**MINI PROJECT REPORT**

**ON**

**WATER QUALITY CHECK**

**BACHELOR OF TECHNOLOGY**

**(Computer Science and Engineering)**

***by***

**T.TANUJA (18NM1A05F7)**

**S.MONIKA (18NM1A05D5)**

**S.CHANDU (18NM1A05D0)**

**R.S.S.V.UMA SRI(18NM1A05C9)**

**Y.S.L.SAHITHI(18NM1A05G7)**

*Under the guidance of*

**MR.R.RAVI**

**(Assistant professor)**

**DEPARTMENT OF COMPUTER SICENCE AND ENGINEERING**

**VIGNAN’S INSTITUTE OF ENGINEERING FOR WOMEN**

(Affiliated to Jawaharlal Nehru Technological University,Kakinada)

Kapujaggarajupeta, Vadlapudi Post, Visakhapatnam, Andhra Pradesh.

**2021**



**DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING**

**CERTIFICATE**

This is to certify that the mini project report titled **“WATER QUALITY CHECK”** is being submitted to Department of CSE by **T.TANUJA(18NM1A05F7), S.MONIKA(18NM1A05D5), S.CHANDU(18NM1A05D0) and R.S.S.V.UMASRI (18NM1A05C9),Y.S.L.SAHITHI (18NM1A05G7)** during the study of IV B.Tech I Semester of **Bachelor of Technology** in **COMPUTER SCIENCE & ENGINEERING** during the period November 2021-March 2022.

**Mini Project Guide**

**Head of the Department**

Mr.R.Ravi

Mr.Dr.K.Vijaya kumar

Asst. Professor

Hod of CSE Dept

**DECLARATION**

We hereby declare that this project entitled **WATER QUALITY CHECK** submitted as mini project during IV B Tech I Semester of **Bachelor of Technology** in **Computer Science & Engineering** is the original work done under the guidance of Mr.R.Ravi,Assistant Professor , Department of Computer Sciene & Engineeering, Vignan’s Institute of Engineering for women, Visakhapatnam.

**T.TANUJA(18NM1A05F7)**

**S.MONIKA(18NM1A05D5)**

**S.CHANDU(18NM1A05D0)**

**R.S.S.V.UMASRI(18NM1A05C9)**

**Y.S.L.SAHITHI(18NM1A05G7)**



**ACKNOWLEDGEMENT**

We express our sincere indebtedness towards our guide **Mr.R.Ravi**, Department of Computer Science & Engineering, VIEW for his invaluable guidance, suggestions and supervision throughout the work. Without his kind patronage and guidance, the mini project would not have taken shape. We would also like to express our gratitude and sincere regards for his kind approval of the mini project, time to time counseling and advices.

We would also like to thank to Dr. K. Vijaya Kumar, Associate Professor and Head of the Department, Department of Computer Science & Engineering, VIEW for his expert advice and counseling from time to time.

We owe sincere thanks to all the faculty members in the Department of Computer Science & Engineering for their kind guidance and encouragement from time to time.

**T.TANUJA**

**S.MONIKA**

**S.CHANDU**

**R.S.S.V.UMASRI**

**Y.S.L.SAHITHI**



ABSTRACT

Water is one of the renewable resources essential for the existence of all forms of life. That’s the reason why water quality monitoring is a key to assuring the health of living beings. Accuracy detection of pollutants in water is critical and expensive. So the brief idea of our project is to present an affordable method to find the quality of water based on machine learning which consists of many classification and prediction models.We need efficient detection algorithms to predict the quality of water. Classification models like decision tree model or random forest model can be used to predict whether the water is purified or not with the help of its properties.Our project will be deployed as a web application so that anyone can access it effortlessly with the help of url.



|  |  |
| --- | --- |
| CONTENTS | Page.No |
| **TITLE** |  |
| * Declaration | *3* |
| * Acknowledgement | 4 |
| * Abstract | *5* |
| **TITLES** |  |
| **Chapter 1: Introduction** | *8* |
| * About Water | *8-9* |
| **Chapter2:Software requirements** | *9* |
| ▪ HTML AND CSS | *9-10* |
| ▪ *FLASK* | *10-11* |

**INDEX**

**INDEX**

CSE DEPT.Page 6

|  |  |
| --- | --- |
| **Chapter 3:Processing models** | 12 |
| ▪ Waterfall Model | 12-14 |
| **Chapter 4:Machine learning** | 15 |
| ▪ Introduction | 15-17 |
| ▪ Data Pre-Processing | 17 |
| ▪ Supervised learning and Unsupervised  learning | 18-19 |
| ▪ Random Forest | 19-21 |
| **Chapter 5:Project View and Code** | 21 |
| ▪ Data Flow and Modules | 21-23 |
| ▪ Pages | 24-25 |
| ▪ Output | 25-27 |
| ▪ Code | 28-42 |

CSE DEPT.Page 7

CSE DEPT.Page 8

CHAPTER1

INTRODUCTION

**1.1 About Water**

Water is the most significant resource of life, crucial for supporting the

life of most existing creatures and human beings. Living organisms need water with

enough quality to continue their lives. There are certain limits of pollutants 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 freshwater, such as ground and surface water, are

natural water resources.

Hence, rapid industrial development has prompted the decay of water quality at a disturbing

rate. Furthermore, infrastructures, with the absence of public awareness, and less hygienic

qualities, significantly affect the quality of drinking water. In developing countries, it is

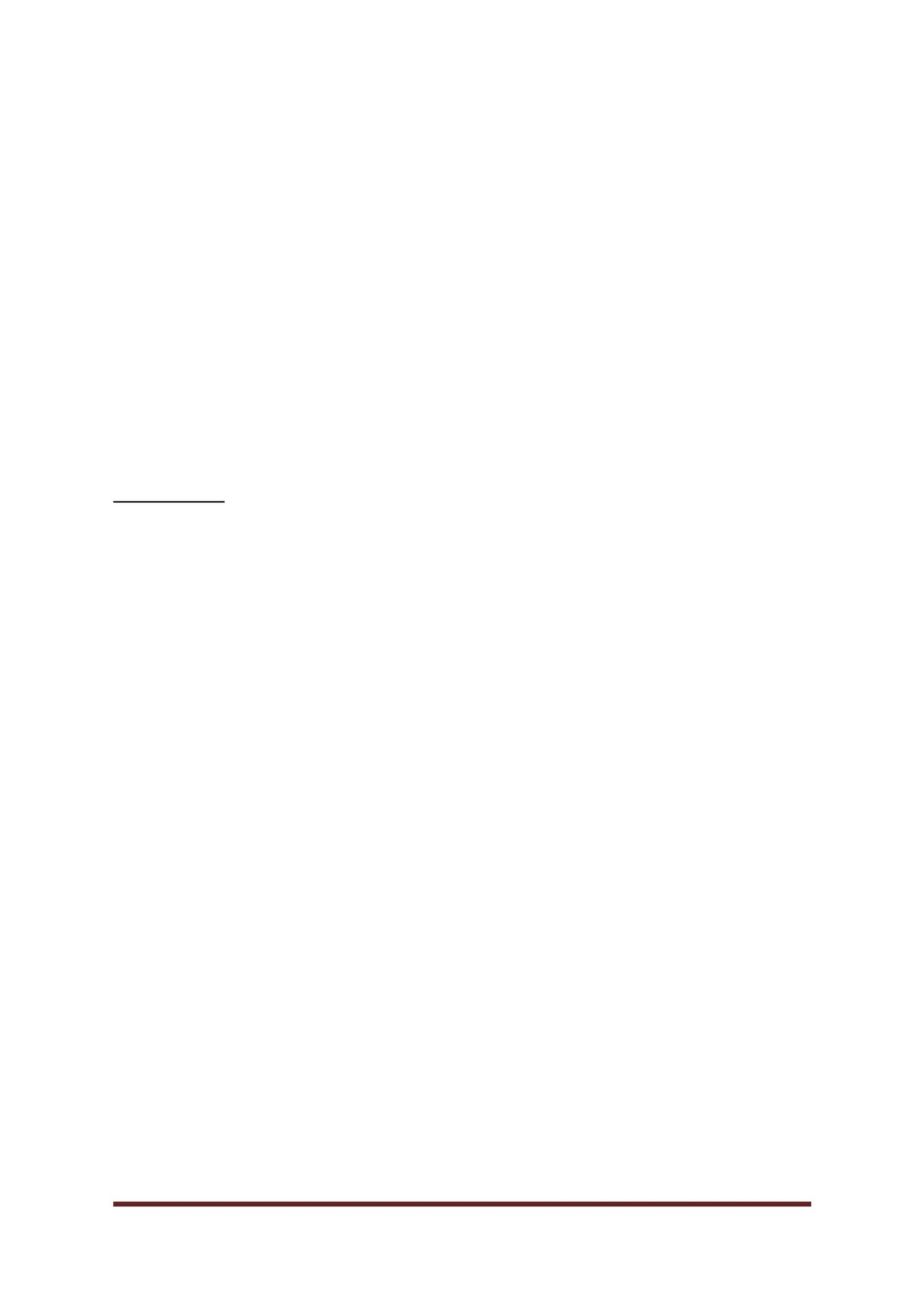
announced that 80% of health problems are caused by contaminated water.

Therefore, it is very important to suggest new approaches to analyze and, if possible, to

predict the water quality.There are several methodologies proposed for the prediction and

modeling of the WQ (Water Quality). These methodologies include statistical approaches,

visual modeling, analyzing algorithms, and predictive algorithms.



Several studies have been performed to model and predict the water quality using different

ML models. These studies have approved the feasibility and effectiveness of employing ML

applications to predict the quality of drinking water

**CHAPTER 2**

**SOFTWARE REQUIREMENTS**

**Fig.No: 2.1**

**2.1 HTML**

**HTML** stands for **HyperText Mark-up Language**, which is the most widely used

language on Web to develop web pages. **HTML** was created by Berners-Lee in late 1991

but "HTML 2.0" was the first standard HTML specification which was published in 1995.

