WATER QUALITY PREDICTION USING MACHINE LEARNING ALGORITHM

A Project Report
Submitted in partial fulfilment of the requirements for the award of the degree

of

BACHELOR OF TECHNOLOGY

In

COMPUTER SCIENCE AND ENGINEERING

by

G. Anusha -2010115

K. Vineeth Kumar -2010121

M. Hemanth Kumar Reddy -2010125

Under the esteemed guidance of

Sri. P.R. RAJESH KUMAR, M. Tech,(Ph.D)
Lecturer & Head of the Department

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

SRI KRISHNADEVARAYA UNIVERSITY

COLLEGE OF ENGINEERING AND TECHNOLOGY

ANANTAPUR – 515003

ANDHRA PRADESH

2020-2024

SRI KRISHNADEVARAYA UNIVERSITY

COLLEGE OF ENGINEERING AND TECHNOLOGY

ANANTAPUR – 515003

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

Certified that this is a bonafide record of the dissertation work entitled, **"WATER QUALITY PREDICTION USING MACHINE LEARNING ALGORITHM",** done by **G. ANUSHA** bearing Admin. No: (2010115), **K. VINEETH KUMAR** bearing Admin. No: (2010121), **M. HEMANTH KUMAR REDDY** bearing Admin. No: (2010125) submitted to the faculty of Computer Science and Engineering, in partial fulfillment of the requirements for the Degree of **BACHELOR OF TECHNOLOGY** with specialization in **COMPUTER SCIENCE AND ENGINEERING** from Sri Krishnadevaraya University College of Engineering and Technology, Anantapur.

Signature of the Supervisor

MR.P.R. Rajesh Kumar, M.Tech,(Ph.D)

Lecturer, Dept of CSE,

SKU College of Engg. & Tech.,

S.K.University,

Ananthapuramu.

Signature of the Head of the Department

MR.P.R. Rajesh Kumar, M.Tech,(Ph.D)

Lecturer & Head, Dept of CSE,

SKU College of Engg. & Tech.,

S.K.University,

Ananthapuramu.

SRI KRISHNADEVARAYA UNIVERSITY

COLLEGE OF ENGINEERING AND TECHNOLOGY

ANANTAPUR – 515003

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

DECLARATION

We hereby declare that the project report entitled **"WATER QUALITY PREDICTION USING MACHINE LEARNING ALGORITHM"** submitted to the Department of Computer Science and Engineering, Sri Krishnadevaraya University, Anantapuramu for the partial fulfilment of the academic requirement for the degree for Bachelor of Technology in Computer Science and Engineering is an authentic record of our work carried out during the final year under the esteemed guidance of **MR.P.R.Rajesh Kumar, M.Tech,(Ph.D)** Lecturer Computer Science and Engineering Department, College of Engineering and Technology, Sri Krishnadevaraya University, Ananthapuramu.

Signature of the students

1.

2.

3.

ACKNOWLEDGEMENT

The satisfaction and euphoria that accompany the successful completion of any task would be incomplete without the mention of people who made it possible whose constant guidance and encouragement crowned our efforts with success. It is a pleasant aspect that I have now the opportunity to express my gratitude for all of them.

It is with immense pleasure that I would like to express my indebted gratitude **to** my Guide **P.R.RAJESH KUMAR, M.Tech, (Ph.D)** in CSE Department, who has guided me a lot and encouraged me in every step of the project work. His valuable moral support and guidance throughout the project helped me to a greater extent. I thank him for his stimulating guidance, constant encouragement and constructive criticism which have made possible to bring out this project work.

I wish to express my deep sense of gratitude to **P.R.RAJESH KUMAR, M.Tech, (Ph.D)** Lecturer and **Head of the Department of Computer Science and Engineering,** for giving me the opportunity of doing the project and for providing a great support in completing my project work. I feel elated to thank him for inspiring me all the way by providing good lab facilities and helping me in providing such good environment.

I wish to convey my acknowledgment to **Dr.R.RAMACHANDRA, M.Tech., Ph.D., Principal,** SKU College of Engineering and Technology, Anantapuramu, for providing such a good environment and facilities.

My special thanks to the **Faculty of CSE Department** for giving the required information in doing my project work. Not to forget, I thank all the non-teaching staff, my friends and class mates who had directly or indirectly helped and supported me in completing my project in time.

Finally I wish to convey my gratitude to my parents who fostered all the requirements and facilities that I need.

G. Anusha	**2010115**
K. Vineeth Kumar	**2010121**
M. Hemanth Kumar Reddy	**2010125**

ABSTRACT

Monitoring and preserving water quality has become one of the most essential activities in many rural and urban areas today. The quality of water is adversely affected due to various forms of pollution caused by industry waste, sewage discharge, harmful chemicals etc.

Access to clean water to drink is crucial for good health, a fundamental human right, and an essential aspect of any health protection strategy. On a national, regional, and local level, this is significant as a development and healthcare issue. As a result, monitoring water quality now heavily relies on modeling and forecasting water quality. In this work, Machine learning algorithms are developed namely Decision tree, Random Forest algorithms for the dataset to predict the safety for human consumption. The used dataset has 10 significant parameters(pH,Hardness,Solids,Chloramines,Sulfate,Conductivity,Organic_carbon,Trihalom ethanes,Turbidity) and the developed model will be evaluated based on some statistical parameters. The goal of this project is to create a model that predicts the value of a target variable (Decision tree, Random Forest) and whether the water is potable or not.

Recently, many researchers began to use the big data analytics approach due to advancements in big data applications and availability of environmental sensing networks and data. The aim of this is to investigate various big-data and machine learning based techniques for water quality forecasting. This reviews the published research results relating to water quality evaluation using methods of artificial intelligence, decision trees, deep learning etc. Furthermore, it throws light on some of the challenges and future research needs.

KEYWORDS: Water quality evaluation, Machine Learning, data-driven water quality evaluation,and water quality prediction.

CONTENTS

LIST OF FIGURES:

LIST OF ABBREVIATIONS

S.NO	ACRONYM	ABBREVIATION
1	WQI	Water Quality Index
2	pH	Potential of Hydrogen
3	TDS	Total Dissolved Solids
4	TOC	Total Organic Carbon
5	S	Sulphate Concentration
6	THMS	Trihalomethanes
7	EC	Electrical Conductivity
8	Ca	Calcium
9	Mg	Magnesium
10	SLM	Supervised Machine Learning
11	DT	Decision Tree
12	RF	Random Forest
13	ML	Machine Learning
14	HTML	Hyper Text Markup Language
15	CSS	Cascading Style Sheets

CHAPTER 1

INTRODUCTION

All life on earth relies on water, making it one of our fundamental necessities. With a surface area of roughly 71% of the entire planet, it controls the majority of the available space. Water that is continuously extracted from the surface or the ground and used to such an extent to where it can no longer be used is sometimes referred to as water consumption and use. Water that has been contaminated by anthropogenic contaminants and is unfit for human consumption is referred to as contaminated water.

Additionally, industries contribute to this pollution Water borne diseases in aquatic organisms can be brought on by pathogens in this polluted water. Several main causes of water contamination are population growth and advancements in technology. If the present state of affairs persists, life on Earth will be untenable since there will be a tremendous demand for water and a potential shortage. As a result, maintaining water potability now heavily relies on the prediction and modeling of water quality. In this study, decision trees, random forests, and simple Bayes algorithms are built for the dataset to predict food safety for human consumption. The created model will be assessed using some statistical variables, and the employed dataset has 10 significant parameters. With the help of decision trees, a random forest this methodology tries to produce a model that can predict the value of the desired variable and whether or not the water is potable.

Each cell in the body receives its energy mostly from water, which also controls all the body's functions. 80% of the cerebrum is made up of water. Extreme dehydration may result in mental impments and a loss of the ability to clearly think. One of the most important regular resources for the survival of all species on Earth is water. Water is used for many different things, such as drinking, washing, and water systems, due to its nature. Water is essential for both living things and plants. Simply put, all organic living things require a huge quantity and exceptional quality of water to exist. Freshwater is a fundamental asset to horticulture and industry for its essential presence. Water quality observation is a key stage in the administration of freshwater assets. As indicated by the yearly report of WHO, many individuals are kicking the bucket because of the absence of unadulterated drinking water parti. It is critical to check the nature of water for its expected reason, whether it be animals watering, compound showering, or drinking water. A tool called water quality testing can be used to locate pure drinking water. This means that for the protection of pure and clean water, the proper water testing is quite important. Water testing is crucial in determining the proper operation of water sources, evaluating the safety of drinking water and deducing the measures to curb the menace

We can respond to questions like whether the water is fit for drinking, washing, or water systems, to name a few applications, by testing the nature of a water body. It can use the results of water quality tests to examine the nature of water in a location, a state, or the entire country, starting with one water body and moving on to the next. Since irresistible illnesses caused by pathogenic bacteria, infections, helminths, and other parasites are the most well-known and pervasive health danger associated with drinking water, microbiological quality is typically the most urgent issue to be addressed during this process. When certain synthetic compounds are present in drinking water in excess, health risks result. These synthetics contain nitrate, fluoride, and arsenic. To the client should be given safe drinking (consumable) water for drinking, meal preparation, personal hygiene, and cleaning. To ensure purity at the point of client supply, the water must adhere to standard quality standards

Water is the major important resource of mankind. In everyday life, people use water frequently. It is one of the most needs of human beings to avoid skin and lung diseases, we must use good-quality water. For this purpose, we have to calculate the value of the Water Quality Index of our daily usage water, Water quality assessment methods differ in their methodology as well as their input parameters.The most frequent Water Quality Index Methods are the National Sanitation Foundation Method, Oregon Water Quality Index Method, Weighted Arithmetic Water Quality Index Method, and the Canadian Council of Ministers of the Environment Water Quality Index Method CCME-WQI)...In this research paper, we adopted the Weighted Arithmetic Water Quality Index Method. We calculated the important parameters: salinity, total suspended solids (TDS), dissolved oxygen (DO), acidity and alkalinity (pH), and biochemical oxygen demand (BOD) and tabulated as a CSV file.

WQI Value	**Water Quality**	**Number of samples**	**Percentage of samples**	**Purpose of uses**
0-25	Excellent	26	38.80	Drinking and irrigation
26-50	Good	24	35.82	Drinking and irrigation
51-75	Moderate	13	19.40	Irrigation and treatment needed for before drinking
76-100	Poor	04	5.98	Need attention for irrigation
>100	Very Poor	-	-	unfit for all uses

Fig 1: WQI Value, Quality and Purpose of Use

Nowadays many problems are solved by machine learning algorithms in an efficient way. The most important algorithms are Decision Tree, Random Forest. A Machine Learning model system learns from data and builds the model. When it receives new data, it predicts the output for new data. The accuracy of predicted output depends upon the volume of data. If the volume of data is high only the model predicts the output as more accurate.

1.1 Literature Survey

In 2018, Ali Heidar Nasrolahi along with, Ca, Mg, etc such components into consideration. Performance was tallied by using several ML and DL algorithms.[1] It was observed that results of SVM was the front runner and gave the best accuracy. ANN gave acceptable accuracy for practical purposes. Amir Hamzeh Haghiabi and Abbas Parsaie predicted the Water Quality of a river bed in Iran Tireh River by taking pH, Na.

In 2019, Umair Ahmed et.al explained ways to efficiently predict water quality using supervised Machine Learning. Harrowing diseases have been in increased proportions due to the depreciation and deterioration of water quality at an alarming rate which was a direct impact of rapid urbanization and industrialisation. Their research monitors and works with supervised Machine Learning algorithms to calculate Water Quality Index (WQI) and Water Quality Class (WQC), the former being a singular index which describes the general quality of water and the latter being the derivative and distinctive class on the basis of WQI.

In 2020, Mohammed Al-Yaari et.al illustrated the use of Artificial Intelligence algorithms along with the performance of each used algorithm. As we know, for the protection of the environment, predicting and modelling of the quality of water is immensely important. In the methodology they proposed, to predict WQI, artificial intelligence algorithms, such as, NARNET and LSTM were used. Along with this, KNN, SVM and Naïve Bayes algorithms were also implemented. They used a dataset with 7 relevant and significant features and statistical parameters were used to develop the model and evaluate them.[3]

In 2020, Navideh Noori et.al explained the water quality prediction using SWAT-ANN coupled approach. For solving environmental problems Machine Learning algorithm such as Artificial Neural Networks is being used widely. They illustrated the application of SWAT-ANN for water quality prediction.[4]
In 2022, Jin-Won Yu et.al explained the use of AI algorithms for the water quality prediction. Combined the power of data decomposition, fuzzy C-means clustering and bidirectional gated recurrent model for the prediction of water quality.re

In 2022, Manisha Koranga et.al discussed the use of Machine Learning Algorithms for water quality prediction for Nanital Lake, Uttarakhand. Analysed the use of machine learning algorithms and used eight regression algorithms and nine classification algorithms. Three algorithms Random Forest, SVM and Stochastic Gradient Descent comes out to be the most effective machine learning algorithms.[6] Reviewing the literature shows that artificial intelligence techniques have been proposed for water conservation

projects for which water quality prediction and assessment plays an important role. Hence, this paper presents a designed algorithm for the prediction of Water Quality considering the concentration, pH, duration factors.

This study looks into the approaches that were used to help solve water quality challenges [1]. In most studies, traditional analyses in the laboratory and data analysis are two types of analyses and utilized to help determine the quality of water, but other studies apply machine learning approaches to help find an optimal solution to the water quality problem.

