SMART INTERNZ

Predicting and Analyzing Urban Water Quality using IBM Watson Machine Learning Service

Index

Topics	Page No
1. Introduction	2
a. Overview	
b. Purpose	
2. Literature Survey	3
a. Existing problem	
b. Proposed solution	
3. Theoretical Analysis	4
a. Block diagram	
b. Hardware/Software designing	
4. Experimental investigations	5
5. Flowchart	6
6. Result	7
7. Advantages & Disadvantages	7

8. Applications 8

9. Conclusion 8

10. Future Scope 8

11. Bibliography 8

12. Appendix 9

a. Source code

b. UI output Screenshots

1. Introduction :

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 by considering all water quality standard indicators.

a. Overview :

User interacts with the UI (User Interface) to enter Data.The entered data is analyzed by the model which is integrated.Once model analyses the input the prediction is showcased on the UI.

To accomplish this, we have to complete all the activities and tasks listed below

- Data Collection.
    - Collect the dataset or Create the dataset
- Data Pre processing.

    - Import the Libraries.
    - Importing the dataset.
    - Checking for Null Values.
    - Data Visualization.
    - Taking care of Missing Data.
    - Label encoding.
    - One Hot Encoding.
    - Feature Scaling.
    - Splitting Data into Train and Test.
- Model Building
    - Training and testing the model
    - Evaluation of Model
- Application Building
    - Create an HTML file
    - Build a Python Code

b. Purpose :

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 by considering all water quality standard indicators.
By the end of this project we'll be able to understand:

- Regression and Classification Problems
- To grab insights from data through visualization.
- Applying different algorithms according
- Evaluation metrics
- how to build a web application using the Flask framework.

2. Literature Survey :

a. Existing Problem :

Predicting urban water quality, however, is very challenging due to the following reasons. First, urban water quality varies by locations non-linearly and depends on multiple factors, such as meteorology, water usage patterns, land

use, and urban structures.Most ambient water bodies such as rivers, lakes, and streams have specific quality standards that indicate their quality. Moreover, water specifications for other applications/usages possess their standards. For example, irrigation water must be neither too saline nor contain toxic materials that can be transferred to plants or soil and thus destroying the ecosystems. Water quality for industrial uses also requires different properties based on the specific industrial processes. Some of the low-priced resources of fresh water, such as ground and surface water, are natural water resources. However, such resources can be polluted by human/industrial activities and other natural processes.

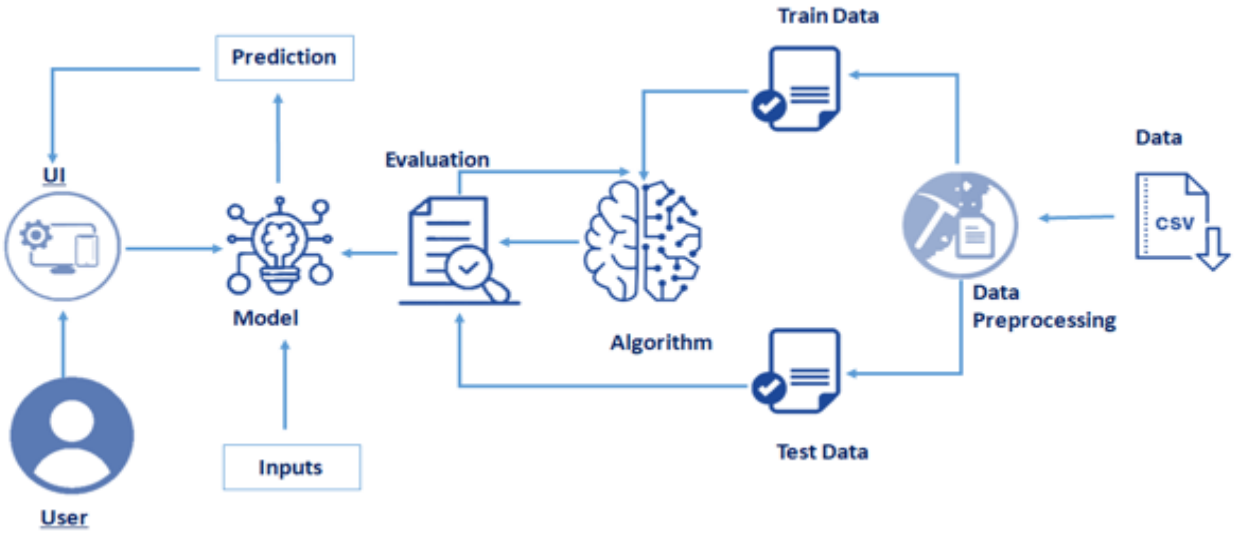
b. Proposed Solution :

