

Project Report

PROJECT REPORT

Project name	Efficient water quality analysis prediction using machine learning
Team id	PNT2022TMID52107
Team members	M.Srisakthi J.Shyja jenifer M.Sindhu K.Sujith

INTRODUCTION

Project Description:

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 non-linearly 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.

Purpose:

- To evaluate the quality of water and determine the safety of water.
- To monitor changes in water quality.
- To determine whether water is suitable for the health of the natural environment.
- To determine whether water is suitable for human consumption and other uses.

LITERATURE SURVEY

Problem Statement:

One in nine people worldwide uses drinking water from unimproved and unsafe sources. 2.4 billion people live without any form of sanitation.

- Water is one of the most essential for the existence of life. The safety and accessibility of drinking-water are major concerns throughout the globe.
- Water makes up about 70% of the surface and is one of the most important sources

vital to sustaining life.

- Water quality has been conventionally estimated through expensive and time consuming lab and statical analysis.
- This system is proposed to check the water quality and warn the user before water gets contaminated using Machine Learning.

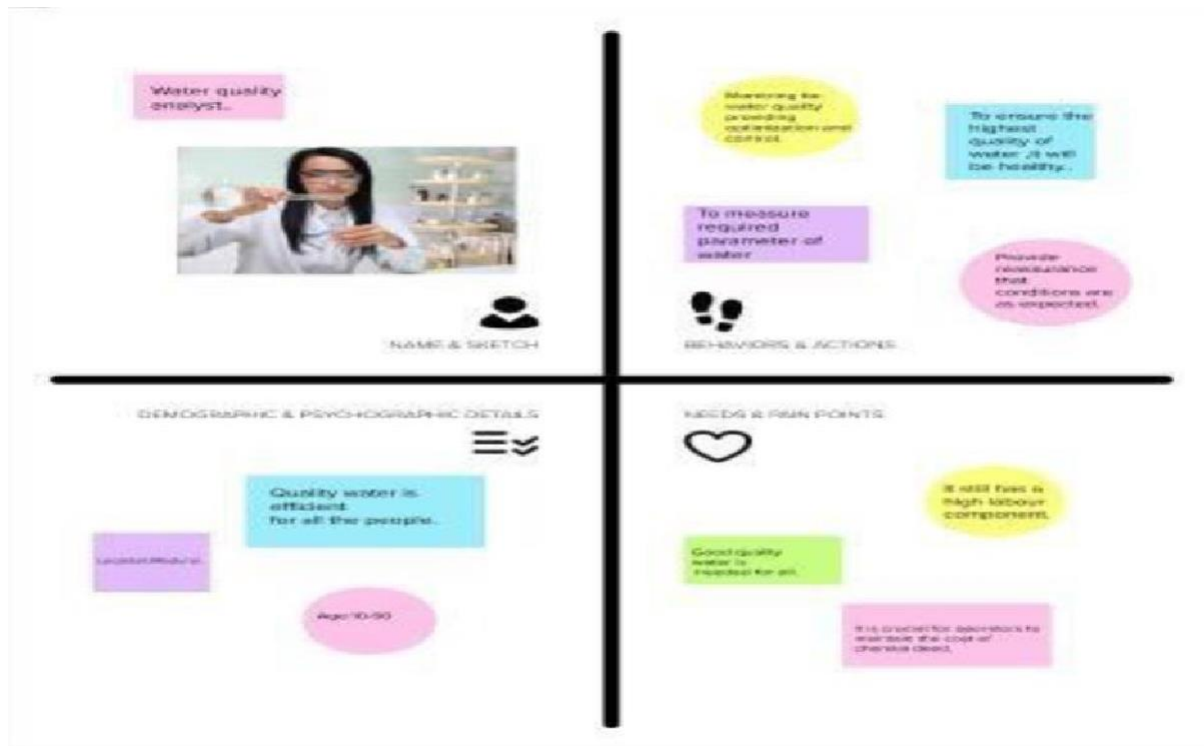
References:

Abobakr Saeed Abobakr Yahya , Ali Najah Ahmed , Faridah Binti Othman , Rusul Khaleel Ibrahim , Haitham Abdulmohsin Afan , Amr El-Shafie , Chow Ming Fai , Md Shabbir Hossain , Mohammad Ehteram and Ahmed Elshafie **“Water Quality Prediction Model Based Support Vector Machine Model for Ungauged River Catchment under Dual Scenarios”**

[Water](#) | [Free Full-Text](#) | [Water Quality Prediction Model Based Support Vector Machine Model for Ungauged River Catchment under Dual Scenarios \(mdpi.com\)](#)

IDEATION AND PROPOSED SOLOUTION

Empathy Map:



Ideation & Brainstorming:

Brain strom.pdf - Adobe Acrobat Reader (64-bit)

File Edit View Sign Window Help

Home Tools Brain strom.pdf x

1 / 1 12.5%

2 Brainstorm

Write down any ideas that come to mind that address your problem statement.