HTML 4.01 was a major version of HTML and it was published in late 1999. Though the

HTML 4.01 version is widely used, currently we are having the HTML-5 version which is

an extension to HTML 4.01, and this version was published in 2012.

CSE DEPT.Page 9

Originally, **HTML** was developed with the intent of defining the structure of

documents like headings, paragraphs, lists, and so forth to facilitate the sharing of scientific

information between researchers. Now, HTML is being widely used to format web pages

with the help of different tags available in HTML language.

**HTML** is a MUST for students and working professionals to become a great Software

Engineer specially when they are working in Web Development Domain.

**2.2 CSS**

**CSS** is used to control the style of a web document in a simple and easy way.

**CSS** is the acronym for **"Cascading Style Sheet"**. This tutorial covers both the versions

CSS1, CSS2 and CSS3, and gives a complete understanding of CSS, starting from its basics

to advanced concepts.

CSE DEPT.Page 10

Why to Learn 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 specially when they are working in Web Development Domain.

.

Prerequisites

You should be familiar with:

●

●

●

●

●

Basic word processing using any text editor.

How to create directories and files.

How to navigate through different directories.

Internet browsing using popular browsers like Internet Explorer or Firefox.

Developing simple Web Pages using HTML or XHTML.

.

**2.3 FLASK**

**What is flask?**

Flask is a web framework. This means flask provides you with tools, libraries and

technologies that allow you to build a web application. This web application can be some

web pages, a blog, a wiki or go as big as a web-based calendar application or a commercial

website.

**What is the use of flask?**

Flask is part of the categories of the micro-framework. Micro-framework is normally a

framework with little to no dependencies to external libraries.Flask is a micro web framework

written in Python. It is classified as a microframework because it does not require particular

tools or libraries. It has no database abstraction layer, form validation, or any other

components where pre-existing third-party libraries provide common functions.

**How does a flask work?**

CSE DEPT.Page 11

If heat can't escape from a vacuum flask, it follows that heat can't penetrate into a flask from

outside either. The sealed stopper stops heat getting in by convection; the vacuum stops

conduction, and the metal lining between the outer case and the inner chamber stops heat

radiating in either.

**Advantages:**

● Scalable. I would argue more scalable than monoliths if using modern methods.

Today, applications are often running in containers or using cloud computing with

auto-scaling. Applications do not typically “scale” themselves. The infrastructure

scales. With a smaller application, it's easier to deploy instances across thousands of

server easily to handle increased traffic/load. That’s part of the reason why Pinterest

needed to migrate from Django to Flask as they grew to support more of a

Micro services pattern.

● Simpler Development. If you understand Python well, then you’ll be able to move

around and contribute to a Flask application pretty easily. It’s less opinionated so there

are fewer standards to learn.

● Flexibility. There are very few parts of Flask that cannot be easily and safely altered

because of its simplicity and minimality.



**CHAPTER 3**

**PROCESS MODELS**

**Waterfall Model**

The Waterfall Model was the first Process Model to be introduced. It is

also referred to as a **linear-sequential life cycle model**. It is very simple to

understand and use. In a waterfall model, each phase must be completed before

the next phase can begin and there is no overlapping in the phases.

The Waterfall model is the earliest SDLC approach that was used for software

development.

The waterfall Model illustrates the software development process in a linear

sequential flow. This means that any phase in the development process begins

only if the previous phase is complete. In this waterfall model, the phases do

not overlap.

CSE DEPT.Page 12

Waterfall Model – Design

The Waterfall approach was the first SDLC Model to be used widely in

Software Engineering to ensure success of the project. In "The Waterfall"

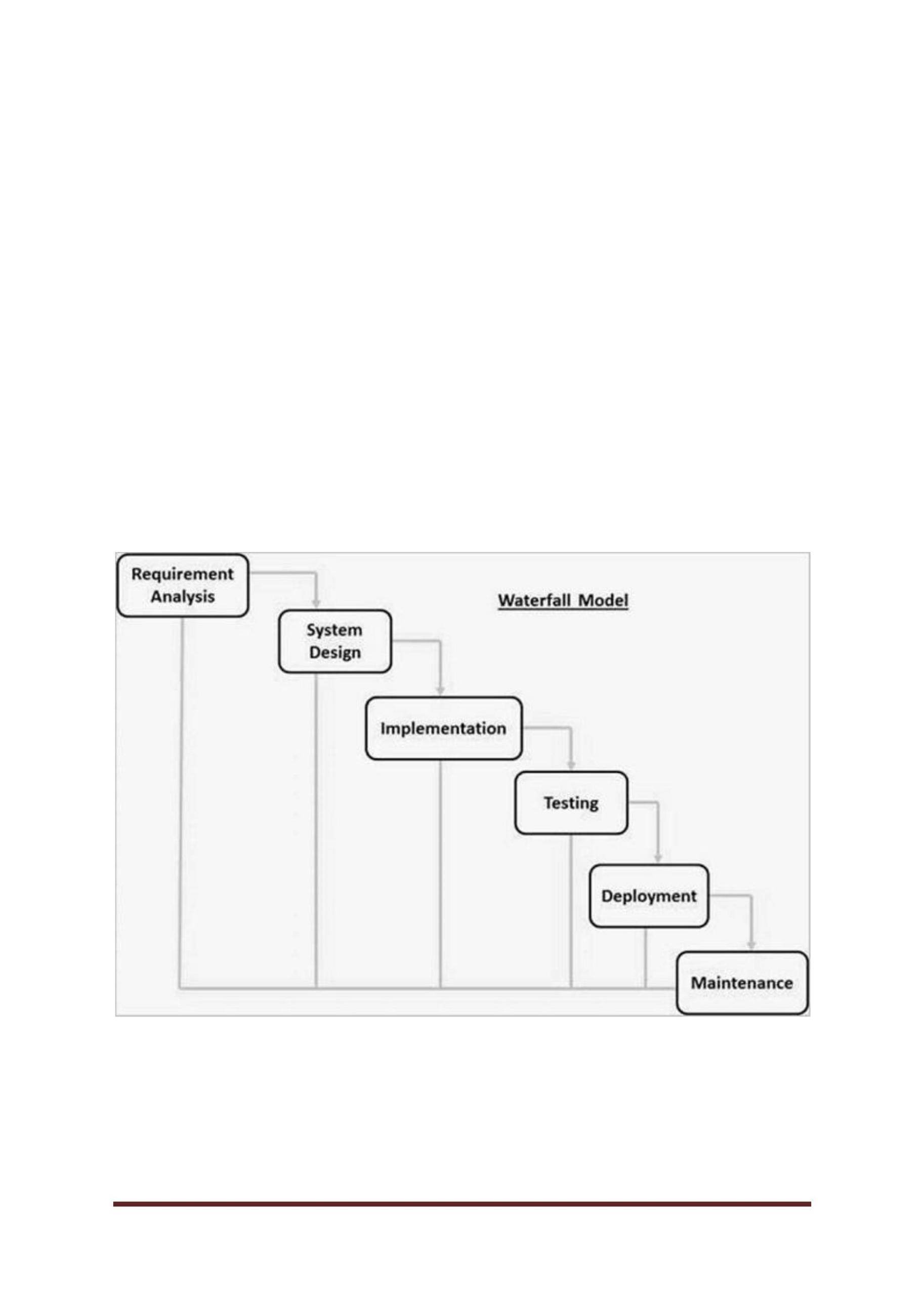
approach, the whole process of software development is divided into separate

phases. In this Waterfall model, typically, the outcome of one phase acts as the

input for the next phase sequentially.

The following illustration is a representation of the different phases of the

Waterfall Model.



CSE DEPT.Page 13

The sequential phases in Waterfall model are −

●

**Requirement Gathering and analysis** − All possible requirements of the system to

be developed are captured in this phase and documented in a requirement

specification document.

●

**System Design** − The requirement specifications from first phase are studied in this

phase and the system design is prepared. This system design helps in specifying

hardware and system requirements and helps in defining the overall system

architecture.

●

●

**Implementation** − With inputs from the system design, the system is first developed

in small programs called units, which are integrated in the next phase. Each unit is

developed and tested for its functionality, which is referred to as Unit Testing.

**Integration and Testing** − All the units developed in the implementation phase are

integrated into a system after testing of each unit. Post integration the entire system

is tested for any faults and failures.

●

●

**Deployment of system** − Once the functional and non-functional testing is done; the

product is deployed in the customer environment or released into the market.

**Maintenance** − There are some issues which come up in the client environment. To

fix those issues, patches are released. Also to enhance the product some better

versions are released. Maintenance is done to deliver these changes in the customer

environment.

All these phases are cascaded to each other in which progress is seen as flowing steadily

downwards (like a waterfall) through the phases. The next phase is started only after the

defined set of goals are achieved for the previous phase and it is signed off, so the name

"Waterfall Model". In this model, phases do not overlap.

Waterfall Model – Application

Every software developed is different and requires a suitable SDLC approach to be

followed based on the internal and external factors. Some situations where the use of

Waterfall model is most appropriate are −

●

●

●

●

●

●

Requirements are very well documented, clear and fixed.

Product definition is stable.

Technology is understood and is not dynamic.

There are no ambiguous requirements.

Ample resources with required expertise are available to support the product.

The project is short.

CSE DEPT.Page 14



CSE DEPT.Page 15

**CHAPTER4**

**Machine Learning**

**Introduction:**

The machine learning model is nothing but a piece of code; an engineer or data scientist

makes it smart through training with data. Machine Learning focuses mainly on the designing

of systems,thereby allowing them to learn and make predictions based on some experience

which is data in case of machines.

Workflow of a Machine learning project This includes all the steps required to build the

proper machine learning project from scratch. We will also go over data pre - processing, data

cleaning, feature exploration and feature engineering and show the impact that it has on

Machine Learning Model Performance.