In this project we are going we are going to develop a machine learning application using the IBM Watson machine learning service to predict and analyze the urban water quality, to know whether the water is safe for the purpose of using or not. We are going to take the input values of D.O, PH, Conductivity, B.O.D, Nitrate, and Total Coliform to calculate the predicted value of water quality. It can be done by calculating the Water Quality Index(WQI).

3. Theoretical Analysis :

User will give the inputs to the UI and the inputs are given to the machine learning model. The model gets trained data that has already been given. The model then evaluates the inputs given by the user and predicts the urban water quality.

a. Block Diagram :



B. Software Designing :

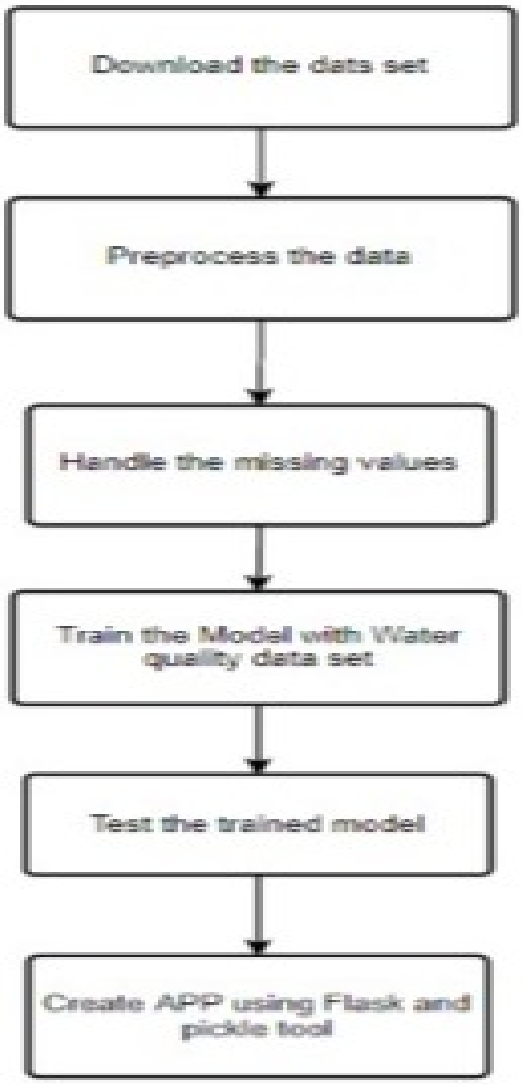
First we need to write the machine learning algorithm in the notebook and then we train the machine learning algorithm using the data. And then we integrate our model with the Flask using the pickle tool. We can run the app using flask framework to get an url and run our project in that webpage to give the input values and to predict the quality of the urban water. We also need to write the HTML codes for the webpage to take the inputs and integrate the codes with the algorithm. By doing so we get these values from the webpages to predict the Water Quality Index.

4. Experimental Investigations :

Living organisms need water with enough quality to continue their lives. There are certain limits of pollution that water species can tolerate. Exceeding these limits affects the existence of these creatures and threatens their lives.Most ambient water

bodies such as rivers, lakes, and streams have specific quality standards that indicate their quality. Moreover, water specifications for other applications/usages possess their standards. For example, irrigation water must be neither too saline nor contain toxic materials that can be transferred to plants or soil and thus destroying the ecosystems. Water quality for industrial uses also requires different properties based on the specific industrial processes. Some of the low-priced resources of fresh water, such as ground and surface water, are natural water resources. However, such resources can be polluted by human/industrial activities and other natural processes. Moreover we can use machine learning algorithms such as SVM, KNN, and Naive Bayes were used to classify the WQI data.