10 minutes

Person 1 Person 2 Person 3 Person 4

Person 5 Person 6 Person 7 Person 8

3 Group ideas

Take turns sharing your ideas while clustering similar or related notes as you go. Once all sticky notes have been grouped, give each cluster a sentence-like label. If a cluster is bigger than six sticky notes, try and see if you can break it up into smaller sub-groups.

20 minutes

monitoring water resources health predict potentially harmful condition communicate impacts of managing and policy decision land drainage and flood protection

Search 'Fill Form'

Export PDF

Adobe Export PDF

Convert PDF Files to Word or Excel Online

Select PDF File

Brain strom.pdf

Convert to

Microsoft Word (*.docx)

Document Language: English (U.S.) Change

Convert, edit and e-sign PDF forms & agreements

Free 7-Day Trial

Brain strom.pdf - Adobe Acrobat Reader (64-bit)

File Edit View Sign Window Help

Home Tools Brain strom.pdf x

1 / 1 12.5%

4 Prioritize

Your team should all be on the same page about what's important moving forward. Place your ideas on this grid to determine which ideas are important and which are feasible.

20 minutes

drinking water

electricity generation

nature land cover

Importance

After you collaborate

You can export the mural as an image or pdf to share with members of your company who might find it helpful.

Quick add-ons

Share the mural

Export the mural

Keep moving forward

Strategy blueprint

Customer experience journey map

Strengths, weaknesses, opportunities & threats

Share template feedback

Proposed Solution:

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	To establish safe drinking water sources in the future, it is imperative to understand the quality and pollution level of existing groundwater. The prediction of water quality with high accuracy is the key to controlling water pollution and the improvement of water management.
2.	Idea / Solution description	In this study , a machine learning (ML) based model is proposed for predicting groundwater quality and compared with three other machine learning(ML) models, namely, random forest, extreme gradient boosting and artifical neural network(ANN).
3.	Novelty / Uniqueness	<p>☐ The main reason behind the success of this ML technique is that it ignore the requirements of selective featuirees that are most representative compared to that of traditional ML algorithm.</p> <p>☐ ML methods can progressively construct highlevel attributes from the given dataset.</p> <p>☐ ML technique is a self deterministic approach that learns features to discover the correct representation required for the given task .</p>
4.	Social Impact / Customer Satisfaction	Generation of unprocessed effluent, municipal refuse, factory wastes, and junking of compostable and non compostable effluents has hugely contaminated nature provided water bodies like rivers,lake ,and ponds. Therefore, there is a necessity to look into the water

5.

Business Model (Revenue Model)

standard before usage .This is a problem that can greatly benefit from artificial intelligence. Traditional methods require human inspection and are time consuming.

Assessment of water quality using conventional methods causes losses in economic value, which in turn affect the decision-making

Problem Solution fit:

problem solution fit.pdf - Adobe Acrobat Reader (64-bit)

File Edit View Sign Window Help

Home Tools Proposed Solution ... problem solution fi... x

1 / 1 61%

Project Title: Efficient Water Quality Analysis and Prediction using Machine Learning