Python Libraries that would be need to achieve the task:

1. Numpy

2. Pandas

3. Sci-kit Le

4. Matplotlib

Understanding the machine learning workflow

We can define the machine learning workflow in 5 stages:

1. Gathering data

2. Data pre-processing

3. Researching the model that will be best for the type of data

4. Training and testing the model

5. Evaluation

**1. Gathering Data:**

The process of gathering data depends on the type of project we desire to make. If we want to

make an ML project that uses real-time data, then we can build an IoT system that uses

different sensor data. The data set can be collected from various sources such as a file,

database, sensor and many other such sources but the collected data cannot be used directly

for performing the analysis process as there might be a lot of missing data, extremely large

values, unorganized text

**2. Data pre-processing:**

Data preprocessing is one of the most important steps in machine learning. It is the most

important step that helps in building machine learning models more accurately. Data

preprocessing is a process of cleaning the raw data i.e. the data is collected in the real world

and is converted to a clean data set. In other words, whenever the data is gathered from

different sources it is collected in a raw format and this data isn’t feasible for the analysis.

**3. Researching the model that will be best for the type of data:**

Our main goal is to train the best performing model possible, using the pre-processed data.We

should select the best machine learning algorithm to perform on our data to get accurate

results. we can either select a supervised or unsupervised or reinforcement machine learning

algorithms based on our requirement.

**4. Training and testing the model on data:**

For training a model we initially split the model into 3 three sections which are ‘Training

data’, ‘Validation data’ and ‘Testing data’.

You train the classifier using ‘training data set’, tune the parameters using ‘validation set’ and

then test the performance of your classifier on unseen ‘test data set’. An important point to

note is that during training the classifier only the training and/or validation set is available.

The test data set must not be used during training the classifier. The test set will only be

available during testing the classifier.

**5. Evaluation:**

Model Evaluation is an integral part of the model development process. It helps to find the

best model that represents our data and how well the chosen model will work in the future.

CSE DEPT.Page 16



To improve the model we might tune the hyper-parameters of the model and try to improve

the accuracy and also look at the confusion matrix to try to increase the number of true

positives and true negatives.

**Fig.no:4.1**

**Data Pre-Processing**

**What is data pre-processing?**

Data preprocessing is a process of cleaning the raw data i.e. the data is collected in the real

world and is converted to a clean data set. In other words, whenever the data is gathered from

different sources it is collected in a raw format and this data isn’t feasible for the analysis.

Therefore, certain steps are executed to convert the data into a small clean data set, this part

of the process is called data pre-processing.

**Why do we need it?**

As we know that data pre-processing is a process of cleaning the raw data into clean data, so

that can be used to train the model. So, we definitely need data pre-processing to achieve

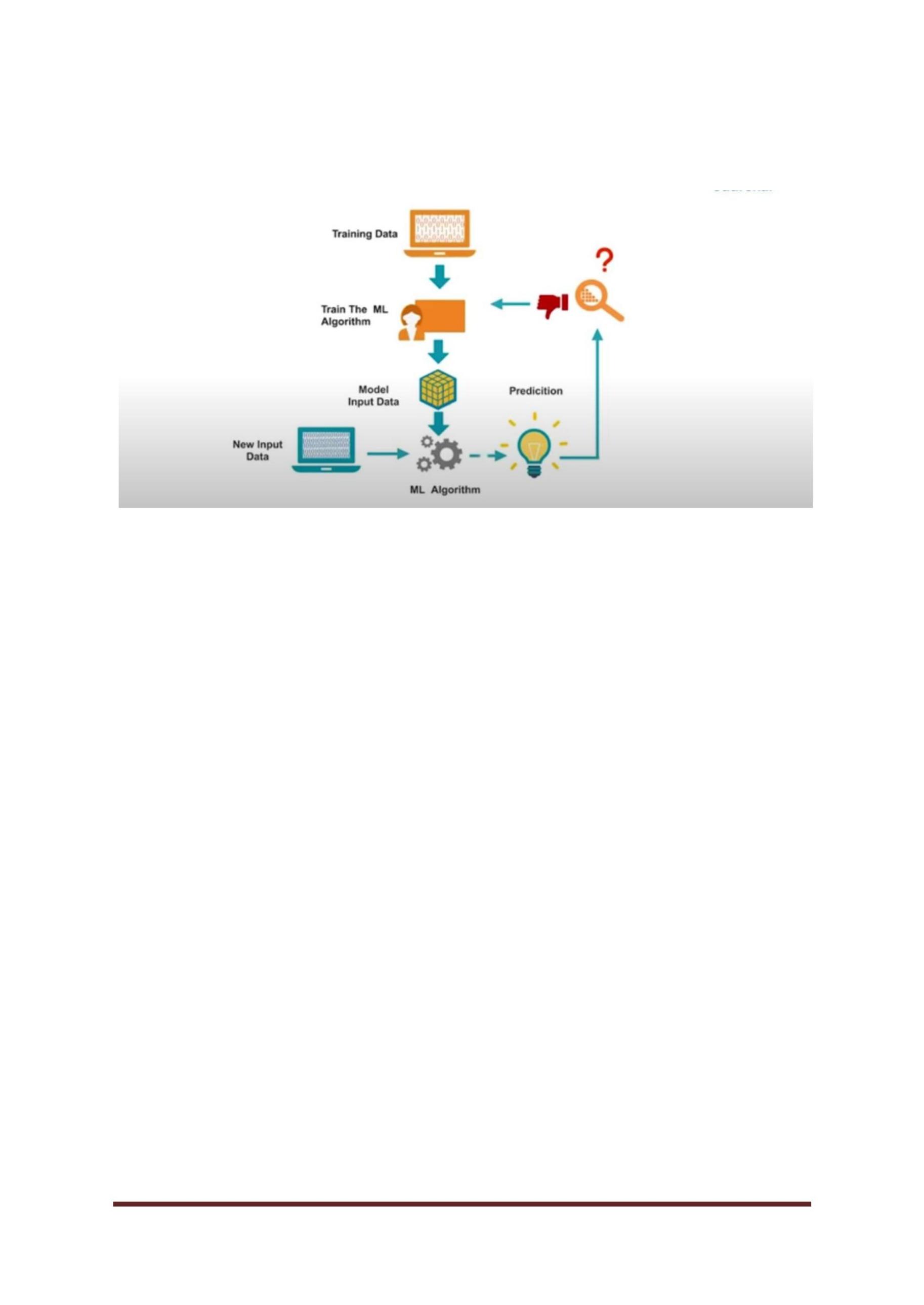
good results from the applied model in machine learning and deep learning projects.

Most of the real-world data is messy, some of these types of data are:

1. **Missing data**: Missing data can be found when it is not continuously created or due to

technical issues in the application (IOT system).

CSE DEPT.Page 17



**Supervised and Unsupervised learning**

**Fig.no:4.2**

**Supervised Learning:**

In Supervised learning, an AI system is presented with data which is labelled, which

means that each data tagged with the correct label.

The supervised learning is categorized into 2 other categories which are

“Classification” and “Regression”.

Classification problem is when the target variable is categorical (i.e. the output could

be classified into classes — it belongs to either Class A or B or something else).

These are some most used classification algorithms.

● K-Nearest Neighbor

● Naive Bayes

● Decision Trees/Random Forest

● Support Vector Machine

● Logistic RegressionRegression:

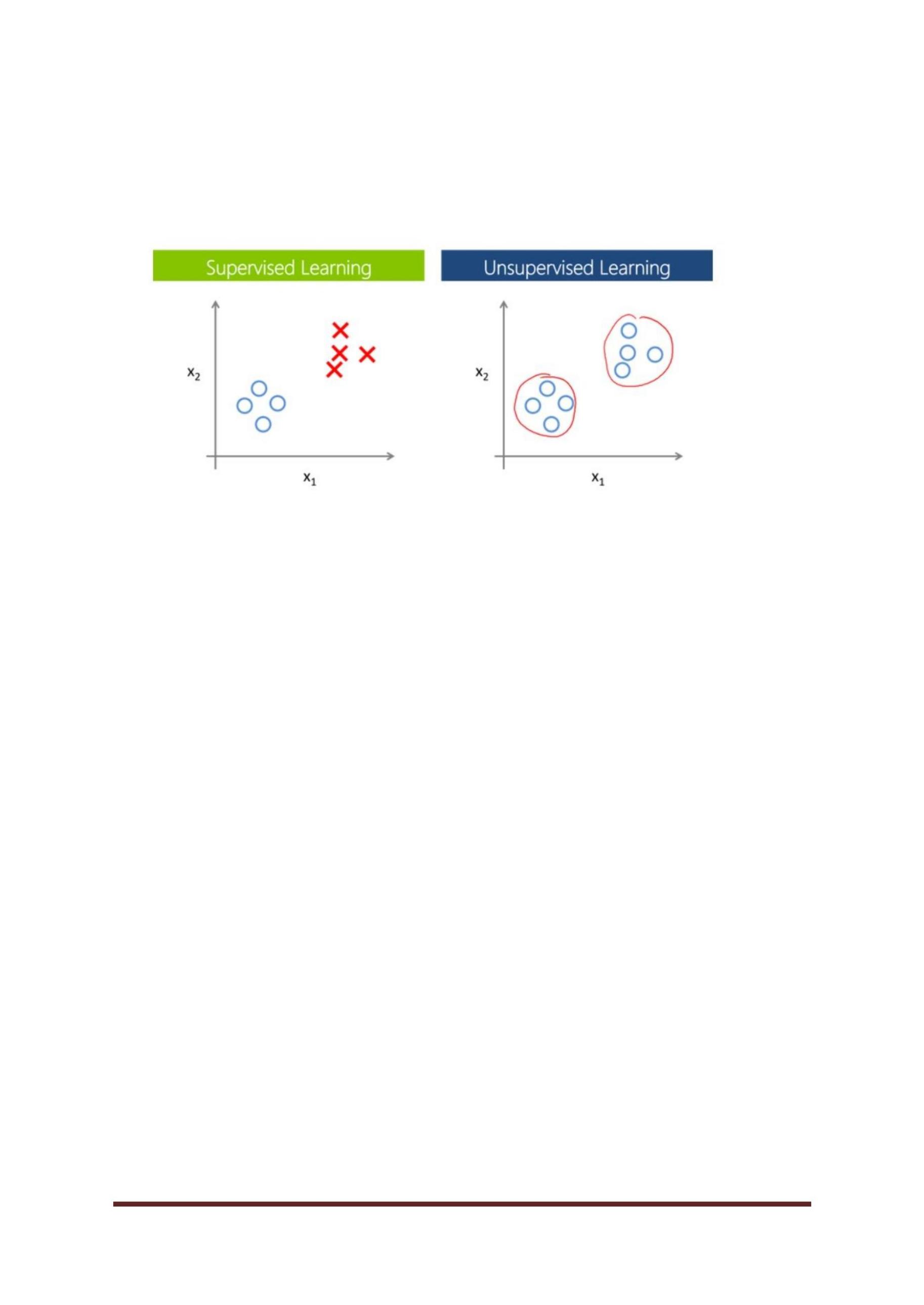
While a Regression problem is when the target variable is continuous (i.e. the output

is numeric).