5. Flowchart :



6. Result :

The result will be a webpage which can be opened through the link provided by the flask app. We need to give the input values of

- Year
- D.O
- P.H
- Conductivity
- B.O.D
- Nitratenen
- Total Coliform

These values are given as inputs to the webpage as inputs, the algorithm then calculates the Water Quality Index(WQI) as the output and shows the predicted value as the output.

7. Advantages & Disadvantages :

Advantages :

- We can find the proper water quality index suitable for organisms.
- By the wqi we get the overall description of the quality of water bodies. It provides a quick and simple methodology to identify the quality of water.
- It indicates the health of water resources and create a yardstick for measuring and assessing the water quantity.
- Helps to translate a wide variety of environmental indicators into a simple system that can be easily communicated.

Disadvantages :

- We can not accurately calculate the water quality all the time, the water quality may vary with many aspects that might not be considered in this project.
- A single index cannot tell the whole story; an index may indicate that a river is unfit for drinking water standards, but the river may be fine for swimming and a healthy habitat for fish and macro invertebrates.
- An index is often limited in terms of time and space-- an unusual reading

at one location or during a particular time-period can skew the index.

8. Applications :

- It is used to for assessing the thermal water discharge that influences the quality of river water.
- Water quality index is used to find the heavy metal pollution of water.
- Water quality index is used to find the fish water pollution.
- Water quality indes is used to find the fresh water habited pollution.

9. Conclusion :

By using the water data, by training the model in the data using pickle tool and flask app we created the api of our model. Which will predict the water quality index(WQI). The water quality index that is predicted is used to analyze the quality of the water resource.

10. Future Scope :

Measuring the quality of water is very important now-a-days. There are many pollutants that reduce the water quality and they will increase in number in the future. So finding the quality of water is a necessary aspect to ensure that the water quality is great. By getting the water quality index we get to know the situation of a particular water body. Then we can take some measures to improve the quality of the water at that particular ecosystem for the organisms to use the good quality water.

11. Bibliography :

https://www.researchgate.net/publication/304188597_Predicting_and_analyzing_water_quality_using_Machine_Learning_A_comprehensive_model

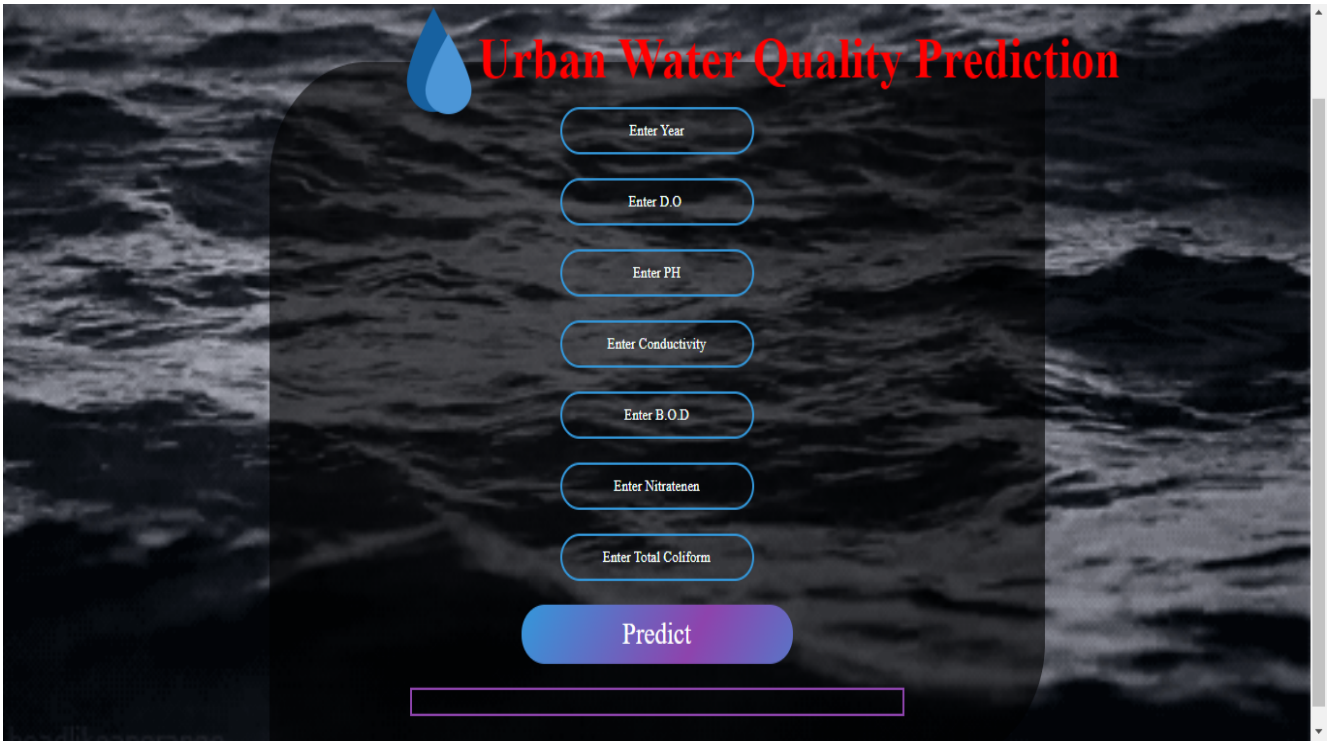
https://iwaponline.com/wqrj/article/53/1/3/38171/Water-quality-prediction-using-machine-learning