Project Design Phase-I - Solution Fit Template

Team ID: PNT2022TMDIS2107

Problem-Solution Fit canvas

Purpose / Vision		Version
1. CUSTOMER SEGMENTS (C-2) <ul style="list-style-type: none"> Urban people's Stakeholder's of RO based companies. Manufacturing companies. 	6. CUSTOMER LIMITATIONS (C-3) <ul style="list-style-type: none"> Spending power Budget Lack of efficient computer system Untrained customers 	5. AVAILABLE SOLUTIONS (P-3) & (C-3) <ul style="list-style-type: none"> Chemical methods AI techniques
2. PROBLEMS / PAINS (P-2) <ul style="list-style-type: none"> Urban people are mostly self-employed their livelihood are not stable. So, this method will be a cost efficient method for them. To check whether the water quality is in compliance with the standards, and hence, suitable or not for the designated use. 	9. PROBLEM ROOT / CAUSE (P-3) <p>People think that testing the water quality for normal usage are bad investment right now because their too expensive, and possible changes to law might influence the return of investment significantly and diminish the benefits.</p>	7. BEHAVIOR (P-3) & (C-3) <ul style="list-style-type: none"> Choosing of efficient providers. When their expected standard of water is achieved we can expect this behaviour
3. TRIGGERS TO ACT (T-2) <ul style="list-style-type: none"> Seeing their neighbours using efficient water quality analysis method for their individual purpose. Reading about innovative and efficient solutions 	10. YOUR SOLUTION (S-3) <p>This ML technique is an extension of the artificial neural network method; it has additional complex architectures that make this approach suitable for managing multi-dimensional inputs because of its high model configuration flexibility, greater generalization power, and robust learning capacity.</p>	8. CHANNELS OF BEHAVIOR (C-3) <p>Extract channels from behaviour block</p>
4. EMOTIONS (E-2) & (P-2) <ul style="list-style-type: none"> Before the implementation of this system people were infuriated about their water needs. After accomplishing this system they will be reimbursed. 		<p>Extract online & offline CO of BE</p> <p>Extract online & offline CO of BE</p>

Search 'Draw Line'

Export PDF

Adobe Export PDF

Convert PDF Files to Word or Excel Online

Select PDF File

problem solution fit.pdf

Convert to

Microsoft Word (*.docx)

Document Language: English (U.S.) Change

Convert, edit and e-sign PDF forms & agreements

Free 7-Day Trial

Type here to search

20:50 18-11-2022

REQUIREMENT ANALYSIS

Functional requirement:

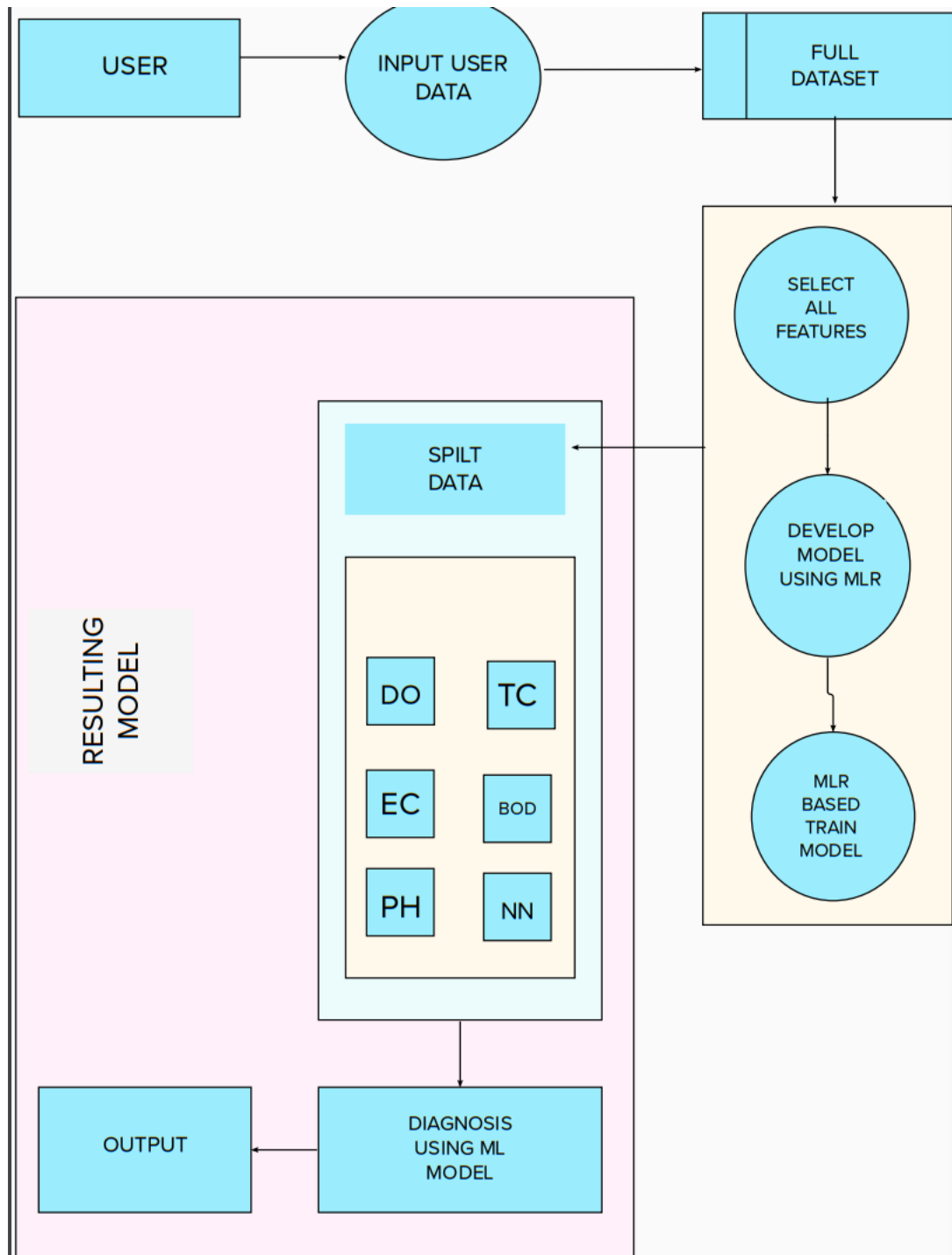
FR. No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User registration	Registration through Gmail Create an account Follow the instructions
FR-2	User Confirmation	Confirmation via Email and it is predicted by water level sensor
FR-3	Interface sensor	Interface sensor and Water level sensor produces the detection of clean drinking water
FR-4	Accessing datasets	Datasets are collected by data preprocessing method.
FR-5	Mobile application	The efficient of water quality is analyzed, the mobile application is not used ..