These are some most used regression algorithms.

● Linear Regression

CSE DEPT.Page 18



● Support Vector Regression

● Decision Trees/Random Forest

● Gaussian Processes Regression

● Ensemble Methods

**Unsupervised Learning:**

In unsupervised learning, an AI system is presented with unlabeled, un-categorized

data and the system’s algorithms act on the data without prior training. The output is

dependent upon the coded algorithms. Subjecting a system to unsupervised learning is

one way of testing AI.

The unsupervised learning is categorized into 2 other categories which are

“Clustering” and “Association”.

**Clustering:**

A set of inputs is to be divided into groups. Unlike in classification, the groups are not

known beforehand, making this typically an unsupervised task.

Methods used for clustering are:

● Gaussian mixtures

● K-Means Clustering

● Boosting

● Hierarchical Clustering

● K-Means Clustering

● Spectral Clustering

**Random Forest**

Random forest is a supervised learning algorithm. The "**forest**" it builds is an ensemble of

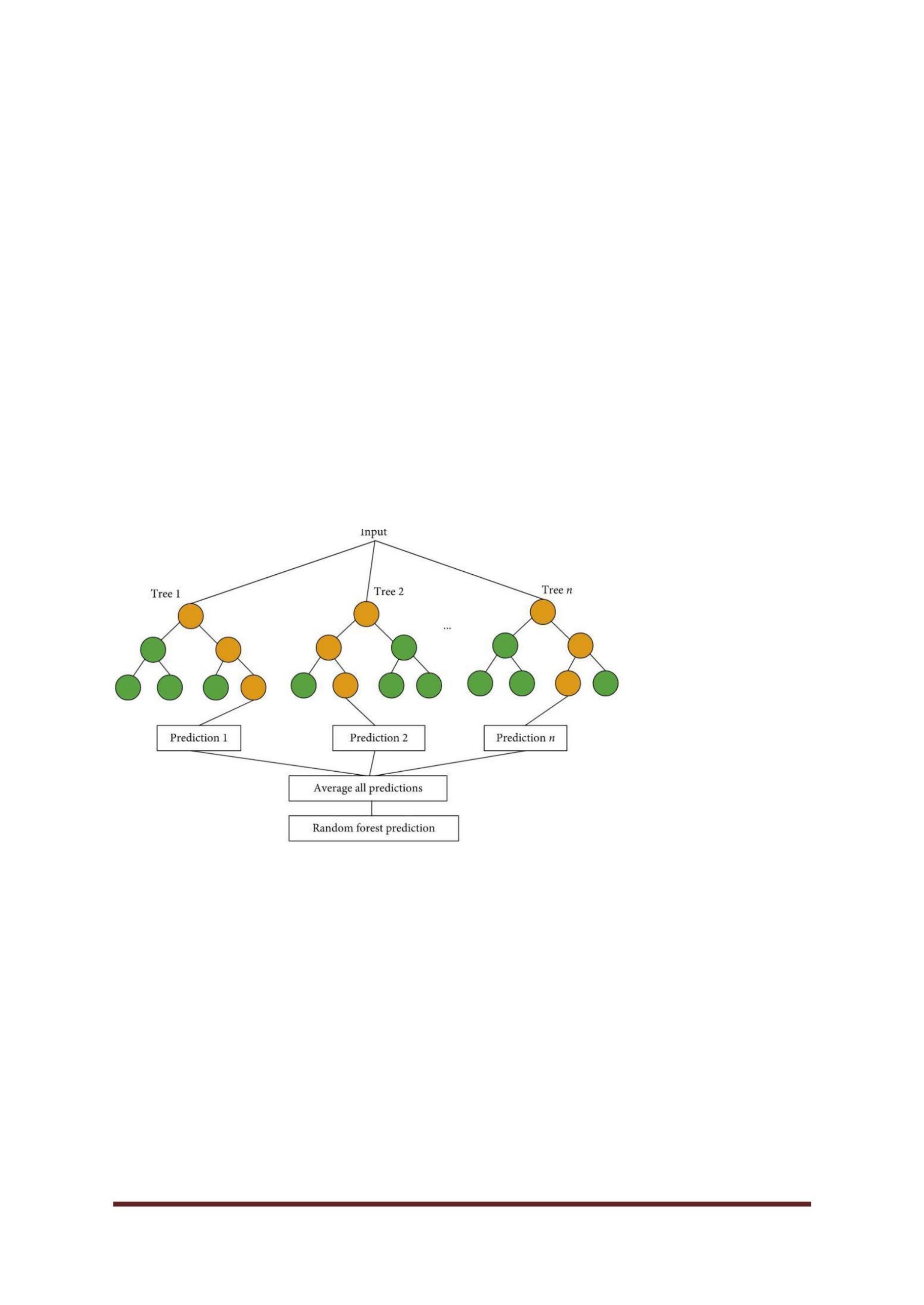
decision trees, usually trained with the “**bagging**” method. The general idea of the bagging

method is that a combination of learning models increases the overall result.

Put simply: random forest builds multiple decision trees and merges them together to get a

more accurate and stable prediction.

CSE DEPT.Page 19



The Random Forest Algorithm is used to solve both regression and classification problems,

making it a diverse model that is widely used by engineers.

Pros:

● Used for regression and classification problems, making it a diverse model.

● Prevents overfitting of data.

● Fast to train with test data.

**Cons:**

● Slow in creating predictions once the model is made.

● Must beware of outliers and holes in the data.

**Fig.no:4.3**

Another great quality of the random forest algorithm is that it is very easy to measure the

relative importance of each feature on the prediction. Sklearn provides a great tool for this

that measures a feature's importance by looking at how much the tree nodes that use that

feature reduce impurity across all trees in the forest. It computes this score automatically for

each feature after training and scales the results so the sum of all importance is equal to one.

CSE DEPT.Page 20

.

.

CSE DEPT.Page 21

• Logical DFD - This type of DFD concentrates on the system process, and flow of data in the

system.For example in a Banking software system, how data is moved between different entities.

**Advantages and Disadvantages of Random Forest**

**Advantages:**

● Powerful and highly accurate

● No need to normalizing

● Can handle several features at once

● Run trees in parallel ways

**Disadvantages:**

● They are biased to certain features sometimes

● Slow

● Can not be used for linear methods

● Worse for high dimensional data

**CHAPTER 5**

**5.PROJECT VIEW AND CODE**

**D ATA FLOW DIAGRAM AND MODULAR PAT HS**

**Data Flow Diagram ?**

Data flow diagram is a graphical representation of flow of data in an information system. It is

capable of depicting incoming data flow, outgoing data flow and stored data. The DFD does not

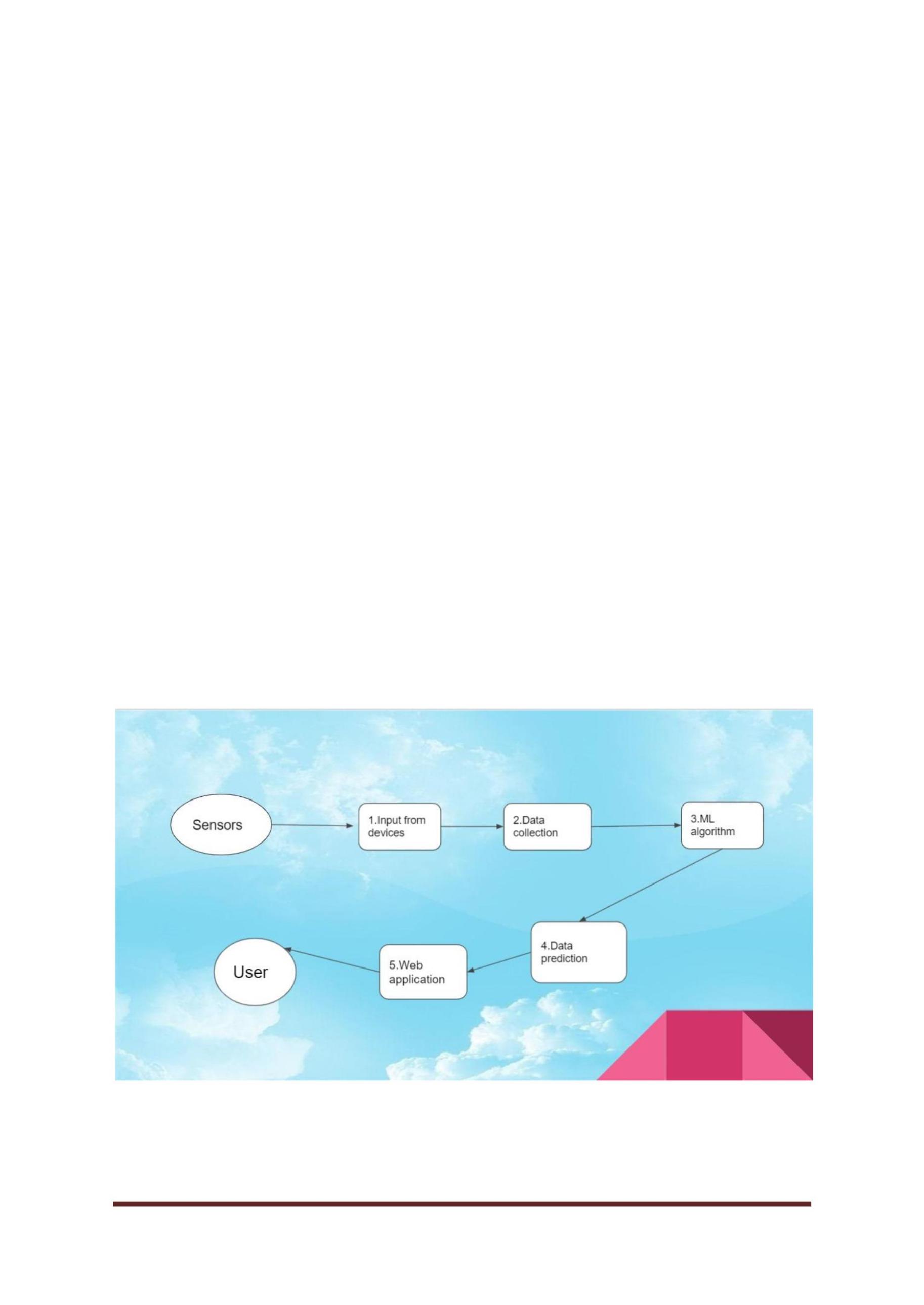
mention anything about how data flows through the system. There is a prominent difference between