Consumers' health is being negatively impacted by poor drinking water quality. At least 2 billion individuals used feces-contaminated drinking water around the world, according to reports. Developing accurate decisions about the control and safeguarding of drinking water quality necessitates an awareness of the factors impacting its purity. Potable water quality is typically impacted by the source water's quality, how it is handled before being delivered, how it is distributed, how it is maintained, and how effectively it is filtered at residence. Furthermore, in rural areas and small municipalities, drinking water is frequently drawn straight from wells or retrieved unfiltered from rivers, lakes, and reservoirs. As a result, the purity of the source water is a significant factor affecting the quality of the drinking water. Many developing nations have achieved waterborne disease reduction and the development of safe water supplies a significant public health aim in recent years, and the situation has improved slightly. However, the situation is far from ideal, particularly in rural regions, and even marginally better conditions may be jeopardized by growing water consumption and reduced water availability as a result of population expansion and economic development. It is vital to use a practical and effective drinking water quality evaluation approach to get trustworthy results and make informed decisions.

Clark RM, Hakim ,Many water quality evaluation approaches have been proposed since Horton produced the first Water Quality Index (WQI) in the 1960s [2]. The two indices for determining the general state of drinking source water quality are straightforward, adaptable, and stable, with little sensitivity to input data. Similarly, to give water quality information, we employed the weighted arithmetic WQI approach. These WQIs convert a huge number of variables into a digital number and aid in the comprehension of water quality, making them the most widely used water quality assessment tool, despite significant flaws. Recent water quality assessments used matter element extension analysis (MEEA) and entropy TOPSIS in a wastewater irrigation area and a rapidly urbanizing area, respectively [2]. Both approaches are mathematical, but they are accurate in estimating overall water quality. These water quality evaluation methods, on the other hand, rely on water quality standards for classification.

PROBLEM IDENTIFICATION

1.2 EXISTING SYSTEM

The existing system is basically used to predict the water quality on the basis of WQI value. The system is trained by using different machine learning algorithms which gives different accuracies. In some areas, this model is predicting the water quality differently i.e., where there is highconcentration of water pollutants, it is showing that water pollutant has less concentration. Due to this,the particular people may get affected because of wrong results. The accuracy of those model liesbetween 60 to 70%.

1.2.1 DISADVANTAGES OF EXISTING SYSTEM

The model failed to give accurate results in some particular areas. The dataset which was taken had shorter period. Short period of data cannot be able predict the accurate results for future. WQI level is not known to some areas in particular city due to inadequate data. The accuracy of this model is 90%.

1.3 PROPOSED SYSTEM

The proposed system is also used to predict the water quality on the basis of WQI value.The proposed system is trained and tested by using random forest algorithm. The dataset that we have taken here contains different cities of India and covered nearly 30,000 areas in those cities. The dataset is standard and verified without any missing data. We developed model that is capable of predicting most accurate results of a particular area. The accuracy of the model ranges from 80-90%.

1.3.1 ADVANTAGES OF PROPOSED SYSTEM

This model is trained and tested to give accurate results of any area in a particular city. The dataset is standard and verified. The dataset taken is of longer period (2018-2023). The dataset is very recent one. Longer period of data is helpful to predict the best accurate results for future. The accuracy of the model is 86.76% by using random forest(Decision Tree) regressor algorithm.

1.4 ECONOMICAL FEASIBILITY

1.4.1 Feasibility Study

Preliminary investigation examines project feasibility; the likelihood the system will be useful to the organization. The main objective of the feasibility study is to test the Technical, Operational and Economical feasibility for adding new modules and debugging old running system. All systems are feasible if they are given unlimited resources and infinite time. There are aspects in the feasibility study portion of the preliminary investigation:

- Technical Feasibility
- Operation Feasibility
- Economic Feasibility

1.4.2 Technical Feasibility

The technical issue usually raised during the feasibility stage of the investigation includes the following:

- Does the necessary technology exist to do what is suggested?
- Do the proposed equipment's have the technical capacity to hold the data required to use the new system?
- Will the proposed system provide adequate response to inquiries, regardless of the number or location of users?
- Can the system be upgraded if developed?
- Are there technical guarantees of accuracy, reliability, ease of access and data security?

1.4.3 Operation Feasibility

The operational feasibility includes User friendly, reliability, security, portability, availability and maintainability of the software used in the project.

1.4.4 Economic Feasibility

Analysis of a project costs and revenue in an effort to determine whether or not it is logical and possible to complete.

1.5 REQUIREMENTS

A software requirements specification (SRS) is a description of a software system to be developed, its defined after business requirements specification (CONOPS) also called stakeholder requirements specification (STRS) other document related is the system requirements specification (SYRS). The software requirements specification (SRS) lays out functional and non-functional requirements, and may include a set of use cases that describe user interactions that the software must provide.

1.5.1 FUNCTIONAL REQUIREMENTS

Functional requirements specify which output file should be produced from the given file they describe the relationship between the input and output of the system, for each functional requirement a detailed description of all data inputs and their source and the range of valid inputs must be specified.

1.5.2 NON-FUNCTIONAL REQUIREMENTS

Describe user-visible aspects of the system that are not directly related with the functional behaviour of the system. Non-Functional requirements include quantitative constraints, such as response time (i.e. how fast the system reacts to user commands.) or accuracy. It involves the development of clear, straightforward problem statements that can be linked directly with the specific goals and objectives. These statements should clarify how the problem might prevent the achievement of these goals and objectives.

1.5.3 HARDWARE AND SOFTWARE REQUIREMENTS

All computer software needs certain hardware components or other software resources to be present on a computer. These prerequisites are known as (computer) system requirements and

are often used as a guideline as opposed to an absolute rule. Most software defines two sets of system requirements: minimum and recommended. With increasing demand for higher processing power and resources in newer versions of software, system requirements tend to increase over time. Industry analysts suggest that this trend plays a bigger part in driving upgrades to existing computer systems than technological advancements. A second meaning of the term of System requirements is a generalization of this first definition, giving the requirements to be met in the design of a system or sub-system.

Hardware Requirements

Processor : Intel i3

Ram : 4GB

Hard Disk Space : 50GB

Any desktop / Laptop system with above configuration or higher level.

Software Requirements

Operating System : Windows8/10

Coding Language : Python

Coding Platform : Jupyter Notebook

Version : Python 3.6.8

IDE : Python 3.6.8 IDLE

ML Packages : Numpy, Pandas, Sklearn, Flask

CHAPTER 2

SYSTEM DESIGN

This chapter provides information of software development life cycle, design model i.e. various UML diagrams and process specification.

2.1 DESCRIPTION

Systems design is the process or art of defining the architecture, components, modules, interfaces, and data for a system to satisfy specified requirements. One could see it as the application of systems theory to product development. There is some overlap and synergy with the disciplines of systems analysis, systems architecture and systems engineering.

The System Design Document describes the system requirements, operating environment, system and subsystem architecture, files and database design, input formats, output layouts, human-machine interfaces, detailed design, processing logic, and external interfaces.

This design activity describes the system in narrative form using non-technical terms. It should provide a high-level system architecture diagram showing a subsystem breakout of the system, if applicable. The high-level system architecture or subsystem diagrams should, if applicable, show interfaces to external systems. Supply a high-level context diagram for the system and subsystems, if applicable. Refer to the requirements trace ability matrix (RTM) in the Functional Requirements Document (FRD), to identify the allocation of the functional requirements into this design document.

This section describes any constraints in the system design (reference any trade-off analyses conducted such, as resource use versus productivity, or conflicts with other systems) and includes any assumptions made by the project team in developing the system design.

This section describes any contingencies that might arise in the design of the system that may change the development direction. Possibilities include lack of interface agreements with outside agencies or unstable architectures at the time this document is produced. Address any possible workarounds or alternative plans.

This section provides the organization code and title of the key points of contact (and alternates if appropriate) for the information system development effort. These points of contact should include the Project Manager, System Proponent, User Organization, Quality Assurance

(QA) Manager, Security Manager, and Configuration Manager, as appropriate. This section describes the organization of the Systems Design Document.

2.2 UML DIAGRAMS:

An object-oriented system is composed of objects. The behavior of the system is achieved through collaboration between these objects, and the state of the system is the combined state of all the objects in it. Collaboration between objects involves those sending messages to each other. The exact semantics of message sending between objects varies depending on what kind of system is being modeled. In some systems, "sending a message" is the same as "invoking a method".

Object Oriented Analysis aims to model the problem domain, the problem we want to solve by developing an object-oriented (OO) System. The source of the analysis is a written requirement statement, and/or written use cases, UML diagrams can be used to illustrate the statements.

The Unified Modeling Language allows the software engineer to express an analysis model using the modeling notation that is governed by a set of syntactic semantic and pragmatic rules.

A UML system is represented using five different views that describe the system from distinctly different perspective. Each view is defined by a set of diagrams, which is as follows.

Diagram: "The group of things and relationship is known as diagrams". To better understand to a model (or) a system there are different in UML each diagram provides different information about system. Every software contains structural aspects as well as behavioural aspects to represent this the diagrams.

Diagrams are categorized into two parts.

- Static/structural diagram
- Dynamic/behavioural diagram

STATIC DIAGRAMS:

2.2.1 Class diagram

2.2.2 Object diagram

2.2.3 Component diagram

DYNAMIC DIAGRAMS:

2.2.4Collaboration diagram

2.2.5 Use Case diagram

2.2.6Sequence diagram

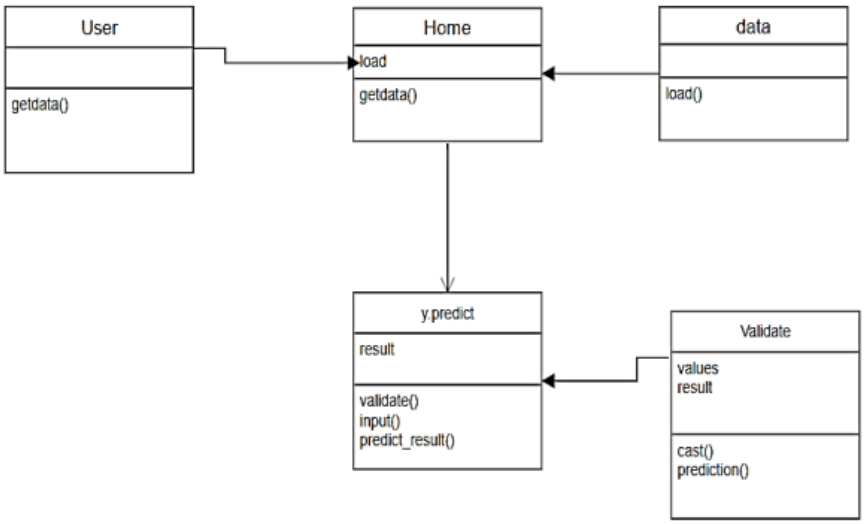
STATIC DIAGRAMS:

2.2.1 CLASS DIAGRAM

Class diagram are the blueprints of our system. You can use this diagram to model the objects they make up the system, to display the relationship between objects, and also to describe what those objects do and services that they provide. These class diagrams are useful in many stages of system design.

A Class diagram gives an overview of a system by showing its classes and the relationships among them. Class diagrams are static they display what interacts but not what happens when they do interact. The class chart delineates the attributes and operations of a class moreover the goals constrained on the structure.

FIG 2.2.1 CLASS DIAGRAM



2.2.2 OBJECT DIAGRAM

An object diagram shows the relation between instantiated classes and defined classes, andthe relation between these objects in the system. These object diagrams are useful to describe smaller portions of system when your system class diagram is complex.

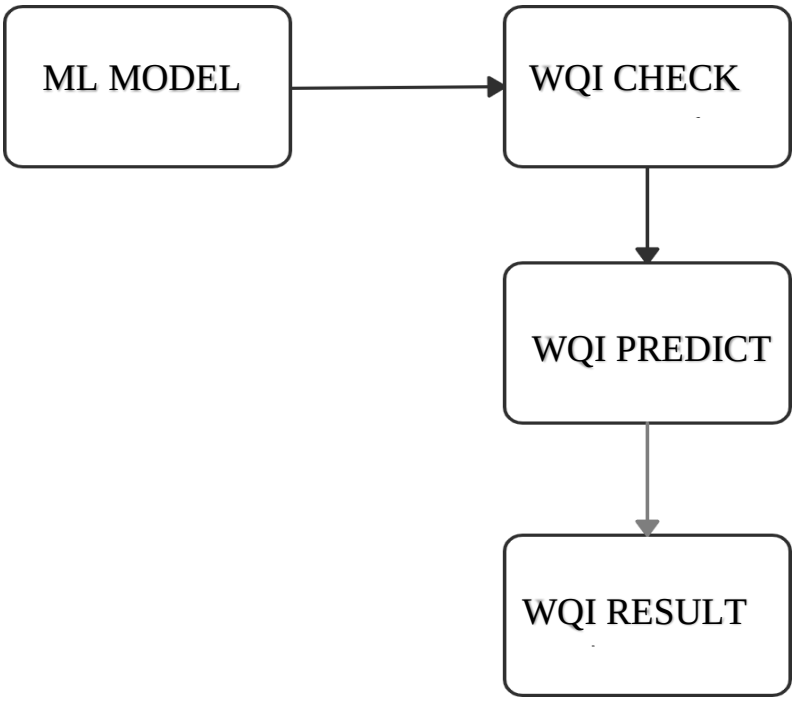


Fig. 2.2.2 OBJECT DIAGRAM

2.2.3 COMPONENT DIAGRAM

Component diagram is a special kind of diagram in UML. The purpose is also different from all other diagrams discussed so far. It does not describe the functionality of the system but it describes the components used to make those functionalities.

