

# **LITERATURE SURVEY ON EFFICIENT WATER QUALITY ANALYSIS AND PREDICTION USING MACHINE LEARNING**

**Mentor:**  
**Mr.A.Gopinath**

**Team Leader:**  
**Vidhya Lakshmi K**

**Team Members:**  
**Jeevitha Sai G**  
**Priyadharshini J**  
**Priyanga P**

# PROBLEM DEFINITION

Water is considered as a vital resource that affects various aspects of human health and lives. The quality of water is a major concern for people living in urban areas. The quality of water serves as a powerful environmental determinant and a foundation for the prevention and control of waterborne diseases. However predicting the urban water quality is a challenging task since the water quality varies in urban spaces non-linearly and depends on multiple factors, such as meteorology, water usage patterns, and land uses, so this project aims at building a Machine Learning (ML) model to Predict Water Quality by considering all water quality standard indicators.



# USE CASES

- To obtain quantitative information on the physical, chemical, and biological characteristics of water to analyze the water quality.
- To analyze the datasets to make predictions
- To determine an efficient and feasible way to obtain the ML predictions from quality analysis datasets.

# SURVEY PAPER-1

**Author:**

Amir Hamzeh Haghiabi, Ali Heidar Nasrolahi, Abbas Parsaie

**Title:**

Water quality prediction using machine learning methods

**Published Journal:**

Water Quality Research Journal

**Year of Published:**

02-2018



## Objective:

The study of water quality of rivers is a common theme in earth sciences. To evaluate the quality of rivers two approaches are considered, including measuring the water quality components and defining the mechanism of pollution transmission. Among water quality components, measuring the dissolved oxygen (DO), chemical oxygen demand (COD), biochemical oxygen demand (BOD), electrical conductivity (EC), pH, temperature, K, Na, Mg, etc. have been proposed.

## Critical Findings:

They stated that for developing the ANN, some steps should be considered to reduce the trial and error process. They stated that for the initial design of ANN model, after dataset division, in the first step one hidden layer consisting of numbers of neurons equal to input features is considered. At this stage, the performance of different transfer functions is evaluated and the best ones are chosen. In the next step, the size of the network is modified to improve the precision of the developed model. To this end, the numbers of neurons or number of hidden layers would increase. The last two stages of this approach are also applicable to the design of SVM

# SURVEY PAPER-2

**Author:**

Mourade Azrour, Jamal Mabrouki, Ghizlane Fattah, Azedine Guezzaz, Faissal Aziz

**Title:**

Machine learning algorithms for efficient water quality prediction

**Year of Published:**

06-2022





## **Objective:**

In this study, we take the advantages of machine learning algorithms to develop a model that is capable of predicting the water quality index and then the water quality class. The method They propose is based on four water parameters: temperature, pH, turbidity and coliforms. The use of the multiple regression algorithms has proven to be important and effective in predicting the water quality index.

## **Critical Findings:**

The method they propose is based on four water parameters: temperature, pH, turbidity and coliforms. The use of the multiple regression algorithms has proven to be important and effective in predicting the water quality index. In addition, the adoption of the artificial neural network provides the most highly efficient way to classify the water quality.

# SURVEY PAPER-3

**Author:**

Yafra Khan, Chai Soo See

**Title:**

Predicting and analyzing water quality using Machine Learning: A comprehensive model

**Published Journal:**

IEEE Long Island Systems, Applications and Technology Conference (LISAT)

**Year of Published:**

06-2016





## Objective:

The goal of this study is to develop a water quality prediction model with the help of water quality factors using Artificial Neural Network (ANN) and time-series analysis. For this paper, the data includes the measurements of 4 parameters which affect and influence water quality. For the purpose of evaluating the performance of model, the performance evaluation measures used are Mean-Squared Error (MSE), Root Mean-Squared Error (RMSE) and Regression Analysis.



# SURVEY PAPER-4

**Author:**

Umair Ahmed, Rafia Mumtaz, Hirra Anwar, Asad A. Shah, Rabia Irfan, Jose García-Nieto

**Title:**

Efficient Water Quality Prediction Using Supervised Machine Learning

**Published Journal:**

MDPI journal

**Year of Published:**

10-2019



## Objective:

This research explores a series of supervised machine learning algorithms to estimate the water quality index (WQI), which is a singular index to describe the general quality of water, and the water quality class (WQC), which is a distinctive class defined on the basis of the WQI. The proposed methodology employs four input parameters, namely, temperature, turbidity, pH and total dissolved solids. The proposed methodology achieves reasonable accuracy using a minimal number of parameters to validate the possibility of its use in real time water quality detection systems.



# SURVEY PAPER-5

**Author:**

water quality analysis using ML

**Title:**

water quality analysis using ML

**Published Journal:**

Science Direct

**Year of Published:**

11-2021



## Objective:

This research explores the Machine Learning (ML) algorithms for comparing AutoML and an expert architecture built by the authors for Water Quality Assessment to evaluate the Water Quality Index, which gives the general water quality, and the Water Quality Class, a term classified on the basis of the Water Quality Index.



# SURVEY PAPER-6

**Author:**

Manya Kakkar , Vansh Gupta , Jai Garg , Dr. Surender Dhiman

**Title:**

Detection of Water Quality using Machine Learning

**Published Journal:**

International Journal Of Engineering Research And Technology  
(IJERT)

**Year of Published:**

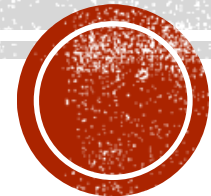
11-2021





## Objective:

The system makes use of IoT and Machine Learning technology. It consists of physical and chemical sensors that detect pH, Turbidity, Color, Dissolved Oxygen, Conductivity to check influencing factors. The data collected by the sensors is saved in a database and then submitted for analysis. The neural network method is used to forecast the outcome. It is employed in order to generate a non-linear connection for projected output. When any of the parameters falls below the standard values, the system sends an alarm notification to the user. This enables the user to be aware of water pollution in their home tanks ahead of time.



**THANK YOU!**