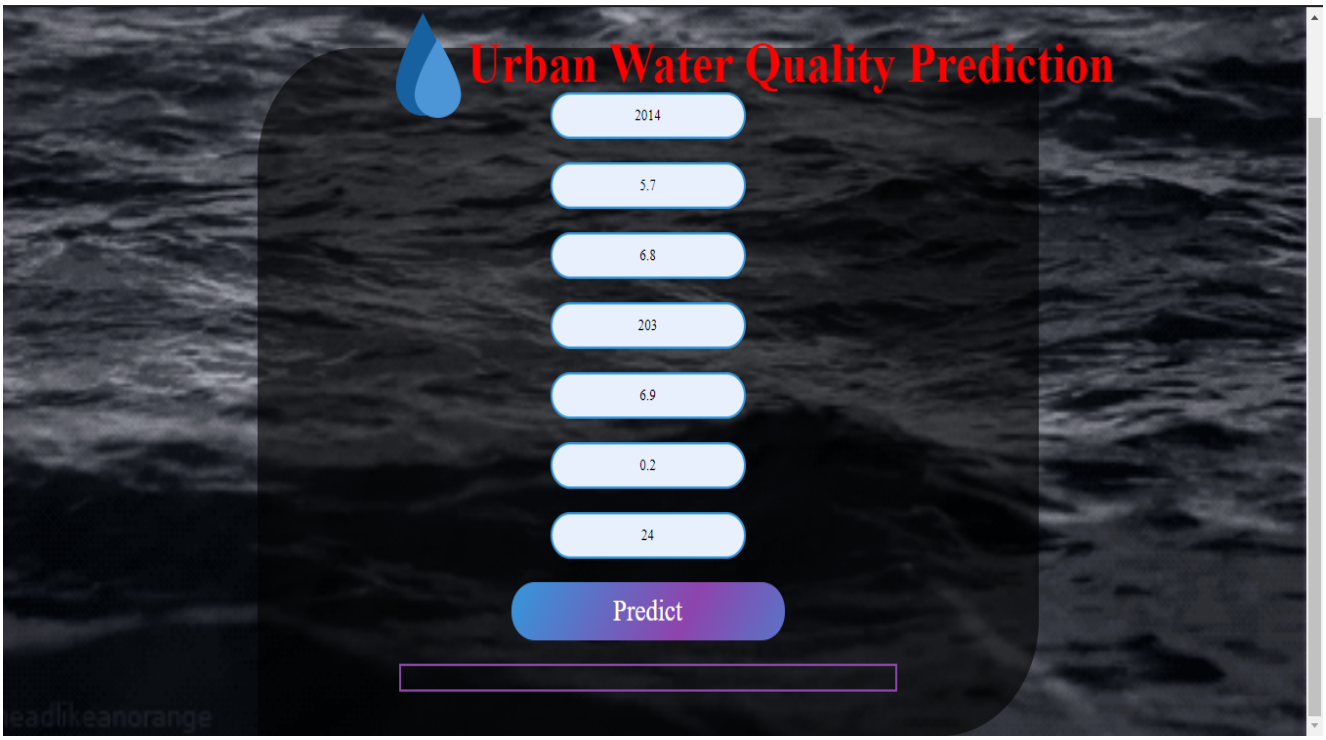
12. Appendix :