Thus from that point of view, component diagrams are used to visualize the physical components in a system. These components are libraries, packages, files, etc.

Component diagrams can also be described as a static implementation view of a system. Static implementation represents the organization of the components at a particular moment.

A single component diagram cannot represent the entire system but a collection of diagrams is used to represent the whole.

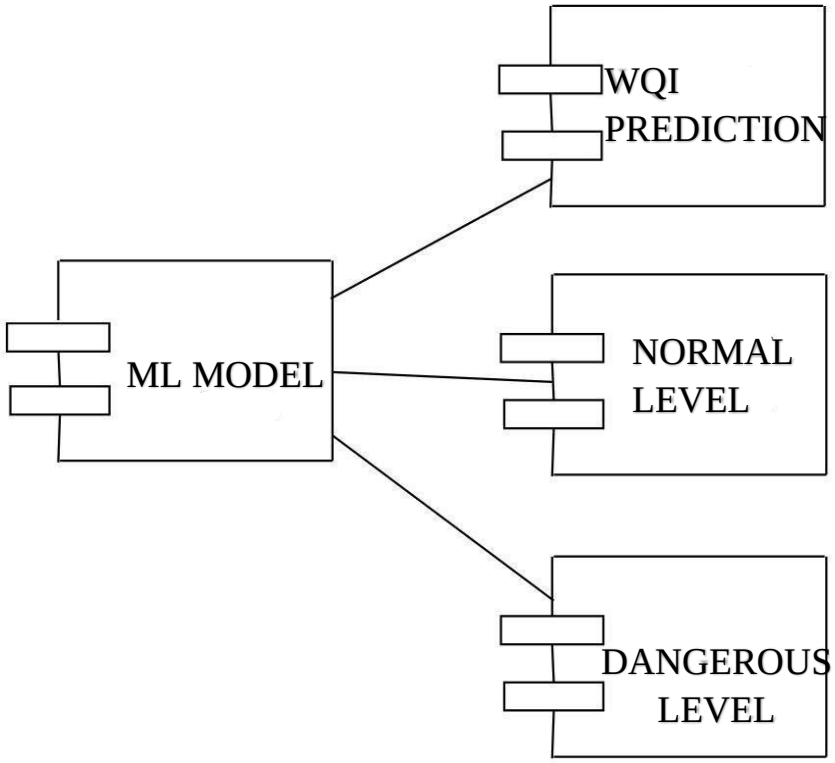


Fig. 2.2.3 COMPONENT DIAGRAM

DYNAMIC DIAGRAMS

2.2.4 SEQUENCE DIAGRAM

The sequence diagrams show the object interaction arranged in time sequence of our designed system. A sequence diagram is an introduction that empathizes the time ordering of messages. Graphically a sequence diagram is a table that shows objects arranged along the X-axis and messages ordered in increasing time along the Y-axis.

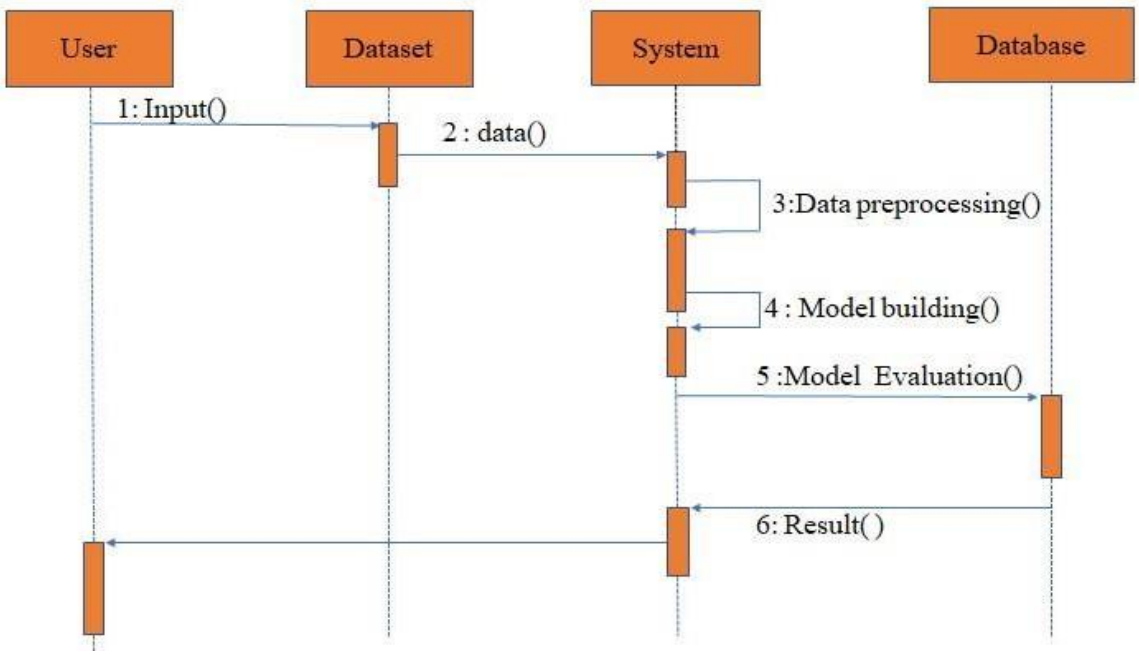


Fig. 2.2.4 SEQUENCE DIAGRAM

2.2.5 USE CASE DIAGRAM

Use case diagrams model behavior within a system and helps the developers understand of what the user require.

Use case diagram can be useful for getting an overall view of the system and clarifying who can do and more importantly what they can't do.

Use case diagram consists of use cases and actors and shows the interaction between the use case and actors.

2.2.5.1 The purpose is to show the interactions between the use case and actor.

2.2.5.2 To represent the system requirements from user's perspective.

2.2.5.3 An actor could be the end-user of the system or an external system.

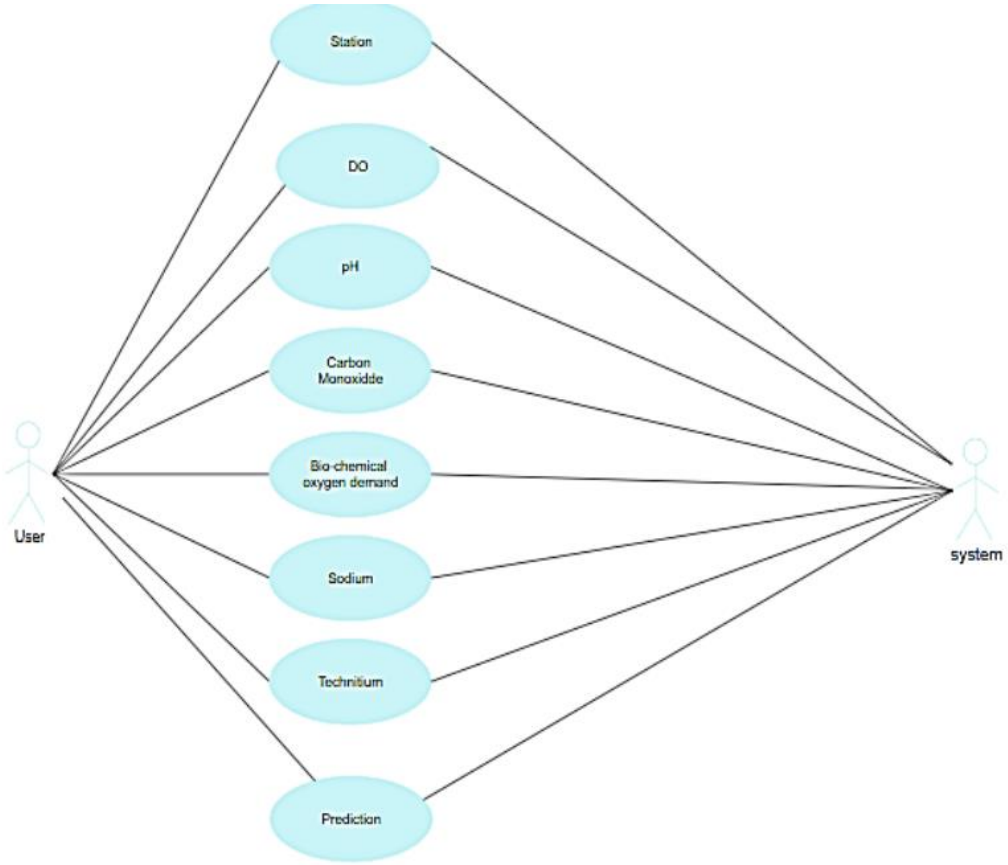


Fig. 2.2.5 USE CASE DIAGRAM

2.2.6 COLLABORATION DIAGRAM

It is also termed as communication diagram. This shows the interaction and relationship among the objects of a designed system.

Communication diagrams formerly known as collaboration diagrams are almost identical to sequence diagrams in UML, but they focus more on relationships of objects, how they associate and connect through messages in a sequence rather than interactions.

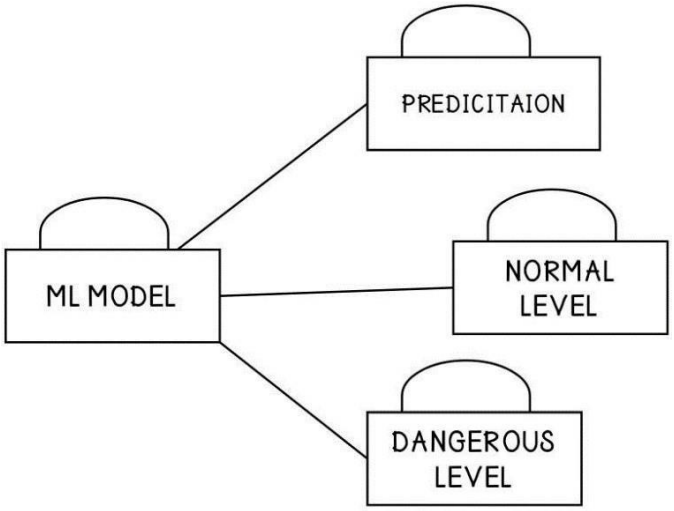


Fig. 2.2.6 COLLABORATION DIAGRAM

2.3 SYSTEM ARCHITECTURE

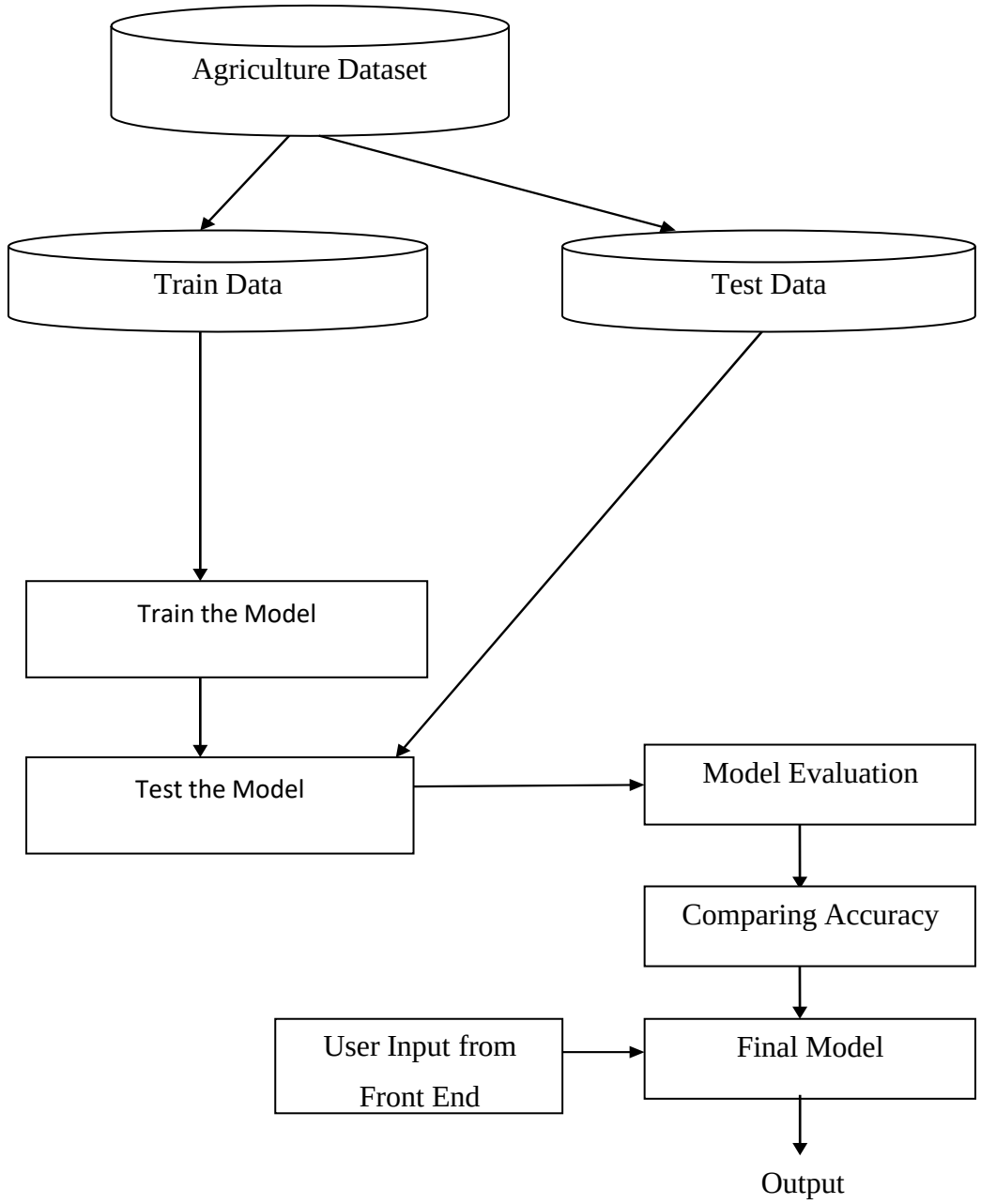


Fig. 2.3 SYSTEM ARCHITECTURE

- Water Quality dataset is taken and loaded.
- The data is pre-processed to clean the data and understand the dataset.