DFD and Flowchart. The flowchart depicts flow of control in program modules. DFDs depict flow of

data in the system at various levels. DFD does not contain any control or branch elements.

**Types of DFD :**

Data Flow Diagrams are either Logical or Physical.



• Physical DFD - This type of DFD shows how the data flow is actually implemented in the system. It

is more specific and close to the implementation.

This is the data flow diagram for the career vision web how the data flow from home page to goal

page is all shown here in the below diagram

**Fig.No:5.1**

CSE DEPT.Page 22

**WHAT IS A MODULAR DESIGN?**

As mentioned earlier, modular design (or modularity) is an approach used to design various

products or applications – by breaking it down into separate or independent parts. These

individual parts (for example, a laptop battery) can then be used for the same functionality in

different systems or products. A creative leader chooses modular designs for their products

because of its many benefits.

In the below diagram we see how the modules are connected to each other and how the flow

passed from one module to another module these are all we find from the below diagram

**Fig.No:5.2**

CSE DEPT.Page 23

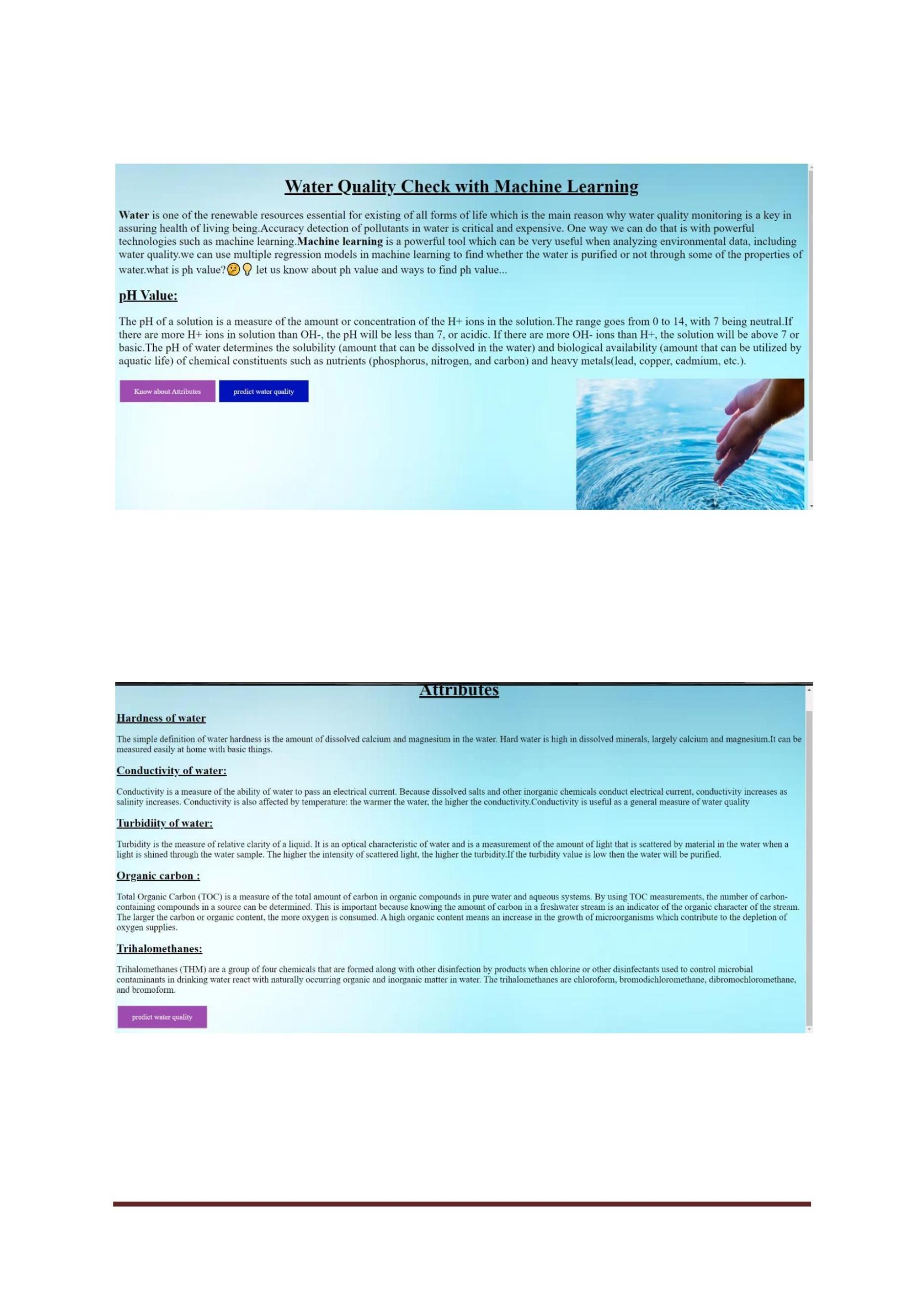


**PAGES**

**Fig.No:5.3**

**Fig.No:5.4**

CSE DEPT.Page 24