Non-functional Requirements:

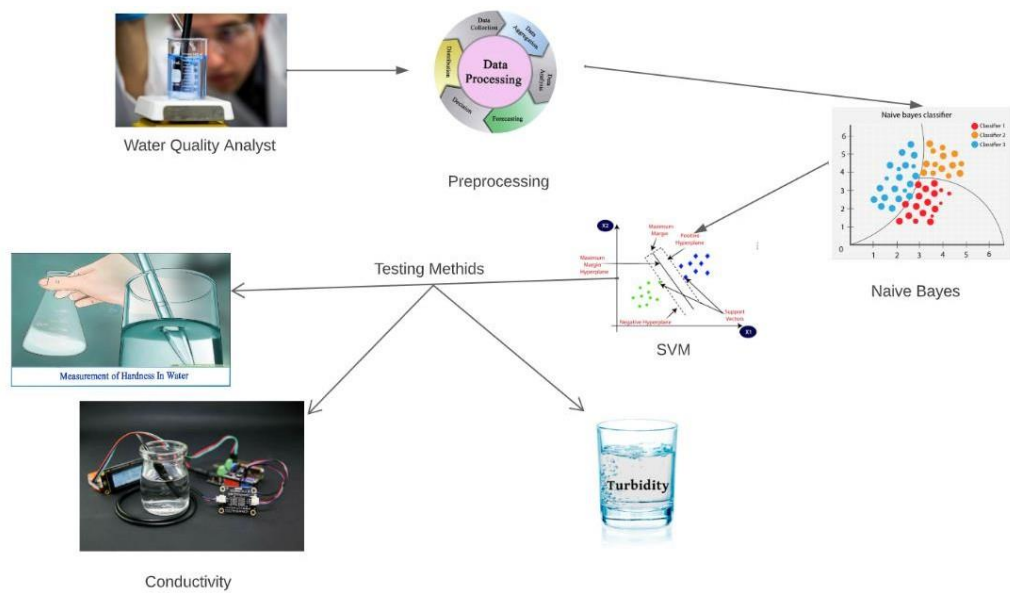
NFR-2	Security	We have designed this project to secure the people from drinking the impurity water.
NFR-3	Reliability	This project will help everyone in protecting their health. Accurate water quality prediction is the basis of water environment management and is of great significance for water environment protection
NFR-4	Performance	This system uses different sensors for monitoring the water quality by determine pH, Turbidity, conductivity and temperature. The data preprocessing access the dataset. With the use of this we predict the quality water.
NFR-5	Availability	By developing and deploying resilient hardware and software we can analyze the drinking water .
NFR-6	Scalability	This project used to measure and determine the quality of water. This provide pollution free and purified water.