- The data is split as training and testing data.
- The model is built using machine learning algorithm like Decision tree .
- The model is trained with the pre-processed data.
- The model is tested and accuracy is calculated for different ML algorithms.
- The algorithm with best accuracy is finalized and that model will predict the water quality index and also based on predicted water quality index value we will display water quality is good or bad based on user given new data from front end.

2.4 DATA FLOW DIAGRAM

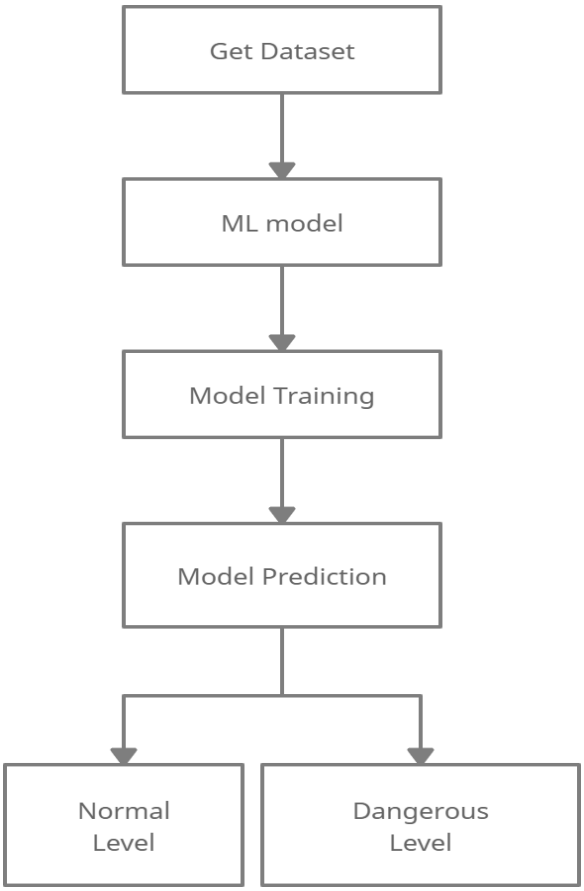


Fig. 2.4 DATA FLOW DIAGRAM

The dataset is read and next pre-processing is done to clean the data and to remove the nullvalues if any present in the dataset. Required machine learning algorithms are applied on the model to finally predict the given water quality level is satisfactory or poisonous.

CHAPTER 3

IMPLEMENTATION

Implementation is the stage of the project when the theoretical design is turned out into a working system. Those it can be considered to be the most critical stage in achieving a successful new system and in giving the user, confidence that the new system will work and be effective. The implementation stage involves careful planning, investigation of the existing system and its constraints on implementations, designing of methods to achieve change over and evaluation of change over methods.

An Implementation is a realization of a technical specification or algorithm as a program, software components, or other computer system though computer programming and deployment. Many implementations may exist for specifications or standards. A special case occurs in object-oriented programming, when a concrete class implements an interface.

Implementation literally means to put into effect or to carry out. In the system implementation phase, the software deals with translation of the design specifications into source code. The ultimate goal of the implementation is to write the source code and the internal documentation so that it can be verified easily. The code and documentation should be written in a manner that eases debugging, testing and modification. System flow charts, simple run on packages, sample output etc., is part of the implementation. This includes requirement analysis scope analysis, customization, system integrations, user policies, and user training. An effort was made to satisfy the following goals in order:

- Clarity and simplicity of the code
- Minimization of Hard Coding
- Minimization of the amount of memory used
- Thorough phased implementation has been done so that we can use our proposed system correctly.

3.1MODULES

3.1.1 DATACOLLECTION

Collecting data for training the ML model is the basic step in the machine learning pipeline. The predictions made by ML systems can only be as good as the data on which they have been trained. Following are some of the problems that can arise in data collection:

- Inaccurate data. The collected data could be unrelated to the problem statement.
- Missing data. Sub-data could be missing. That could take the form of empty values in columns or missing images for some class of prediction.
- Data imbalance. Some classes or categories in the data may have a disproportionately high or low number of corresponding samples. As a result, they risk being under-represented in the model.
- Data bias. Depending on how the data, subjects and labels themselves are chosen, the model could propagate inherent biases on gender, politics, age or region, for example. Data bias is difficult to detect and remove.

The experimental data is collected from Kaggle.

3.1.2 DATA PRE-PROCESSING

Real-world raw data and images are often incomplete, inconsistent and lacking in certain behaviours or trends. They are also likely to contain many errors. So, once collected, they are pre-processed into a format the machine learning algorithm can use for the model.

Pre-processing includes a number of techniques and actions:

Data cleaning. These techniques, manual and automated, remove data incorrectly added or classified.

Data imputations. Most ML frameworks include methods and APIs for balancing or filling in missing data. Techniques generally include imputing missing values with standard deviation, mean, median of the data in the given field.

Oversampling. Bias or imbalance in the dataset can be corrected by generating more observations/samples with methods like repetition, bootstrapping or and then adding them to the under-represented classes.

Data integration. Combining multiple datasets to get a large corpus can overcome incompleteness in a single dataset.

Data normalization. The size of a dataset affects the memory and processing required for iterations during training. Normalization reduces the size by reducing the order andmagnitude of data.

3.1.3CALCULATING WQI BUCKET

Calculating Sub-Index for each Pollutant & Filling the Null values of WQI column by taking maximum values out of Sub-indexes. Calculating WQI bucket and filling the Null value present in the WQI bucket.

3.1.4 DATA ANALYSIS AND VISUALIZATION

Data visualization is an essential specialist knowledge in applied statistics and the statistics of machine learning rely practically on quantitative explanations, and software analyses are valuable instruments for obtaining qualitative understanding. This helps to discover and know about a dataset and help to identify trends in fraudulent data transfers and more. Information visualizations are feasible with low domain knowledge to communicate and display core connections through graphs and maps which are more emotional and consumers than associative or substantive steps.

Evaluation of water quality and Average WQI in last 5 years.

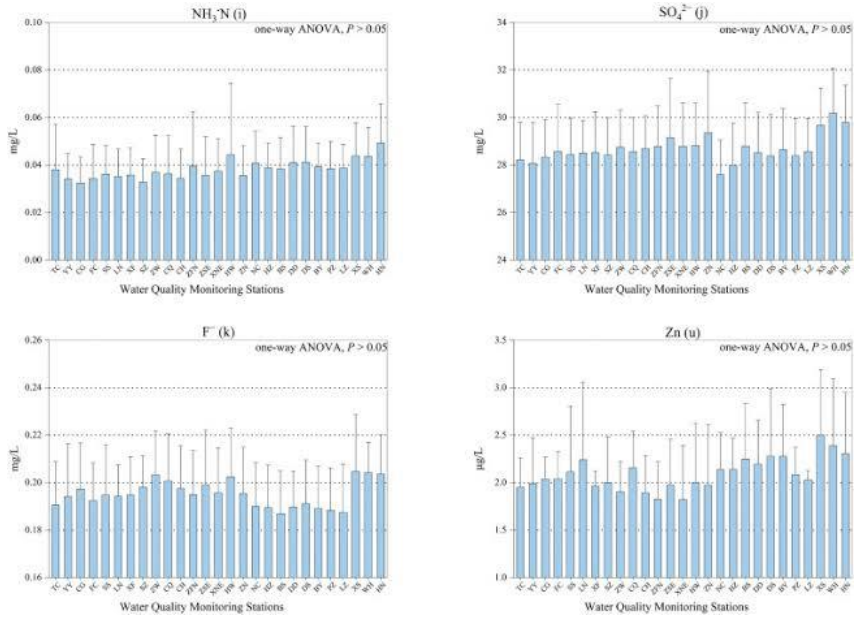


Figure 3.1.4(a): Evaluation Of Water Quality

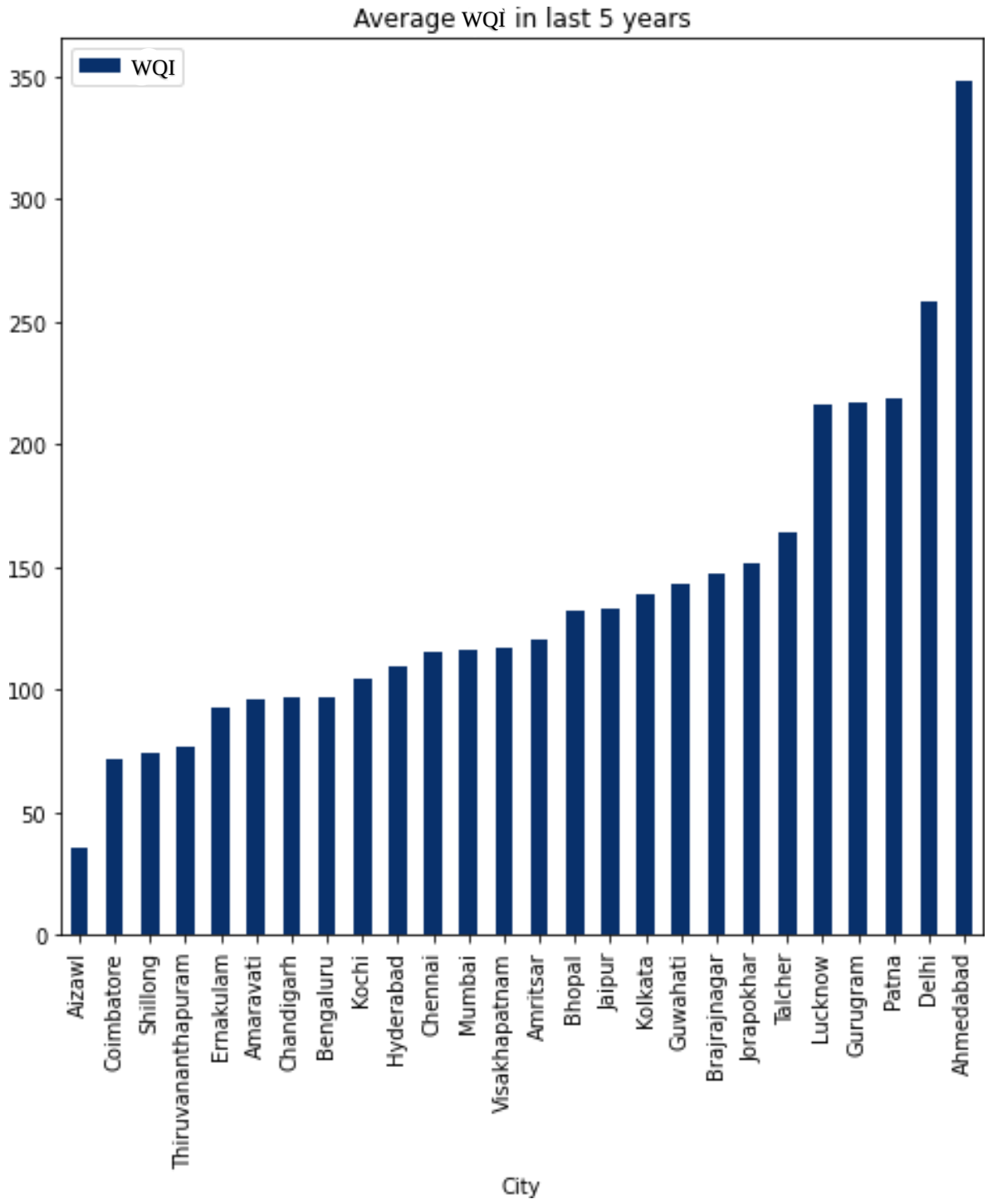


Figure 3.1.4(b): Average WQI in last 5 years

3.1.5 PREDICTING THE OUTPUT

- The collected data from the user is given as input to our finalized algorithm to predict the Water Quality Index.
- It takes the input values for each pollutant from the user, then it calculates the WQI value.
- Then it predicts the WQI bucket, and the WQI value.
- Finally, the predicted Water Quality Index output is displayed on the front endand also displays the Water quality is Good or Bad based on predicted the water quality index value is the front end.

3.2 TECHNOLOGIES USED

The following are the technologies used in the project:

3.2.1 HYPER TEXT MARKUP LANGUAGE

Hyper Text Markup Language (HTML), the languages of the World Wide Web (WWW), allows users to produces Web pages that include text, graphics and pointer to other Web pages (Hyperlinks).HTML is not a programming language but it is an application of ISO Standard 8879, SGML (Standard Generalized Markup Language), but specialized to hypertext and adapted to the Web. The idea behind Hypertext is that instead of reading text in rigid linear structure, we can easily jump from one point to another point.