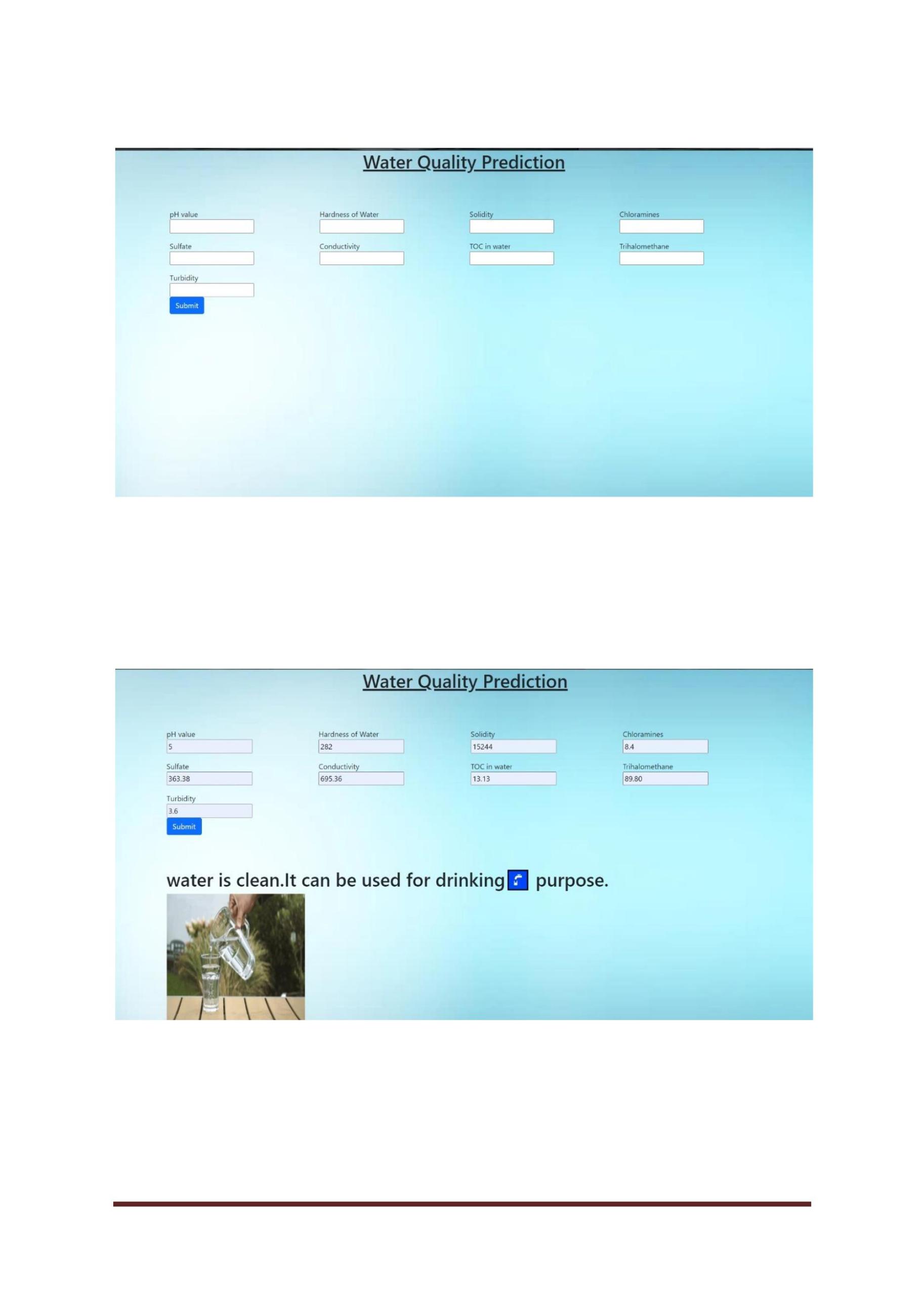


**Fig.no:5.5**

**5.4 OUTPUT**

**Fig.no:5.6**

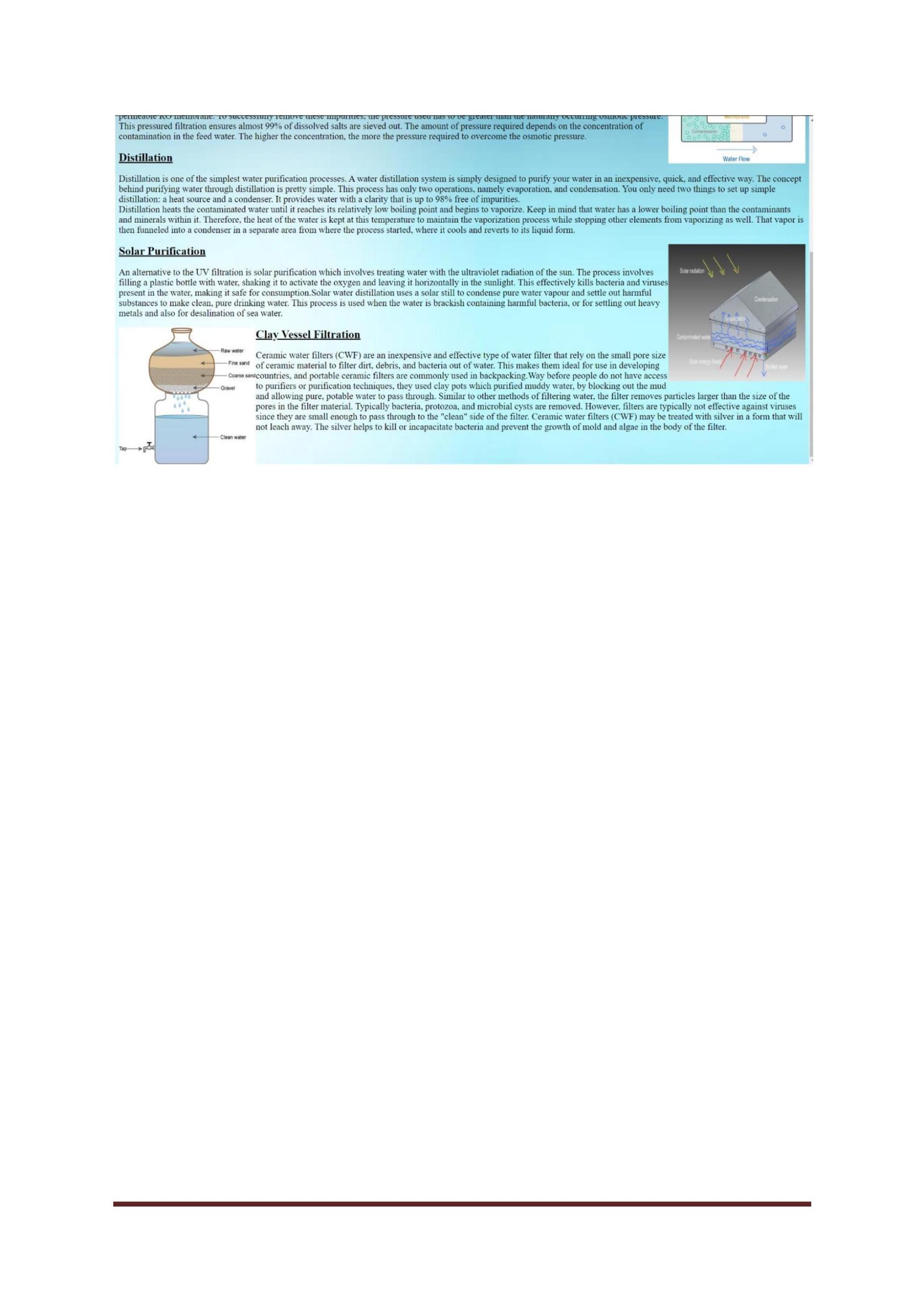
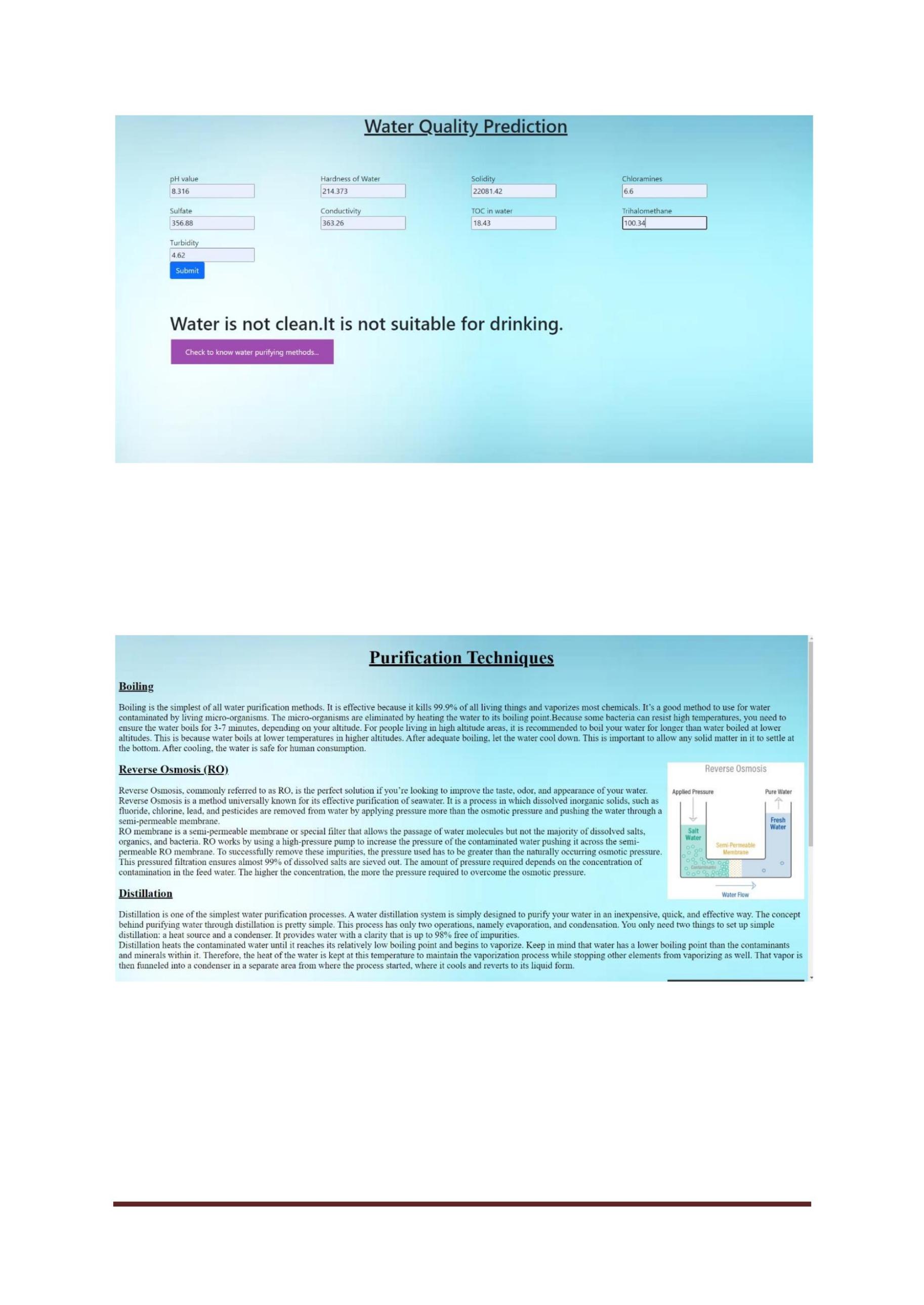
CSE DEPT.Page 25



**Fig.no:5.7**

**Fig.no:5.8**

CSE DEPT.Page 26



**Fig.no:5.9**

CSE DEPT.Page 27

**CODE**

**Water Quality Check with Machine Learning**

<!DOCTYPE html>

<html>

<head>

<title>Water Quality Check</title>

<style>

h1

{

text-align: center;

}

body

{

background-image:url("https://image.shutterstock.com/image-illustration/light-blue-gradient-

color-background-260nw-1040861230.jpg");

background-repeat: no-repeat;

background-attachment:fixed;

background-size:100% 150%;

}

.button {

border: none;

color: white;

padding: 15px 32px;

text-align: center;

text-decoration: none;

display: inline-block;

font-size: 16px;

margin: 4px 2px;

CSE DEPT.Page 28



cursor: pointer;

}

.button1 {background-color: #9e4caf;}

.button2 {background-color: #0013ba;}

</style>

</head>

<body>

<h1 style="font-size: 40px"><u>Water Quality Check with Machine Learning</u></h1>

<p style="font-size:25px"><b>Water</b> is one of the renewable resources essential for

existing of all forms of life which is the main

reason why water quality monitoring is a key in assuring health of living being.Accuracy

detection of pollutants in water is critical and expensive.