PROJECT DESIGN

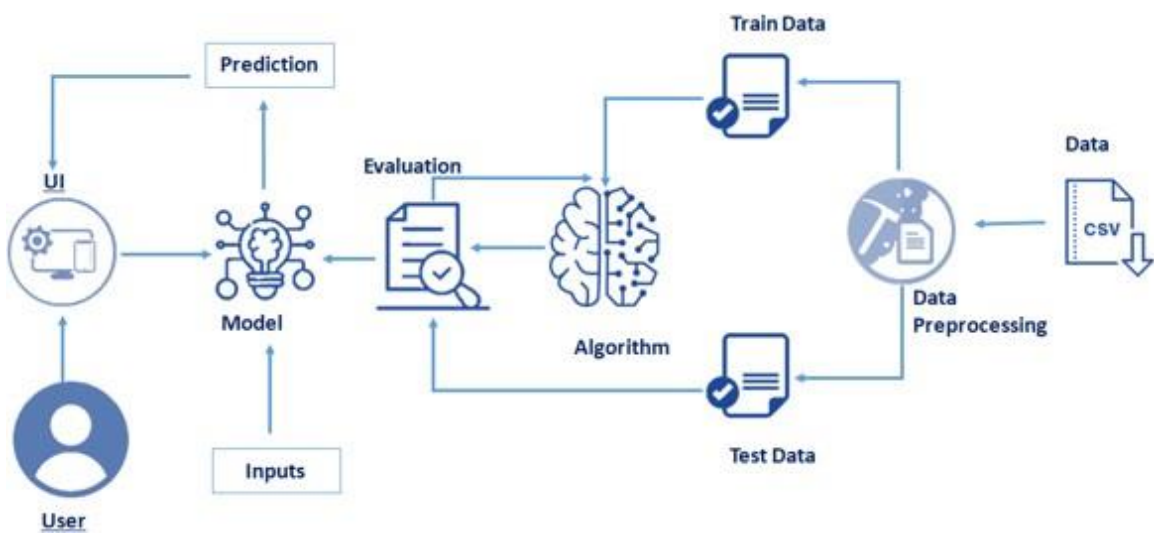
Data Flow Diagrams:



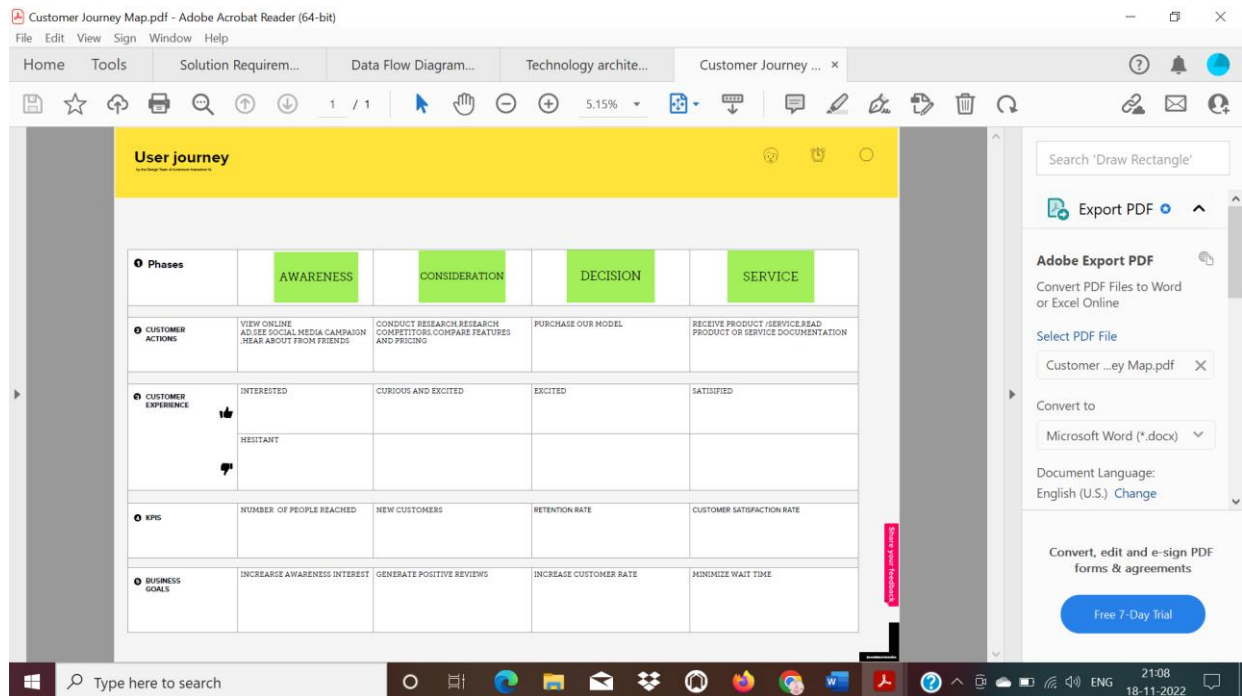
Solution Architecture:



Technical Architecture



User Journey



PROJECT PLANNING & SCHEDULING

Sprint	Function al Require ment (Epic)	User Story Number	User Story / Task	Story Point
Sprint-1	Data Collection	USN-1	Collect the appropriate dataset predicting the water quality.	10
Sprint-1		USN-2	Data Preprocessing – Used to transform the data into useful format.	7
Sprint-2	Model Building	USN-3	Calculate the Water Quality Index (WQI) using Regression algorithm of Machine Learning.	10

Project Tracker, Velocity & Burndown Chart: (4 Marks)

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	24 Oct 2022	27 Oct 2022	20	29 Oct 2022
Sprint-2	20	5 Days	28 Oct 2022	01 Nov 2022	20	04 Nov 2022
Sprint-3	20	8 Days	02 Nov 2022	09 Nov 2022	20	11 Nov 2022
Sprint-4	20	9 Days	10 Nov 2022	18 Nov 2022	20	19 Nov 2022

S.NO	MILESTONE	DESCRIPTION	DURATION	Working Status
1.	Prerequisites	Prerequisites are all the needs at the requirement level needed for the execution of the different phases of a project.	1 WEEK	Completed
2.	Ideation	Ideation process is where you generate ideas and solutions through sessions such as Sketching, Prototyping, Brainstorming, Worst Possible ideas and wealth of other techniques.	1 WEEK	Completed
3.	Project design phase	Project design is an early phase of a project where the project's key features, structure, criteria for success, and major deliverables are planned out. The	1 WEEK	Completed

		aim is to develop one or more designs that can be used to achieved the desire goals.		
4.	Project Planning Phase	In the Planning Phase, the Project Manager works with the team to create the technical design, task list, resources, communication plan, budget and initial schedule for project.	1 WEEK	Completed
5.	Data Collection and Data pre - processing	A Data collection is a process of gathering and measuring information on variables to ensure accuracy and facilitate analysis. It help to solve the critical workloads.	1 WEEK	Completed
6.	Model Building	Model Building is used for project visualization to provide information about the proposed state. It helps to identify the quality of objectives and it formulate the conceptual model.	4 WEEKS	Completed
7.	Develop Application	A web application is application software that runs in a web browser, unlike software programs that	4 WEEKS	Completed

		run locally and natively on the operating system of the device.		
8.	Project development phase	Project development is the process of planning and allocating resources to fully develop a project or product from concept to go - live.	4 WEEKS	Completed

CONCLUSION:

The prediction of water quality is very vital in monitoring the pollution and in sustaining the availability of potable water resources. Undoubtedly, it can afford early warnings when the water quality changes as well as it can reduce the adverse consequences resulting from the poor water quality. Herein, the SVM approach was introduced to estimate the water quality of Langat River Basin using six parameters. The presented model accurately estimated the water quality factors with relatively minor prediction errors, proving a quite efficient and robust performance. The model also can help in the optimization of water quality monitoring plans by decreasing the frequency, quantity of sampling sites, and water quality factors. Prediction precision with the maximum error was equal to 1% and CC was equal to 0.9987. Even though the outcomes seem to be reasonable, the application of water quality parameters is quite sensitive to the error level. 1% as a maximum error is comparatively on the higher side in such an application which triggers the need to improve it. In this regard, it is suggested to deploy the optimal kernel parameters determined and choose the Nu-RBF model as the optimal model.

GitHub & Project Demo Link

<https://github.com/IBM-EPBL/IBM-Project-47808-1660802465>

<https://youtu.be/1vX3cpRx4T0>