We can navigate through the information based on our interest and preference. A mark-up language is simply a series of elements, each delimited with special characters that define how text or other items enclosed within the elements should be displayed. Hyperlinks are underlined or emphasized works that load to other documents or some portions of the same document.

It is a versatile language and can be used on any platform or desktop.HTML provides tags (special codes) to make the document look attractive. HTML tags are not case sensitive. Using graphics, fonts, different sizes, color, etc., can enhance the presentation of the document.

Anything that is not a tag is part of the document itself.

FEATURES OF HTML:

1) It is a very easy and simple language. It can be easily understood and modified.

2) It is very easy to make an effective presentation with HTML because it has a lot of formatting tags.

3) It is a markup language, so it provides a flexible way to design web pages along with the text.

4) It facilitates programmers to add a link on the web pages (by html anchor tag), so it enhances the interest of browsing of the user.

5) It is platform-independent because it can be displayed on any platform like Windows, Linux, and Macintosh, etc.

6) It facilitates the programmer to add Graphics, Videos, and Sound to the web pages which makes it more attractive and interactive.

7) HTML is a case-insensitive language, which means we can use tags either in lower-case or upper-case.

3.2.2 CSS:

Cascading Style Sheets, fondly referred to as CSS, is a simple design language intended to simplify the process of making web pages presentable. CSS is a MUST for students and working professionals to become a great Software Engineer especially when they are working in Web Development Domain. I will list down some of the key advantages of learning CSS.

Create Stunning Web site - CSS handles the look and feel part of a web page. Using CSS, you can control the colour of the text, the style of fonts, the spacing between paragraphs, how columns are sized and laid out, what background images or colors are used, layout designs, and variations in display for different devices and screen sizes as well as a variety of other effects. Become a web designer - If you want to start a career as a professional web designer, HTML and CSS designing is a must skill. Control web - CSS is easy to learn and understand but it provides powerful control over the presentation of an HTML document. Most commonly, CSS is combined with the markup languages HTML or XHTML. Learn other languages - Once you understand the

basic of HTML and CSS then other related technologies like JavaScript, php, or angular are become easier to understand.

APPLICATIONS:

1. CSS is one of the most widely used style language over the web. I'm going to list few of them here:
2. CSS saves time - You can write CSS once and then reuse same sheet in multiple HTML pages. You can define a style for each HTML element and apply it to as many Web pages as you want.
3. Pages load faster - If you are using CSS, you do not need to write HTML tag attributes every time.
4. Just write one CSS rule of a tag and apply it to all the occurrences of that tag. So, less code means faster download times.
5. Easy maintenance - To make a global change, simply change the style, and all elements in all the web pages will be updated automatically.
6. Superior styles to HTML - CSS has a much wider array of attributes than HTML, so you can give a far better look to your HTML page in comparison to HTML attributes.
7. Multiple Device Compatibility - Style sheets allow content to be optimized for more than one type of device. By using the same HTML document, different versions of a website can be presented for handheld devices such as PDAs and cell phones or for printing.
8. Global web standards - Now HTML attributes are being deprecated and it is being recommended to use CSS. So, it’s a good idea to start using CSS in all the HTML pages to make them compatible to future browsers.

3.2.3 FLASK

Flask (source code) is a Python web framework built with a small core and easy-to-extend philosophy. Flask is considered more Pythonic than the Django web framework because in common situations the equivalent Flask web application is more explicit. Flask is also easy to get started with as a beginner because there is little boilerplate code for getting a simple app

up and running. Flask is a popular Python web framework, meaning it is a third-party Python library used for developing web applications.

Example:

```
from flask import Flask app = Flask(__name__)

    @app.route('/') def hello_world():
    return 'Hello World'

    if__name__== '_main_':
    app.run()
```

Importing flask module in the project is mandatory. An object of Flask class is our **WSGI** application.

Flask constructor takes the name of **current module (__name__)** as argument.

The **route()** function of the Flask class is a decorator, which tells the application which URL should call the associated function.

app.route(rule, options)

- The **rule** parameter represents URL binding with the function.
- The **options** is a list of parameters to be forwarded to the underlying Rule object.

In the above example, **'/'** URL is bound with **hello_world()** function. Hence, when the home page of web server is opened in browser, the output of this function will be rendered.

Finally the **run()** method of Flask class runs the application on the local development server.

app.run(host, port, debug, options)

All parameters are optional

Sr.No.	Parameters & Description
1	**Host** Hostname to listen on. Defaults to 127.0.0.1 (localhost). Set to '0.0.0.0' to have server available externally
2	**Port** Defaults to 5000
3	**Debug** Defaults to false. If set to true, provides a debug information
4	**Options**
	To be forwarded to underlying Werkzeug server.

The above given **Python** script is executed from Python shell.

Python Run.py

A message in Python shell informs you that

* Running on http://127.0.0.1:5000/ (Press CTRL+C to quit)

Open the above URL **(localhost:5000)** in the browser. **'Water Quality'** message will be displayed on it.

Debug mode

A **Flask** application is started by calling the **run()** method. However, while the application is under development, it should be restarted manually for each change in the code. To avoid this inconvenience, enable **debug support**. The server will then reload itself if the code changes. It will also provide a useful debugger to track the errors if any, in the application.

The **Debug** mode is enable by setting the **debug** property of the **application** object to **True** before running or passing the debug parameter to the **run()** method.

```
app.debug = True app.run()
app.run(debug = True)
```

Flask – Routing

Modern web frameworks use the routing technique to help a user remember application URLs. It is useful to access the desired page directly without having to navigate from the home page.

The **route()** decorator in Flask is used to bind URL to a function. For example −

```
@app.route('/hello') def
hello_world(): return 'hello
world'
```

Here, URL **'/hello'** rule is bound to the **hello_world()** function. As a result, if a user visits **http://localhost:5000/hello** URL, the output of the **hello_world()** function will be rendered in the browser.

The **add_url_rule()** function of an application object is also available to bind a URL with a function as in the above example, **route()** is used.

Flask – HTTP methods

Http protocol is the foundation of data communication in world wide web. Different methods of data retrieval from specified URL are defined in this protocol.

The following table summarizes different http methods –

Sr.No.	Methods & Description
1	**GET**
	Sends data in unencrypted form to the server. Most common method.
2	**HEAD** Same as GET, but without response body
3	**POST** Used to send HTML form data to server. Data received by POST method is not cached by server.
4	**PUT** Replaces all current representations of the target resource with the uploaded content.
5	**DELETE** Removes all current representations of the target resource given by a URL

By default, the Flask route responds to the **GET** requests. However, this preference can be altered by providing methods argument to **route()** decorator.

In order to demonstrate the use of **POST** method in URL routing, first let us create an HTML form and use the **POST** method to send form data to a URL.

Flask – Templates

It is possible to return the output of a function bound to a certain URL in the form of HTML. For instance, in the following script, **hello()** function will render **'Hello World'** with **<h1>** tag attached to it.

```
from flask import Flask
```

```
app = Flask(__name__)


@app.route('/')      def
index():
   return '<html><body><h1>Hello World</h1></body></html>'


if__name__== '_main_':
   app.run(debug = True)
```

However, generating HTML content from Python code is cumbersome, especially when variable data and Python language elements like conditionals or loops need to be put. This would require frequent escaping from HTML.

This is where one can take advantage of **Jinja2** template engine, on which Flask is based. Instead of returning hardcode HTML from the function, a HTML file can be rendered by the **render_template()** function.

```
from flask import Flask app = Flask(_name__)


@app.route('/') def index():
   return render_template(‘hello.html’)


if__name__== '_main_':
   app.run(debug = True)
```

Flask will try to find the HTML file in the templates folder, in the same folder in which this script is present.

The term **‘web templating system’** refers to designing an HTML script in which the variable data can be inserted dynamically. A web template system comprises of a template engine, some kind of data source and a template processor.

Flask uses **Jinja2** template engine. A web template contains HTML syntax interspersed placeholders for variables and expressions (in these case Python expressions) which are replaced values when the template is rendered.

The following code is saved as **hello.html** in the templates folder.

```
<!doctype html>
<html>
  <body>

    <h1>Hello {{ name }}!</h1>

  </body>
```

```
</html>
```

Next, run the following script from Python shell.

```
from flask import Flask, render_template app =
Flask(__name__)


@app.route('/hello/<user>')         def
hello_name(user):
   return render_template('hello.html', name = user)



if__name__== '__main__':
   app.run(debug = True)
```

As the development server starts running, open the browser and enter URL as

− **http://localhost:5000/hello/mvl**

The **variable** part of URL is inserted at **{{ name }}** place holder.

Flask – Sending Form Data to Template

We have already seen that the http method can be specified in URL rule. The **Form** data received by the triggered function can collect it in the form of a dictionary object and forward it to a template to render it on a corresponding web page.

In the following example, **'/'** URL renders a web page (student.html) which has a form. The data filled in it is posted to the **'/result'** URL which triggers the **result()** function.

The **results()** function collects form data present in **request.form** in a dictionary object and sends it for rendering to **result.html**.

The template dynamically renders an HTML table of **form** data.

3.2.4 MACHINE LEARNING

- Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed.
- Machine Learning is the science of getting computers to learn and act like humans do,
- Machine learning is like farming or gardening. Seeds is the algorithms, nutrients is the data, the gardner is you and plants is the programs.

Applications of Machine learning

- Web search: ranking page based on what you are most likely to click on.
- Computational biology: rational design drugs in the computer based on past experiments.
- Finance: decide who to send what credit card offers to. Evaluation of risk on credit offers. How to decide where to invest money.
- Robotics: how to handle uncertainty in new environments. Autonomous. Self-driving car.
- E-commerce: Predicting customer churn. Whether or not a transaction is fraudulent.
- Space exploration: space probes and radio astronomy.
- Information extraction: Ask questions over databases across the web.
- Social networks: Data on relationships and preferences. Machine learning to extract value from data.
- Debugging: Use in computer science problems like debugging. Labor intensive process. Could suggest where the bug could be.
- Machine learning algorithms also help to improve user experience and customization for online platforms.
- Facebook, Netflix, Google, and Amazon all use recommendation systems to prevent content glut and provide unique content to individual users based on their likes and dislikes.
- Facebook utilizes recommendation engines for its news feed on both Facebook and Instagram, as well as for its advertising services to find relevant leads.

- Netflix collects user data and recommends various movies and series based on the preferences of the user. Google utilizes machine learning to structure its results and for YouTube's recommendation system, among many other applications.
- Amazon uses ML to place relevant products in the user's field of view, maximizing conversion rates by recommending products that the user actually wants to buy.

Key elements of machine learning

1. **Representation:** how to represent knowledge.

    a. Examples include decision trees, sets of rules, instances, graphical models, neural networks, support vector machines, model ensembles and others.

2. **Evaluation:** the way to evaluate candidate programs (hypotheses).

    a. Examples include accuracy, prediction and recall, squared error, likelihood, posterior probability, cost, margin, entropy k-L divergence and others.

Types of Machine learning

There are four types of machine learning:

- **Supervised learning**: (also called inductive learning) Training data includes desired outputs. This is spam this is not, learning is supervised.
- **Unsupervised learning:** Training data does not include desired outputs. Example is clustering. It is hard to tell what is good learning and what is not.
- **Semi-supervised learning:** Training data includes a few desired outputs.
- **Reinforcement learning:** Rewards from a sequence of actions. It is the most ambitious type of learning.

Machine Learning Packages:

Numpy:

- NumPy is the fundamental package for scientific computing in Python.

Features of Numpy:

- A powerful N-dimensional array object.
- Sophisticated (broadcasting) functions.
- Tools for integrating C/C++ and Fortran code.
- Useful linear algebra, Fourier transform, and random number capabilities.

Applications of Numpy

- Numpy provides a high-performance multidimensional array and basic tools to compute with and manipulate these arrays.
- It is an alternative for lists and arrays in Python.
- It maintains minimal memory.
- We can perform different types of mathematical operations.
- Shape Manipulations.
- It is used with other libraries like Pandas, SciPy, Matplotlib and Tkinter etc.

Pandas:

- Pandas is a Python package that provides fast, flexible, and expressive data structures.
- Pandas is a fast, powerful, flexible and easy to use open source data analysis and manipulation tool, built on top of the Python programming language.
- It aims to be the fundamental high-level building block for doing practical, real world data analysis in Python.

Features of Pandas

- Dataframe object for data manipulation with integrated indexing.
- Tools for reading and writing data between in-memory data structures and different file formats.
- Data alignment and integrated handling of missing data.
- Reshaping and pivoting of data sets.
- Label-based slicing, fancy indexing, and subsetting of large data sets.
- Data structure column insertion and deletion**.**
- Group by engine allowing split-apply-combine operations on data sets.
- Data set merging and joining.
- Hierarchical axis indexing to work with high-dimensional data in a lower-dimensional data structure.
- Time series-functionality: Date range generation and frequency conversion, moving window statistics, moving window linear regressions, date shifting and lagging.
- Provides data filtration.

Pandas deals with the following three data structures

1. Series

- Series is a one-dimensional array like structure with homogeneous data.
- For example, the following series is a collection of integers 10, 23, 56,….