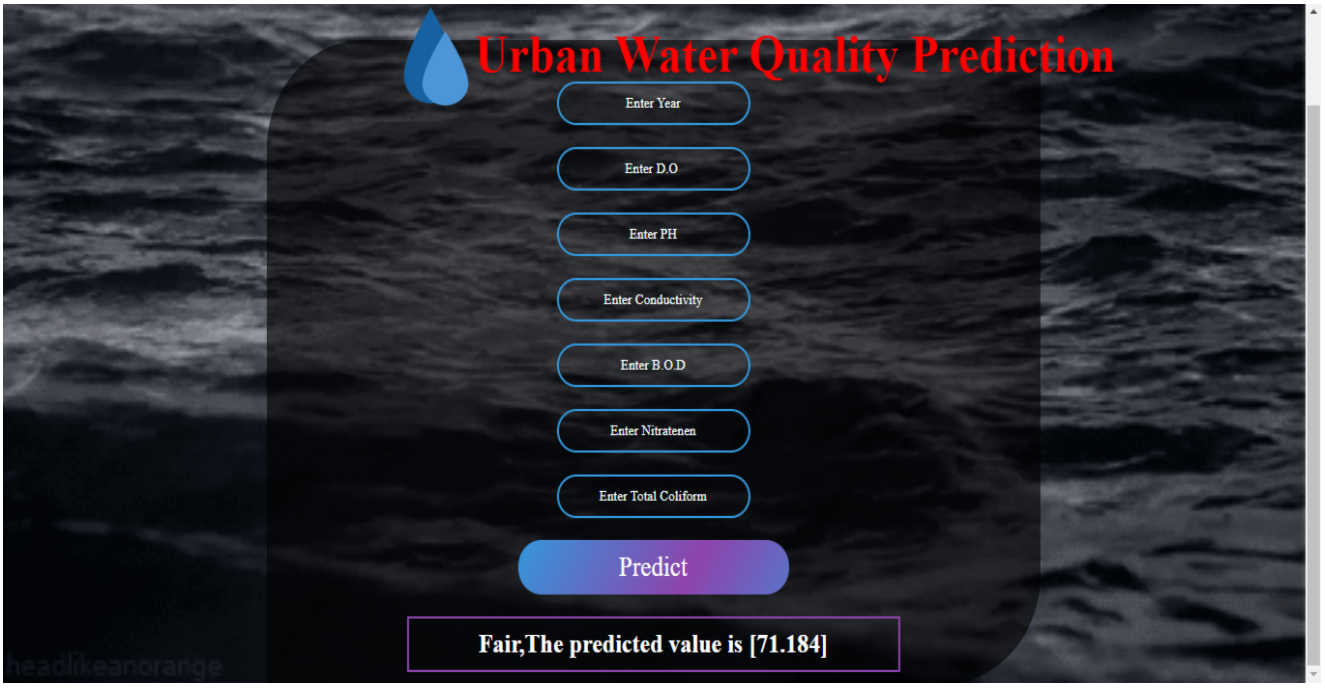
a. Source Code :

- Installing the required python packages and anaconda software.
- Data collection.
- Data pre-processing.
- Model building.
- Application building.
- Train the model on IBM.

b. UI Output Screenshots :




Urban Water Quality Prediction
2014
5.7
6.8
203
6.9
0.2
24
Predict


Urban Water Quality Prediction
Enter Year
Enter D.O
Enter PH
Enter Conductivity
Enter B.O.D
Enter Nitratenen
Enter Total Coliform
Predict
Fair,The predicted value is [71.184]