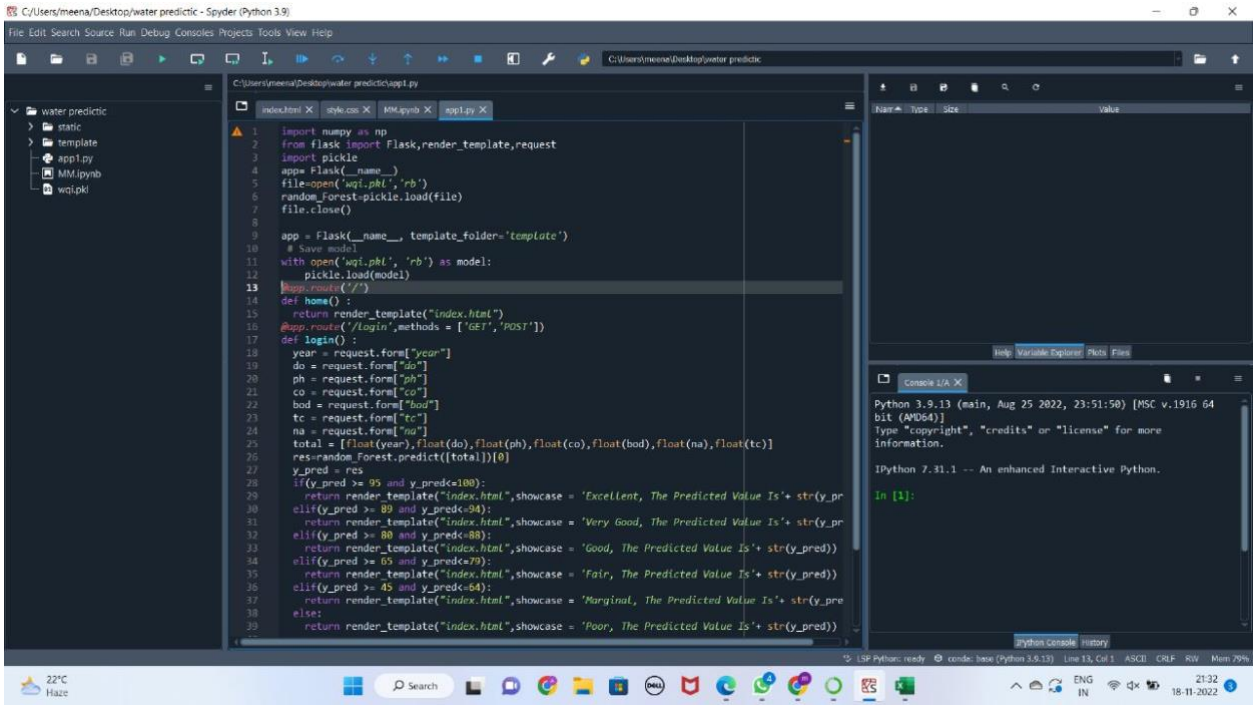
Machine learning models fail silently, which means they will make predictions even if the incoming data looks nothing like the data they were trained against. it allow businesses to make highly accurate guesses as to the likely outcomes of a question based on historical datathese historical data is used to build a mathematical model that captures important trends**.**

14. APPENDIX

Source Code


Search or jump to...
Pull requests
Issues
Marketplace
Explore
IBM-EPBL / IBM-Project-20715-1659761232
Public
Watch 1
Fork 0
Star 0
Code
Issues
Pull requests
Actions
Projects
Wiki
Security
Insights
Settings
main
IBM-Project-20715-1659761232 / Emerging methods for early detection of forest fire / Project development phase /
Go to file
Add file
ashwinirajan Add files via upload
eb97a1a now
History
Sprint 1
Add files via upload
8 days ago
Sprint 2
Add files via upload
yesterday
Sprint 3
Add files via upload
now
© 2022 GitHub, Inc.
Terms
Privacy
Security
Status
Docs
Contact GitHub
Pricing
API
Training
Blog
About
https://github.com/IBM-EPBL/IBM-Project-20715-1659761232
CANAL JOI...
2 1475 ZUARI AT PANCHAWADI GOA 29.5 6.3 6.9 179 1.7 0.1 3243 5330 2014
RIVER ZUARI AT
72°F
Mostly cloudy
Search
ENG
IN
20:59
18-11-2022


Anaconda Navigator
File Help
ANACONDA.NAVIGATOR
Upgrade Now
Connect
Home
Environments
Learning
Community
All applications
on
base (root)
Channels
DataSpell
CMD.exe Prompt
JupyterLab
Notebook
Powershell Prompt
Qt Console
Spyder
VS Code
Datalore
Deepnote
IBM Watson Studio Cloud
Oracle Cloud Infrastructure
Oracle Data Science Service
Glueviz
Orange 3
PyCharm Professional
RStudio
Anaconda Notebooks
Cloud notebooks with hundreds of packages ready to code.
Learn More
Documentation
Anaconda Blog
Launch
Install
72°F
Mostly cloudy
Search
ENG
IN
20:59
18-11-2022



GitHub & Project Demo Link

https://drive.google.com/file/d/1ceDPBP4zZ3ObLyyy5uTSFU4dS6CL58Wn/view?usp=share_link