10	23	56	17	52	61	73	90	26	72

2. DataFrame

- DataFrame is a two-dimensional array with heterogeneous data.
- For example, the following table show the data is represented in rows and columns. Each column represents an attribute and each row represents a person**.**

Name	Age	Gender	Rating
Steve	32	Male	3.45
Lia	28	Female	4.6
Vin	45	Male	3.9
Katie	38	Female	2.78

3. Panel

- Panel is a three-dimensional data structure with heterogeneous data.
- It is hard to represent the panel in graphical representation. But a panel can be illustrated as a container of DataFrame**.**

Applications

- Economics
- Recommendation Systems
- Stock Prediction.
- Neuroscience.
- Statistics.
- Advertising.
- Analytics.
- Natural Language Processing.
- Big Data
- Data Science

SkLearn:

- Scikit-learn (SkLearn) is the most useful and robust library for machinelearning in Python.
- It provides a selection of efficient tools for machine learning and statistical modelling.
- It includes classification, regression, clustering and dimensionality reduction via a consistence interface in Python.
- This library, which is largely written in Python, is built upon NumPy, SciPy and Matplotlib.

Features of Sklearn

- Datasets
- Feature extraction
- Feature selection

- Cross-Validation
- Supervised Models
- Unsupervised Models
- Dimensionality Reduction
- Ensemble methods

Benifits of SciKit-Learn

- BSD license: Scikit-learn has a BSD license; hence, there is minimal restriction on the use and distribution of the software, making it free to use for everyone.
- Easy to use: The popularity of Scikit-learn is because of the ease of use it offers.
- Document detailing: It also offers document detailing of the API that users can access at any time on the website, helping them integrate Machine Learning into their own platforms.
- Extensive use in the industry: Scikit-learn is used extensively by various organizations to predict consumer behaviour, identify suspicious activities, and much more.
- Machine Learning algorithms: Scikit-learn covers most of the Machine Learning algorithms Huge community support.
- Algorithms flowchart: Unlike other programming languages where users usually face a problem of having to choose from multiple competing implementations of same algorithms, Scikit-learn has an algorithm cheat sheet or flowchart to assist the users.

Applications

- Scikit-learn is a library that contains several implementations of machine learning algorithms.
- financial cyber security analytics, product development, neuro imaging, barcode scanner development, medical modelling.
- Regression modelling
- Decision tree pruning and induction
- Comprehensive and neural network training with regression and

classification algorithms

- Decision boundary learning with SVMs
- Advanced probability modelling
- Feature analysis and selection
- Reduction of dimensionality
- Outlier detection and rejection

DECISION TREE:

- Decision Tree is a **Supervised learning technique** that can be used for both classification and Regression problems, but mostly it is preferred for solving Classification problems. It is a tree-structured classifier, where **internal nodes represent the features of a dataset, branches represent the decision rules** and **each leaf node represents the outcome.**
- In a Decision tree, there are two nodes, which are the **Decision Node** and **Leaf Node.** Decision nodes are used to make any decision and have multiple branches, whereas Leaf nodes are the output of those decisions and do not contain any further branches.
- The decisions or the test are performed on the basis of features of the given dataset.
- ***It is a graphical representation for getting all the possible solutions to a problem/decision based on given conditions.***
- It is called a decision tree because, similar to a tree, it starts with the root node, which expands on further branches and constructs a tree-like structure.
- In order to build a tree, we use the **CART algorithm,** which stands for **Classification and Regression Tree algorithm.**
- A decision tree simply asks a question, and based on the answer (Yes/No), it further split the tree into subtrees.

Decision Tree Terminologies:

- **Root Node:** Root node is from where the decision tree starts. It represents the entire dataset, which further gets divided into two or more homogeneous sets.
- **Leaf Node:** Leaf nodes are the final output node, and the tree cannot be segregated further after getting a leaf node.
- **Splitting:** Splitting is the process of dividing the decision node/root node into sub-nodes according to the given conditions.
- **Branch/Sub Tree:** A tree formed by splitting the tree.

- **Pruning:** Pruning is the process of removing the unwanted branches from the tree.
- **Parent/Child node:** The root node of the tree is called the parent node, and other nodes are called the child nodes.

Advantages of the Decision Tree:

- It is simple to understand as it follows the same process which a human follow while making any decision in real-life.
- It can be very useful for solving decision-related problems.
- It helps to think about all the possible outcomes for a problem.
- There is less requirement of data cleaning compared to other algorithms.

3.3 Coding

3.3.1 Machine Learning Code

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
df = pd.read_csv('water_potability.csv')
df.head()
df.shape
df.isnull().sum()
df.info()
df.describe()
df.fillna(df.mean(), inplace=True)
df.isnull().sum()
df.Potability.value_counts()
sns.countplot(df['Potability'])
plt.show()
sns.distplot(df['ph'])
plt.show()
df.hist(figsize=(14,14))
plt.show()
plt.figure(figsize=(13,8))
sns.heatmap(df.corr(),annot=True,cmap='terrain')
plt.show()
df.boxplot(figsize=(14,7))
X = df.drop('Potability',axis=1)
Y= df['Potability']
from sklearn.model_selection import train_test_split
X_train, X_test, Y_train, Y_test = train_test_split(X,Y, test_size= 0.2,
random_state=101,shuffle=True)
from sklearn.tree import DecisionTreeClassifier
from sklearn.metrics import accuracy_score,confusion_matrix,classification_report
dt=DecisionTreeClassifier(criterion= 'gini', min_samples_split= 10, splitter= 'best')
dt.fit(X_train,Y_train)
```

```
prediction=dt.predict(X_test)
print(f"Accuracy Score = {accuracy_score(Y_test,prediction)*100}")
print(f"Confusion Matrix =\n {confusion_matrix(Y_test,prediction)}")
print(f"Classification Report =\n {classification_report(Y_test,prediction)}")
res = dt.predict([[5.735724,
158.318741,25363.016594,7.728601,377.543291,568.304671,13.626624,75.952337,4.732954]])[
0]
res
from sklearn.model_selection import RepeatedStratifiedKFold
from sklearn.model_selection import GridSearchCV

# define models and parameters
model = DecisionTreeClassifier()
criterion = ["gini", "entropy"]
splitter = ["best", "random"]
min_samples_split = [2,4,6,8,10,12,14]

# define grid search
grid = dict(splitter=splitter, criterion=criterion, min_samples_split=min_samples_split)
cv = RepeatedStratifiedKFold(n_splits=10, n_repeats=3, random_state=1)
grid_search_dt = GridSearchCV(estimator=model, param_grid=grid, n_jobs=-1, cv=cv,
                    scoring='accuracy',error_score=0)
grid_search_dt.fit(X_train, Y_train)
print(f"Best: {grid_search_dt.best_score_:.3f} using {grid_search_dt.best_params_}")
means = grid_search_dt.cv_results_['mean_test_score']
stds = grid_search_dt.cv_results_['std_test_score']
params = grid_search_dt.cv_results_['params']

for mean, stdev, param in zip(means, stds, params):
   print(f"{mean:.3f} ({stdev:.3f}) with: {param}")

print("Training Score:",grid_search_dt.score(X_train, Y_train)*100)
print("Testing Score:", grid_search_dt.score(X_test, Y_test)*100)
```

3.3.2 Html & CSS Code

```
<!doctype html>
<html lang="en">
  <head>
    <!-- Required meta tags -->
    <meta charset="utf-8">
    <meta name="viewport" content="width=device-width, initial-scale=1">

    <!-- JQUERY -->
    <script src="https://cdnjs.cloudflare.com/ajax/libs/jquery/3.6.0/jquery.js"></script>

    <!-- Bootstrap CSS -->
    <link href="https://cdn.jsdelivr.net/npm/bootstrap@5.1.3/dist/css/bootstrap.min.css" rel="stylesheet"
integrity="sha384-
1BmE4kWBq78iYhFldvKuhfTAU6auU8tT94WrHftjDbrCEXSU1oBoqyl2QvZ6jIW3"
crossorigin="anonymous">
    <!-- Main CSS -->
    <link href="../static/css/main.css" rel="stylesheet">

    <title>PROJECT!</title>
  </head>
  <body>

    <div class="container">
       <div class="heading text-center fw-light">Water Quality <span class="fw-
bold">Prediction.</span></div>
       <div class="row justify-content-center pt-5">
          <div class="col-md-10 col-lg-8">
             <div class="mx-auto mt-3" style="max-width:400px;">

                <div class="mb-4 p-3 border-purple rounded">
                   <div class="form-group">
                      <label class="form-label" for="ph">pH</label>


                            <input class="form-control" id="ph" name="ph" required="" type="text" value=""
placeholder="Ideal: 7">


                   </div>

                   <div class="form-group">
                      <label class="form-label" for="temp">Temperature</label>


                            <input class="form-control" id="temp" name="temp" required="" type="text"
value="" placeholder="Ideal: 76">

                   </div>

                   <div class="form-group">
```

```
                <label class="form-label" for="turbidity">Turbidity</label>


                    <input class="form-control" id="turbidity" name="turbidity" required=""
type="text" value="" placeholder="Ideal: 4">

                </div>

                <div class="form-group">
                  <label class="form-label" for="hardness">Hardness</label>


                    <input class="form-control" id="hardness" name="hardness" required=""
type="text" value="" placeholder="Ideal: 264">

                </div>

                <div class="form-group">
                  <label class="form-label" for="hardness">Valve Pressure</label>


                    <input class="form-control" id="valvepressure" name="valvepressure"
required="" type="text" value="" placeholder="Ideal: 300">

                </div>
                <div class="d-flex">

                </div>
                <div class="form-group">
                  <label class="form-label" for="hardness">Day</label>


                    <input class="form-control" id="day" name="day" required="" type="text"
value="" placeholder="Range: 0-6">

                </div>
                <div class="form-group">
                  <label class="form-label" for="hardness">Time (hr)</label>


                    <input class="form-control" id="time" name="time" required="" type="text"
value="" placeholder="Ideal: 0-23">

                </div>
                <div class="form-group">
                  <label class="form-label" for="hardness">Flow Rate (L/s)</label>

                    <input class="form-control" id="fr" name="fr" required="" type="text" value=""
placeholder="Range: 0.0-5.0">

                </div>
                <div class="form-group">
```

```
                    <label class="form-label" for="hardness">Water Pressure (atm)</label>

                        <input class="form-control" id="wp" name="wp" required="" type="text"
value="" placeholder="Range: 90.0-400.0">

                </div>
                <p class="text-muted text-end mt-n2">
                    <a class="form-link" href="#">Browse Code Here</a>
                </p>

                <div class="text-center" id="predict" onclick="prediction()">
                    <button class="btn btn-primary btn-lg w-100 px-5 mt-2" > Predict Now!</button>
                </div>
            </div>

            <div class="d-flex text-muted small w-100 mx-auto mb-4">
                <span class="flex-fill border-top-purple" style="margin-top: 1.1rem;"></span>
                <p class="my-1 px-2">The Prediction Will Appear Here</p>
                <span class="flex-fill border-top-purple" style="margin-top: 1.1rem;"></span>
            </div>

            <div class="text-center alert alert-dark px-5 mt-2 py-2" id="predictor">
                <a class="alert-link" href="#">NOT TESTED</a>
            </div>

        </div>
    </div>

  </div>
</div>

<!-- Predict JS -->
<script>
    function prediction(){
        ph = document.getElementById('ph').value;
        temp = document.getElementById('temp').value;
        turbidity = document.getElementById('turbidity').value;
        hardness = document.getElementById('hardness').value;
        day = document.getElementById('day').value;
        time = document.getElementById('time').value;
        fr = document.getElementById('fr').value;
        wp = document.getElementById('wp').value;

        $.ajax(
            {
                type: "GET",
                url:
"/predict?ph="+ph+"&temperature="+temp+"&turbidity="+turbidity+"&hardness="+hardness+"&day="
+day+"&time="+time+"&fr="+fr+"&wp="+wp,
                success: function(updated_data){
                    update = document.getElementById('predictor')
                    if (updated_data == "[0] 1"){
```

```
                update.classList.remove('alert-dark')
                update.classList.remove('alert-danger')
                update.classList.remove('alert-success')
                update.classList.add('alert-danger')



                $("#predictor").html('<a class="alert-link" href="#">Water quality is BAD <br> </a>')
            }
            if (updated_data == "[0] 0"){

                update.classList.remove('alert-dark')
                update.classList.remove('alert-danger')
                update.classList.remove('alert-success')
                update.classList.add('alert-danger')

            $("#predictor").html('<a class="alert-link" href="#">Water quality is BAD <br></a>')
        }
            if (updated_data == "[1] 1"){

                update.classList.remove('alert-dark')
                update.classList.remove('alert-danger')
                update.classList.remove('alert-success')
                update.classList.add('alert-danger')

            $("#predictor").html('<a class="alert-link" href="#">Water quality is GOOD <br></a>')
            }
            if (updated_data == "[1] 0"){

                update.classList.remove('alert-dark')
                update.classList.remove('alert-danger')
                update.classList.remove('alert-success')
                update.classList.add('alert-success')

            $("#predictor").html('<a class="alert-link" href="#">Water quality is GOOD <br></a>')
            }