One way we can do that is with powerful technologies such as machine learning.<b>Machine

learning</b> is a powerful tool which can be very useful

when analyzing environmental data, including water quality.we can use multiple regression

models in machine learning to find whether the water

is purified or not through some of the properties of water.what is ph value?

about ph value and ways to find ph value...</p>

let us know

<h2 style="font-size:30px"><u>pH Value:</u></h2>

<p style="font-size:25px">The pH of a solution is a measure of the amount or concentration

of the H+ ions in the solution.The range goes from 0 to 14,

with 7 being neutral.If there are more H+ ions in solution than OH-, the pH will be less than

7, or acidic.

If there are more OH- ions than H+, the solution will be above 7 or basic.The pH of water

determines the solubility (amount that can be dissolved in the water)

and biological availability (amount that can be utilized by aquatic life) of chemical

constituents such as nutrients (phosphorus, nitrogen, and carbon)

and heavy metals(lead, copper, cadmium, etc.).</p>

<image align="right" src="https://www.niehs.nih.gov/health/assets/images/safewater\_og.jpg"

width=500 height=400></image>

<a href="/mini2" <button class="button button1" type="button">Know about

Attributes</button></a>

CSE DEPT.Page 29



<a

href="/input"<button

class="button

button2"

type="button">predict

water

quality</button>

</body>

</html>

**Attributes**

<!DOCTYPE html>

<html>

<head>

<title>Attributes</title>

<style>

body

{

background-image:url("https://image.shutterstock.com/image-illustration/light-blue-gradient-

color-background-260nw-1040861230.jpg");

background-repeat: no-repeat;

background-attachment:fixed;

background-size:100% 150%;

}

.button {

border: none;

color: white;

padding: 15px 32px;

text-align: center;

text-decoration: none;

display: inline-block;

font-size: 16px;

margin: 4px 2px;

CSE DEPT.Page 30



cursor: pointer;

}

.button1 {background-color: #9e4caf;}

</style>

</head>

<body>

<h1 style="font-size: 40px"align="center"><u>Attributes</u></h1>

<h2 style="font-size:25px"style="color:rgb(15, 15, 15)"><u>Hardness of water</u></h2>

<p style="font-size:20px">The simple definition of water hardness is the amount of

dissolved calcium and magnesium in the water. Hard water is high in dissolved minerals,

largely calcium and magnesium.It can be measured easily at home with

basic things.</p>

<h2 style="font-size:25px"style="color:rgb(15, 15, 15)"><u>Conductivity of

water:</u></h2>

<p style="font-size:20px">Conductivity is a measure of the ability of water to pass an

electrical current. Because dissolved salts

and other inorganic chemicals conduct electrical current, conductivity increases as salinity

increases. Conductivity is also affected

by temperature: the warmer the water, the higher the conductivity.Conductivity is useful as

a general measure of water quality</p>

<h2 style="font-size:25px"style="color:rgb(15, 15, 15)"><u>Turbidiity of

water:</u></h2>

<p style="font-size:20px">Turbidity is the measure of relative clarity of a liquid. It is an

optical characteristic of water

and is a measurement of the amount of light that is scattered by material in the water when

a light is shined through the water sample.

The higher the intensity of scattered light, the higher the turbidity.If the turbidity value is

low then the water will be purified.</p>

<h2 style="font-size:25px"style="color:rgb(15, 15, 15)"><u>Organic carbon :</u></h2>

CSE DEPT.Page 31



<p style="font-size:20px">Total Organic Carbon (TOC) is a measure of the total amount

of carbon in organic compounds in pure water and aqueous systems.

By using TOC measurements, the number of carbon-containing compounds in a source can

be determined. This is important because knowing the amount of carbon

in a freshwater stream is an indicator of the organic character of the stream. The larger the

carbon or organic content, the more oxygen is consumed.

A high organic content means an increase in the growth of microorganisms which

contribute to the depletion of oxygen supplies. </p>

<h2 style="font-size:25px"style="color:rgb(15, 15, 15)"><u>Trihalomethanes:</u></h2>

<p style="font-size:20px">Trihalomethanes (THM) are a group of four chemicals that are

formed along with other disinfection by products when chlorine or other

disinfectants used to control microbial contaminants in drinking water react with naturally

occurring organic and inorganic matter in water.

The trihalomethanes are chloroform, bromodichloromethane, dibromochloromethane, and

bromoform. </p>

<a href="/input" <button class="button button1" type="button">predict water

quality</button></a>

</body>

</html>

**Water Quality Prediction**

<!DOCTYPE html>

<html lang="en">

<head>

<title>Inputs</title>

<style>

body {

background-image:

url("https://image.shutterstock.com/image-illustration/light-blue-gradient-color-background-2

60nw-1040861230.jpg");

background-repeat: no-repeat;

CSE DEPT.Page 32



background-attachment: fixed;

background-size: cover;

}

.button {

border: none;

color: white;

padding: 15px 32px;

text-align: center;

text-decoration: none;

display: inline-block;

font-size: 16px;

margin: 4px 2px;

cursor: pointer;

}

.button1 {background-color: #9e4caf;}

</style>

</head>

<body>

<h1 style="font-size: 40px " align = center><u>Water Quality Prediction</u></h1>

<br>

<br>

<br>

<link href="https://cdn.jsdelivr.net/npm/bootstrap@5.0.0/dist/css/bootstrap.min.css"

rel="stylesheet"

integrity="sha384-wEmeIV1mKuiNpC+IOBjI7aAzPcEZeedi5yW5f2yOq55WWLwNGmvv

x4Um1vskeMj0" crossorigin="anonymous">

<form action='/predict' method="POST" class="maza">

CSE DEPT.Page 33



<div class="container">

<div class="row mb-3">

<div class="col">

<label for="ph">pH value</label></br>

<input type="text" id="pH" name="ph"/></br>

</div>

<div class="col">

<label for="Hardness">Hardness of Water</label></br>

<input type="text" id="Hardness" name="Hardness"/>

</div>

<div class="col">

<label for="Solids">Solidity</label></br>

<input type="text" name="Solids"/>

</div>

<div class="col">

<label for="Chloramines">Chloramines</label></br>

<input type="text" name="Chloramines"/>

</div>

</div>

<div class="row mb-3">

<div class="col">

<label for="Sulfate">Sulfate</label></br>

<input type="text" id="Sulfate" name="Sulfate"/></br>

</div>

<div class="col">

<label for="Conductivity">Conductivity</label></br>

<input type="text" id="Conductivity" name="Conductivity"/></br>

CSE DEPT.Page 34



</div>

<div class="col">

<label for="Organic\_carbon">TOC in water</label></br>

<input type="text" name="Organic\_carbon"/>

</div>

<div class="col">

<label for="Trihalomethane">Trihalomethane</label></br>

<input type="text" name="Trihalomethane"/>

</div>

</div>

<div class="row mb-3">

<div class="col">

<label for="Turbidity">Turbidity</label></br>

<input type="text" name="Turbidity"/>

</div>

<!--div class="col">

<label for="oldpeak">Exercise Relative To Rest</label></br>

<input type="text" name="oldpeak"/>

</div>

<div class="col">

<label for="slope">Peak Exercise</label></br>

<input type="text" name="slope"/>

</div>

<div class="col">

<label for="ca">Major Vessels</label></br>

<input type="text" name="ca"/>

</div>

CSE DEPT.Page 35



</div>

<div class="row mb-3">

<div class="col">

<label for="thal">Defect</label></br>

<input type="text" name="thal"/>

</div>

</div-->

<br>

<div class="row mb-3">

<div class="col">

<button type="submit" class="btn btn-primary">Submit</button>

</div>

</div>

<div class="row mb-3">

<div class="col">

<br>

</div>

</div>

</div>

</form>

{% block content %} {% if values==1 %}

<h1>Water is clean.It can be used for drinking purpose.</h1>

<image

align="center"

src="https://i0.wp.com/thumbs.gfycat.com/CelebratedEuphoricFantail-size\_restricted.gif"