        }
      }
    )
  }
</script>

<!-- SCRIPTS -->
<!-- Bootstrap JS -->
<script src="https://cdn.jsdelivr.net/npm/@popperjs/core@2.10.2/dist/umd/popper.min.js"
integrity="sha384-7+zCNj/IqJ95wo16oMtfsKbZ9ccEh31eOz1HGyDuCQ6wgnyJNSYdrPa03rtR1zdB"
crossorigin="anonymous"></script>
<script src="https://cdn.jsdelivr.net/npm/bootstrap@5.1.3/dist/js/bootstrap.min.js" integrity="sha384-
QJHtvGhmr9XOIpI6YVutG+2QOK9T+ZnN4kzFN1RtK3zEFEIsxhlmWl5/YESvpZ13"
```

```
crossorigin="anonymous"></script>

  </body>
</html>
/* GENERALS BEGIN */
html{
    max-width: 100%;
    overflow-x: hidden;
    scroll-behavior: smooth;
    padding:0;
    margin:0;
}
body{
    background-color: #e5e5e5; /#CCCCCC/
    padding:0;
    margin:0;
    font-size: 1.125rem;
    color: black;
}
a{
    color:black;
    text-decoration: none;
    background-color: transparent;
}
blockquote{
    font-style:italic;
    padding-left: 40px;
    margin: 0 0 1rem;
    border-left: 2px solid #4717f6;
}
th, td{
    padding: 8px;
}
mark{
    background-color:rgba(225, 224, 0, 0.7);
}
a:hover{
    color:#4717f6;
}
.heading{
    font-size: 80px;
}
@media (min-width: 576px){

    .container, .container-sm{
        max-width: 540px;
    }
}
@media (min-width: 768px){

    .container, .container-md, .container-sm{
        max-width: 720px;
```

```
    }
}
@media (min-width: 992px){

    .container, .container-lg, .container-md, .container-sm{
        max-width: 960px;
    }
}
@media (min-width: 1200px){

    .container, .container-lg, .container-md, .container-sm, .container-xl{
        max-width: 1080px;
    }
}

.col, .col-1, .col-10, .col-11, .col-12, .col-2, .col-3, .col-4, .col-5, .col-6, .col-7, .col-8, .col-9, .col-auto, .col-lg, .col-lg-1, .col-lg-10, .col-lg-11, .col-lg-12, .col-lg-2, .col-lg-3, .col-lg-4, .col-lg-5, .col-lg-6, .col-lg-7, .col-lg-8, .col-lg-9, .col-lg-auto, .col-md, .col-md-1, .col-md-10, .col-md-11, .col-md-12, .col-md-2, .col-md-3, .col-md-4, .col-md-5, .col-md-6, .col-md-7, .col-md-8, .col-md-9, .col-md-auto, .col-sm, .col-sm-1, .col-sm-10, .col-sm-11, .col-sm-12, .col-sm-2, .col-sm-3, .col-sm-4, .col-sm-5, .col-sm-6, .col-sm-7, .col-sm-8, .col-sm-9, .col-sm-auto, .col-xl, .col-xl-1, .col-xl-10, .col-xl-11, .col-xl-12, .col-xl-2, .col-xl-3, .col-xl-4, .col-xl-5, .col-xl-6, .col-xl-7, .col-xl-8, .col-xl-9, .col-xl-auto {
    padding-right: 15px;
    padding-left: 15px;
}
/* GENERALS END */

/* COLOR SCHEMES */
.black{
    color: black!important;
}
.blue-bg{
    background-color: #b2b2ff;
}
.secondary-bg{
    background-color: #fafafa;
}
.purple-bg-secondary{
    background-color: #d8b2d8;
}
.purple-bg{
    background-color: #a239ca;
}
.btn-primary {
    border-style: solid;
    font-size: inherit;
    font-weight: inherit;
    text-transform: inherit;
    box-shadow: rgb(0 0 0 / 30%) 0 0 10px;
    background-color: #a239ca;
    border-color: transparent;
    color: #fff;
```

```
}
.btn:hover{
   background-color: #920ec7;
   border-color:transparent;
}
.bg-light{
   background-color: white!important;
}
.blue-bg-secondary{
   background-color: #ccccff;
}
.button-no-style{
   border: 0;
}
/* COLOR SCHEME ENDS */
/* TEXT-DECORATION BEGINS */
.display-3{
   font-size:72px;
}
.link-unstyled{
   color:inherit;
}
.marker-highlight{
   border-radius: 1em 0 1em 0;
   background-image: linear-gradient(-100deg, rgba(225, 224, 0, 0.3), rgba(225, 224, 0, 0.7) 95%,
rgba(225, 224, 0, 0.1))
}
.border-top-purple{
   border-top: 1px solid #a239ca!important;
}
/* TEXT-DECORATION ENDS */

/* NAVBAR BEGINS */
.navbar-light{
   background-color: white;
}
.navbar li{
   transition: 0.3s;
   margin: 0 8px 0 0;
   font-weight: 600;
}
.navbar li:hover{
   background-color: #b2b2ff;
   border-radius: 0.375rem;
}
/* NAVBAR ENDS */

/* CONTENT SECTION BEGINS */
.container > .main-content{
   margin-top: 5rem;
}
.main-content{
```

```
    margin-top:2rem;
}
.card {
    position: relative;
    display: flex;
    flex-direction: column;
    min-width: 0;
    word-wrap: break-word;
    background-color: #fff;
    background-clip: border-box;
    border: 1px solid rgba(0,0,0,.125);
    border-radius: 0.25rem;
}
.border-0 {
    border: 0!important;
}
.border-black{
    border: 1px solid black;
}
.border-purple{
    border: 2.5px dashed purple;
}
/* FORM SECTION BEGINS */
.form-group{
    margin-bottom: 1rem;
}
.form-link{
    color:rgb(0, 132, 255);
    text-decoration: underline;
}
/* FORM SECTION ENDS */
```

CHAPTER 4
SYSTEM TEST

4.1 INTRODUCTION TO TESTING

Testing is a process, which reveals the errors in the program. The WQI value is tested whether it is valid or not. By giving input in jupyter notebook, the testing is done. It represents the ultimate review of specification, design, and generation.

4.1.1Training Dataset

In this project, to train the machine learning model we have taken large dataset of the time period of five years. In this dataset we have composition weightage of 12 pollutant component present atmosphere. In this large dataset we need to provide correct data without any null values. We fill those null values with respective mean values of 12 pollution components. After that we find the sub-indices for each and every component. Then we fill WQI values with maximum sub-indices that particular row. After getting all WQI values we provide WQI level to respective WQI value. Then we train the model based on dataset. We are using 75-80% data for training.

4.1.2Testing Dataset

The test model we have separated the dataset. The remaining 20-25% data which we have separated for testing is used to test. The model is evaluated by using this test data. This evolution had done. This evolution performance will display by using confusion matrix. The accuracy is displayed. We also get the how many are correct and wrong prediction in the data.

4.2 TESTING METHODS

We have different types of testing methods that our model should be tested so that we can verify if there any errors in it. The testing methods are

4.2.1 Unit testing

Unit testing is the first level of testing and is often performed by the developers itself. It is the process of ensuring unique components of a piece of software at the coding level are functional and work as they were designed to. Developers in the test driven environment will typically write and run the tests prior to the software or feature being passed over to the test team. Unit testing will also make debugging very easier because finding issues earlier means they take very less time to fix than if they found later in the testing process.

4.2.2 White Box Testing

White Box Testing is a testing in which in which the software tester has knowledge of the inner workings, structure and language of the software, or at least its purpose. It is purpose. It is used to test areas that cannot be reached from a black box level.

4.2.3 Black Box Testing

Black Box Testing is testing the software without any knowledge of the Inner workings, structure or language of the module being tested. Black box tests, as most other kinds of tests, must be written from a definitive source document, such as specification or requirements document, such as specification or requirements document. It is a testing in which the software under test is treated, as a black box. you cannot

"see" into it. The test provides inputs and responds to outputs without considering how the software works. **Unit Testing**

Unit testing is usually conducted as part of a combined code and unit test phase of the software lifecycle, although it is not uncommon for coding and unit testing to be conducted as two distinct phases.

Test strategy and approach

Field testing will be performed manually and functional tests will be written in detail.

Test objectives

- All field entries must work properly.
- Pages must be activated from the identified link.
- The entry screen, messages and responses must not bedelayed.

Features to be tested

- Verify that the entries are of the correct format
- No duplicate entries should be allowed.
- All links should take the user to the correct page.

4.2.4 Integration testing

After each unit is tested, it is integrated with other units to create components that are designed to perform specific activities. These are then tested as a group through

integration testing to ensure whole segments of an application behave as expected.

4.2.5 System testing

System testing is a black box testing method used to evaluate the completed and integrated system, as a whole, to ensure it meets specified requirements. The functionality of the software is tested from end-to-end and is typically conducted by a separate testing team before the product is pushed into production.

4.2.6 Acceptance testing

Acceptance testing is the last phase of functional testing and is used to assess whether or not the final piece of software is ready for delivery. It involves ensuring that the product is in compliance with all of the original business criteria and that it meets the end user's needs. This requires the product be tested both internally and externally. Beta testing is the real key for getting real feedback from the potential customers and can even address any final usability concerns.

4.3 ACCURACY AND VALIDATION

Accuracy is the measurement which is used to determine which model is best at identifying relationships and patterns between variables in the dataset based on the input or training data. The better a model can generalize to unseen data in the prediction and insights it can produce which in turn deliver more business value. In this project, we are trained a model for prediction of Water quality which is having the accuracy rate. When the input WQI value is given, it needs to be checked for validation which checks if the output is what we areexpecting. It prevents third party users from mishandling the code accidentally or intentionally. It can be used to check if they are no invalid values in the given input data. In this project, it is validated whether the WQI value is valid or not.

The testing is accurate as much efficient to the input WQI value for getting the desired output. Once the output is displayed, it validates whether the output which is displayed on the screen is correct or not. Based on that, the accurate level of the project is estimated.

CHAPTER 5

SCREENSHOTS

This chapter provides various views of the project output, through screenshots with brief description.

5.1 HOME PAGE

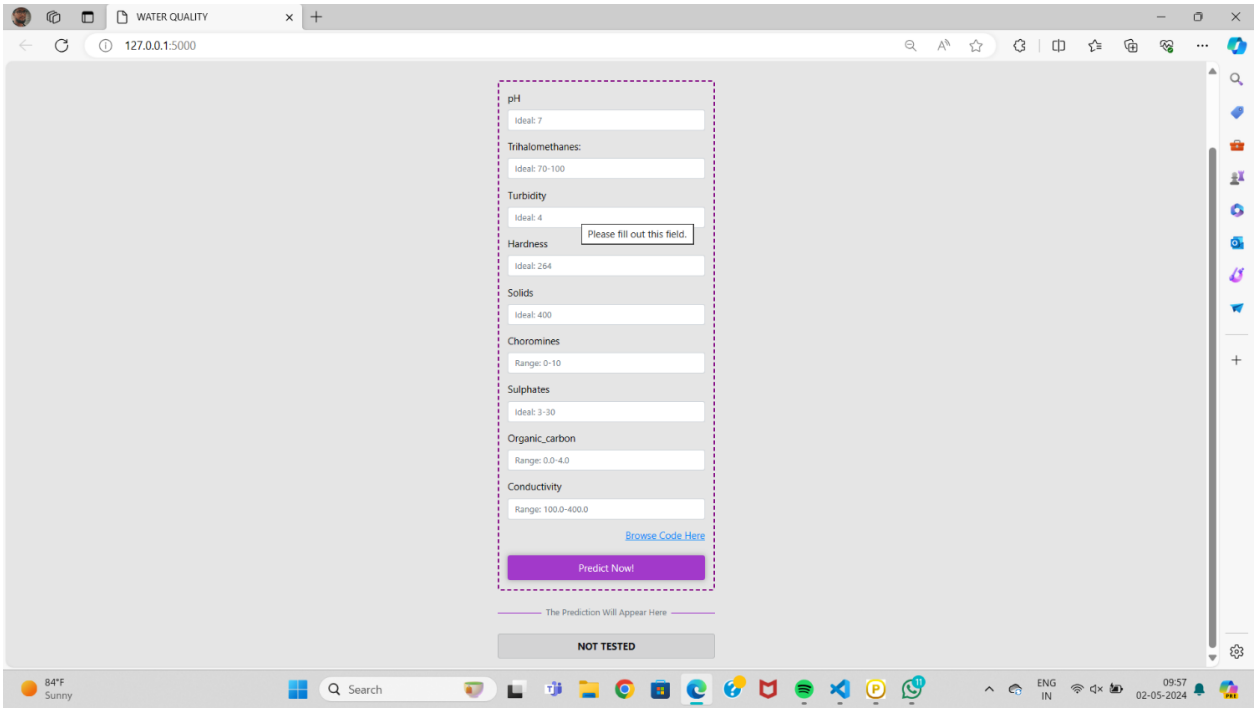


fig 5.1 HOME PAGE

The above picture represents the home page of the developed application, it works as an interface which allows the user to interact with the WATER QUALITY PREDICTION SYSTEM that ultimately give us the prediction.

The home page consists of several input fields which are considered to be capable of affecting the Water quality and those fields namely pH , Trihalomethanies , Tutbidity , Hardness , Solids , Choromines , Sulphates , Organic_carbon ,Conductivity the at the bottom of the page a button is provide that says predict, The user have to give the input to the prediction system by entering the values according to his will and have to click the below provide button then the prediction system willgive the prediction based on the values which are provided by the user.

5.2 INPUT PAGE

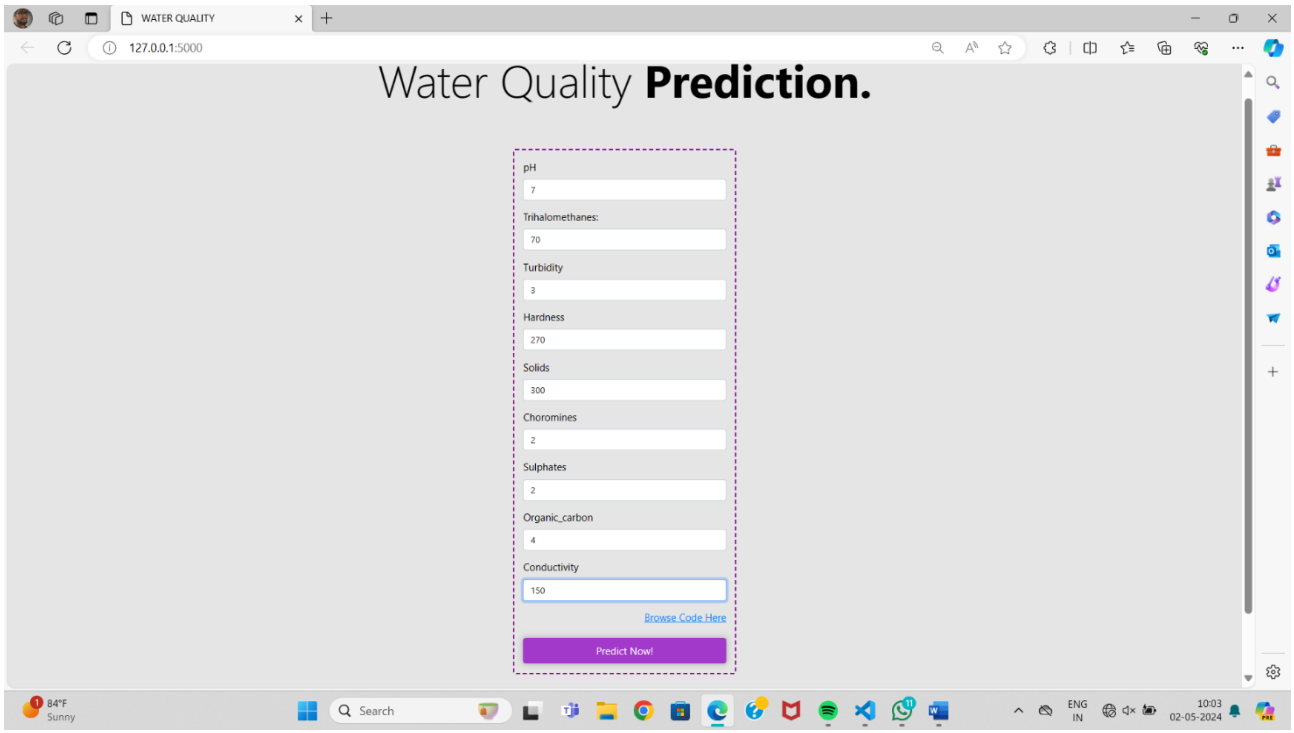


Fig 5.2 INPUT PAGE

The above picture gives us the insights of how the user can give input to the prediction system through the interface, the user can give the variety of values including the user can give the inputs as fraction of numbers that is in floating points, if the user wants not to give the input for some specific input fields the user can choose to enter zero, the zero in the particular input field gives the result.

5.1 OUTPUT PAGE

The final output which is shown to the user is shown as follows:

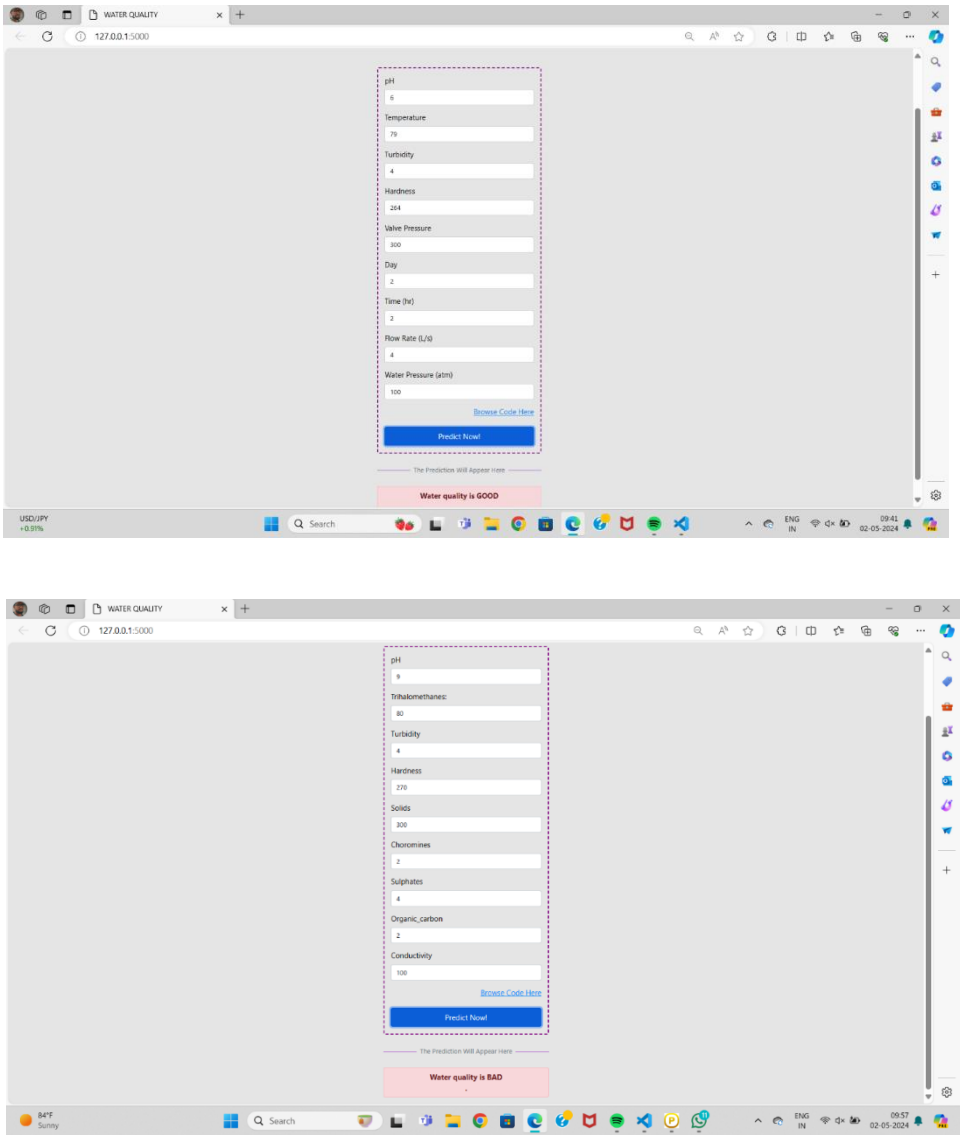


Fig 5.3 OUTPUT PAGE

The user need to refresh the page and each time when we refresh the page the new info will be displayed to the user at a page.

CHAPTER 6

CONCLUSION

6.1 CONCLUSION

According to World Health Organization, 7 million people are at health risk due to water pollution. It is a leading risk factor for majority of health problems like cholera , typhoid , arsenicosis , hepatitis ,Stomach cramps , Respiratory infections , Kidney damage , endocrine damages . Besides the health problems related to Water pollution, it also poses a serious threat to our planet. In order reduce this water pollution predicting the water quality in early stage accuratelyhelps to take the better decisions to maintain the water quality.

In our proposed model, we had taken verified and large period dataset which is used to predict and give better results. The Decision Tree is well trained and for prediction,it gives best output. The proposed system is tested on a large amount of data under different pollutants.. The program developed here in this project is performing better than similar worksproduced till date

6.2 SCOPE FOR THE FUTURE WORK

India meteorological department wants to automate the detecting of Water quality whether good or not from eligibility process (real time). Many efforts from both local and state government are done in order to understand and predict Water quality index aiming improved public health. In order to optimize the Artificial Intelligence environment, the prediction system was implemented. With the advancement of IoT infrastructures, big data technologies, and machine learning techniques, real-time Water quality monitor and evaluation is desirable for future smart cities. For future work, the commonalities between nearby meteorology stationsand combine them in a MTL framework, which may provide a further boosting for the prediction can be considered. This work can be extended by considering data changes over time for real-time forecasting. This can be achieved by building online models that adapt automatically to changes in environment. Also, more data can be included to increase data seasonality.

CHAPTER 7

REFERENCES

1 In 2018, Ali Heidar Nasrolahi along with, Ca, Mg, etc such components into consideration. Performance was tallied by using several ML and DL algorithms. Amir Hamzeh Haghiabi and Abbas Parsaie predicted the Water Quality of a river bed in Iran Tireh River by taking pH, Na.

2. Clark RM, Hakim S, Ostfeld A (2011) Handbook of water and wastewater systems protection. In: Pro-tecting Critical Infrastructure. Springer, pp 1–29. https:// doi. org/ 10. 1007/ 978-1- 4614- 0189-6

3. In 2020, Mohammed Al-Yaari et.al illustrated the use of Artificial Intelligence algorithms along with the performance of each used algorithm. As we know, for the protection of the environment, predicting and modelling of the quality of water is immensely important. In the methodology they proposed, to predict WQI, artificial intelligence algorithms, such as, NARNET and LSTM were used.

4. In 2020, Navideh Noori et.al explained the water quality prediction using SWAT-ANN coupled approach. For solving environmental problems Machine Learning algorithm such as Artificial Neural Networks is being used widely. They illustrated the application of SWAT-ANN for water quality prediction.

5. Waqas M, Tu S, Halim Z et□al (2022) The role of artificial intelligence and machine learning in wire-less networks security: principle, practice and challenges. Artif Intell Rev 55:5215–5261. https:// doi. org/ 10. 1007/ s10462- 022- 10143-2

6. Halim Z, Waqar M, Tahir M (2020) A machine learning-based investigation utilizing the in-text fea-tures for the identification of dominant emotion in an email. Knowl Based Syst 208:106443. https:// doi. org/ 10. 1016/j. knosys. 2020. 106443

7. Wu J, Wang Z (2022) A Hybrid Model for Water Quality Prediction Based on an Artificial Neural Network, Wavelet Transform, and Long Short-Term Memory. Water 14:610

8. Lee S, Lee D (2018) Improved prediction of harmful algal blooms in four Major South Korea's Riv-ers using deep learning models. Int J Environ Res Public Health 15:1322

9. Liu P, Wang J, Sangaiah AK et□al (2019) Analysis and prediction of water quality using LSTM deep neural networks in IoT environment. Sustainability 11:2058

CHAPTER 8

STUDENT BIO-DATA

Student-1

Name: G.Anusha

Father Name: G.Venkata Ramana

Roll. No: 2010115

Date of Birth:15/03/2003

Nationality: Indian

Communication Address:

Town/Village: Thirumalaiahgari palli Mandal:Thanakal

District: Anantapur

PIN Code: 515571

Ph. No: 6302826604

e-mail: anushaanil2003@gmail.com

Permanent Address:

Town/Village: Thirumalaiahgari Palli Mandal: Thanakal

District: Anantapur

PIN Code: 515571

Ph. No: 6302826604

e-mail: anushaanil2003@gmail.com

Qualifications: B. Tech – Computer Science and Engineering

Technical Skills: Java, Mysql , Manual Testing ,Html, CSS,

Area of Interest: Java Full Stack Development

Declaration:

I hereby declared that all the information provided above is accurate to the best of my knowledge.

Signature

Student-2

Name: K. Vineeth Kumar

Father Name: K. Venkataramana

Roll. No: 2010121

Date of Birth: 22/08/2003

Nationality: Indian

Communication Address:

Town/Village: Sadum Mandal: Sadum

District: Chittoor

PIN Code: 517123

Ph. No: 7995892827

e-mail: katarivineethkumar@gmail.com

Permanent Address:

Town/Village: Sadum Mandal: Sadum

District: Chittoor

PIN Code: 517123

Ph. No: 7995892827

e-mail: katarivineethkumar@gmail.com

Qualifications: B. Tech – Computer Science and Engineering

Technical Skills: Python, Mysql, Html, CSS

Area of Interest: Python Full Stack Development

Declaration:

I hereby declared that all the information provided above is accurate to the best of my knowledge.

Signature

Student-3

Name: M. Hemanth Kumar Reddy

Father Name: M. Sudhakar Reddy

Roll. No: 2010125

Date of Birth: 30/06/2002

Nationality: Indian

Communication Address:

Town/Village: N.Kothapalli Mandal: Kalakada

District: Annamaya

PIN Code: 517305

Ph. No: 6303076922

e-mail: hemanthmajjiga2002@gmail.com

Permanent Address:

Town/Village: N.Kothapalli Mandal: Kalakada

District: Annamaya

PIN Code: 517305

Ph. No: 6303076922

e-mail: hemanthmajjiga2002@gmail.com

Qualifications: B.Tech - Computer Scinece and Engineering

Technical Skills: Core Java, Data Structures, Html, CSS , Java Script

Area of Interest: Java Full Stack Development

Declaration:

I hereby declared that all the information provided above is accurate to the best of my knowledge.

Signature