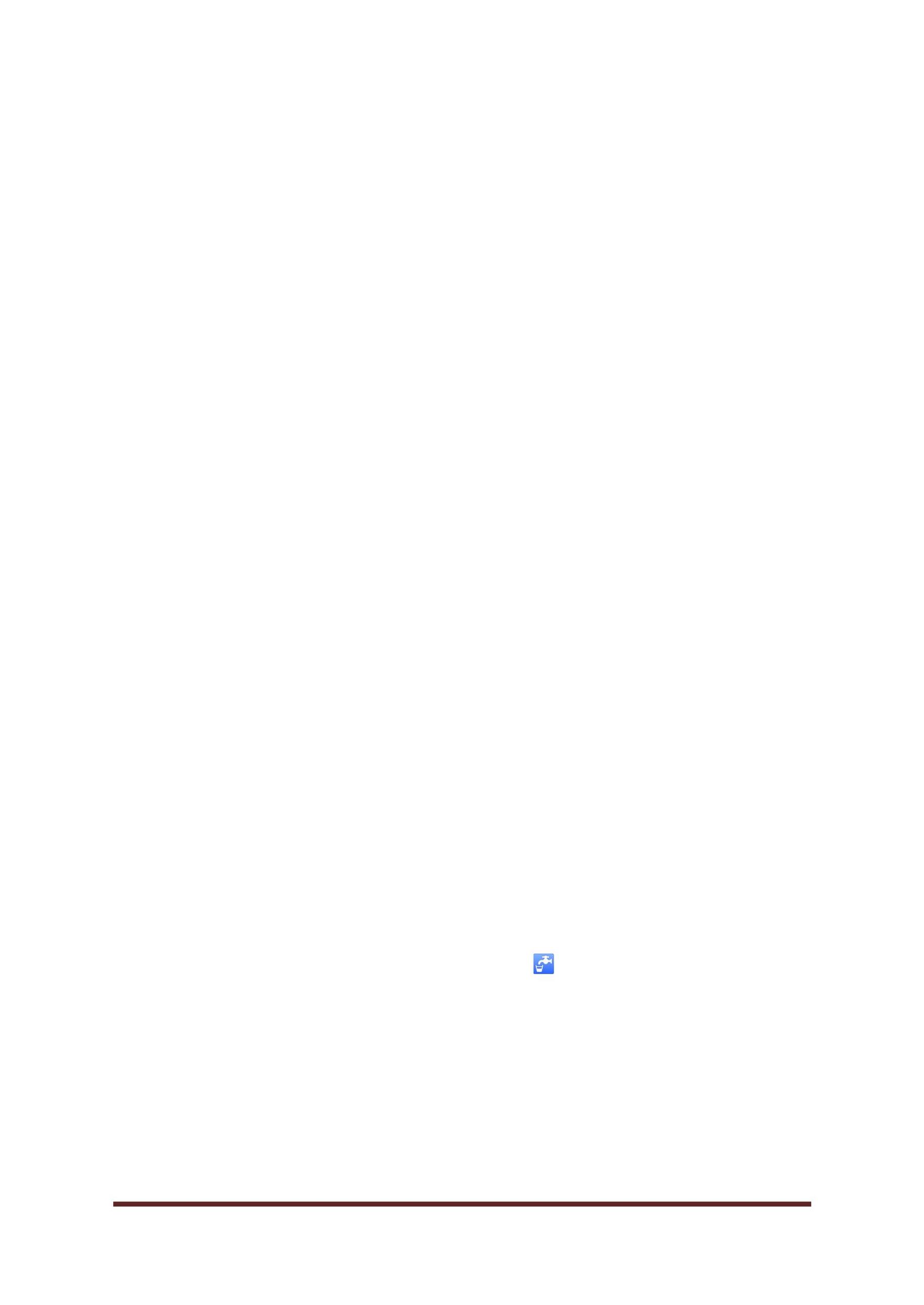
width=300 height=300></image>

{% endif %}

{% if values==0 %}

<h1>Water is not clean.It is not suitable for drinking. </h1>

CSE DEPT.Page 36



<a href="/purification" <button class="button button1" type="button">Check

to know water purifying methods...</button></a>

{% endif %} {% endblock content %}

</body>

</html>

**Purification Techniques**

<!DOCTYPE html>

<html>

<head>

<title>Purification Techniques</title>

<style>

body

{

background-image:url("https://image.shutterstock.com/image-illustration/light-blue-gradient-

color-background-260nw-1040861230.jpg");

background-repeat: no-repeat;

background-attachment:fixed;

background-size:100% 150%;

}

</style>

</head>

<body>

<h1 style="font-size: 40px"align="center"><u>Purification Techniques</u></h1>

<h2 style="font-size:25px"style="color:rgb(15, 15, 15)"><u> Boiling</u></h2>

<p style="font-size:20px">Boiling is the simplest of all water purification methods. It is

effective because

CSE DEPT.Page 37



it kills 99.9% of all living things and vaporizes most chemicals. It’s a good method to use

for water contaminated by living microorganisms.

The micro-organisms are eliminated by heating the water to its boiling point.Because

some bacteria can resist high temperatures,

you need to ensure the water boils for 3-7 minutes, depending on your altitude. For people

living in high altitude areas,

It is recommended to boil your water for longer than water boiled at lower altitudes. This

is because water boils at lower temperatures in higher altitudes.

After adequate boiling, let the water cool down. This is important to allow any solid

matter in it to settle at the bottom.

After cooling, the water is safe for human consumption.</p>

<image align=right src="https://puretecwater.com/images/Reverse-Osmosis-Color.png"

width=300 height=300></image>

<h2 style="font-size:25px"><u> Reverse Osmosis (RO)</u></h2>

<p style="font-size:20px">Reverse Osmosis, commonly referred to as RO, is the perfect

solution if you’re looking to improve the taste, odor,

and the appearance of your water. Reverse Osmosis is a method universally known for its

effective purification of seawater.

It is a process in which dissolved inorganic solids, such as fluoride, chlorine, lead, and

pesticides are removed from water by applying

pressure more than the osmotic pressure and pushing the water through a semipermeable

membrane.</br>

RO membrane is a semi-permeable membrane or special filter that allows the passage of

water molecules but not the majority of dissolved salts, organics, and bacteria.

RO works by using a high-pressure pump to increase the pressure of the contaminated

water pushing it across the semipermeable RO membrane.

To successfully remove these impurities, the pressure used has to be greater than the

naturally occurring osmotic pressure.

This pressured filtration ensures almost 99% of dissolved salts are sieved out. The amount

of pressure required depends on the concentration of contamination in the feed water.

The higher the concentration, the more the pressure required to overcome the osmotic

pressure.</p>

<h2 style="font-size:25px"style="color:rgb(15, 15, 15)"><u>Distillation</u></h2>

CSE DEPT.Page 38



<p style="font-size:20px">Distillation is one of the simplest water purification processes.

A water distillation system is simply designed

to purify your water in an inexpensive, quick, and effective way. The concept behind

purifying water through distillation is pretty simple.

This process has only two operations, namely evaporation, and condensation. You only

need two things to set up simple distillation: a heat source and a condenser.

It provides water with a clarity that is up to 98% free of impurities.</br>Distillation heats

the contaminated water until it reaches its relatively low boiling point and begins to vaporize.

Keep in mind that water has a lower boiling point than the contaminants and minerals

within it. Therefore, the heat of the water is kept

at this temperature to maintain the vaporization process while stopping other elements

from vaporizing as well.

That vapor is then funneled into a condenser in a separate area from where the process

started, where it cools and reverts to its liquid form.</p>

<image

align=right

src="https://www.aosmithindia.com/sites/default/files/1200px-Distillation\_Concept\_2.jpg"

width=300 height=300></image>

<h2 style="font-size:25px"style="color:rgb(15, 15, 15)"><u> Solar Purification</u></h2>

<p style="font-size:20px">An alternative to the UV filtration is solar purification which

involves treating water with the ultraviolet radiation of the sun.

The process involves filling a plastic bottle with water, shaking it to activate the oxygen

and leaving it horizontally in the sunlight.

This effectively kills bacteria and viruses present in the water, making it safe for

consumption.Solar water distillation uses a

solar still to condense pure water vapour and settle out harmful substances to make clean,

pure drinking water. This process is used when the water is

brackish containing harmful bacteria, or for settling out heavy metals and also for

desalination of sea water.</br></p>

<image

src="https://www.aosmithindia.com/sites/default/files/m3\_ss10\_fig10.6.jpg"

height=300></image>

align=left

width=300

<h2 style="font-size:25px"style="color:rgb(15, 15, 15)"><u>Clay Vessel

Filtration</u></h2>

CSE DEPT.Page 39



<p style="font-size:20px">Ceramic water filters (CWF) are an inexpensive and effective

type of water filter that rely on the small pore

size of ceramic material to filter dirt, debris, and bacteria out of water. This makes them

ideal for use in developing countries,

and portable ceramic filters are commonly used in backpacking.Way before people do not

have access to purifiers or purification techniques,

They used clay pots which purified muddy water, by blocking out the mud and allowing

pure, potable water to pass through.

Similar to other methods of filtering water, the filter removes particles larger than the size

of the pores in the filter material.

Typically bacteria, protozoa, and microbial cysts are removed. However, filters are

typically not effective against viruses

since they are small enough to pass through to the "clean" side of the filter. Ceramic water

filters (CWF) may be treated with silver

in a form that will not leach away. The silver helps to kill or incapacitate bacteria and

prevent the growth of mold and algae in the body of the filter. </p>

</body>

</html>

**Machine Learning Code**

import pandas as pd

from sklearn.utils import resample

from sklearn.model\_selection import train\_test\_split

from sklearn.model\_selection import RandomizedSearchCV, GridSearchCV

from sklearn.ensemble import RandomForestClassifier

import warnings

import pickle

warnings.filterwarnings("ignore")

dataset=pd.read\_csv('water\_potability.csv')

dataset= dataset.dropna()

CSE DEPT.Page 40



zero = dataset[dataset['Potability']==0]

one = dataset[dataset['Potability']==1]

dataset\_resample = resample(one, replace = True, n\_samples = 1200)

dataset = pd.concat([zero, dataset\_resample])

x=dataset.iloc[:,0:9].values

y=dataset.iloc[:,-1].values

x\_train,x\_test,y\_train,y\_test=train\_test\_split(x,y,test\_size=0.3,random\_state=0)

rf = RandomForestClassifier(criterion='entropy',random\_state=5)

params\_rf = {'n\_estimators':[100,200, 350, 500], 'min\_samples\_leaf':[2, 10, 30]}

grid\_rf = GridSearchCV(rf, param\_grid=params\_rf, cv=5)

rf.fit(x\_train,y\_train)

pickle.dump(rf,open('model.pkl','wb'))

model=pickle.load(open('model.pkl','rb'))

**Flask Code**

from flask import Flask,render\_template,request

import pickle

import numpy as np

app=Flask(\_name\_)

model=pickle.load(open('model.pkl','rb'))

@app.route('/')

def m1():

return render\_template('mini1.html')

@app.route('/mini2')

def m2():

return render\_template('mini2.html')

@app.route('/input')

CSE DEPT.Page 41



def m3():

return render\_template('input.html')

@app.route('/purification')

def m4():

return render\_template('purification.html')

@app.route('/predict', methods=['POST','GET'])

def predict():

ph=request.form.get('ph')

Hardness=request.form.get('Hardness')

Solids=request.form.get('Solids')

Chloramines=request.form.get('Chloramines')

Sulfate=request.form.get('Sulfate')

Conductivity=request.form.get('Conductivity')

Organic\_carbon=request.form.get('Organic\_carbon')

Trihalomethane=request.form.get('Trihalomethane')

Turbidity=request.form.get('Turbidity')

int\_features=[ph,Hardness,Solids,Chloramines,Sulfate,Conductivity,Organic\_carbon,Trihalo

methane,Turbidity]

final=[np.array(int\_features)]

prediction=model.predict(final)

output=round(prediction[0])

print(prediction)

return render\_template("input.html",values=output)

if \_name\_ == "\_main\_":

app.run(debug=True)

CSE DEPT.Page 42