LITERATURE SURVEY ON

EFFICIENT WATER QUALITY ANALYSIS AND PREDICTION USING MACHINE LEARNING

Domain : Applied Data Science

Team ID : PNT2022TMID29722

Batch No : B11-5A1E

Team Leader:

Swetha V (513119106090)

Team Members:

Nubaish Ahamed S (513119106061)

Vigneshwar V (513119106098)

Madhan Kumar M (513119106048)

PAPER 1

A SUPERVISED LEARNING APPROACH TO WATER QUALITY PARAMETER PREDICTION AND FAULT DETECTION

Publication Year: 2018

- **Author:** Kathleen Joslyn, John Lipon
- **Journal Name:** Institute of Electrical and Electronics Engineers
- **Summary:**

Water quality parameters such as dissolved oxygen and turbidity play a key role in policy decisions regarding the maintenance and use of the nation's major bodies of water. In particular, the United States Geological Survey (USGS) maintains a massive suite of sensors throughout the nation's waterways that are used to inform such decisions, with all data made available to the public. However, the corresponding measurements are regularly corrupted due to sensor faults, fouling, and de-calibration, and hence USGS scientists are forced to spend costly time and resources manually examining data to look for anomalies. It presents a method of automatically detecting such events using supervised machine learning. It first presents an extensive study of which water quality parameters can be reliably predicted, using support vector machines and gradient boosting algorithms for regression. It then shows that the trained predictors can be used to automatically detect sensor de-calibration, providing a system that could be easily deployed by the USGS to reduce the resources needed to maintain data fidelity.

- **Methodology Used:**

In this work, it has shown that two supervised regression algorithms methodology (support vector regression and gradient boosting) can be used to predict a variety of water quality factors with a high degree of accuracy.

- **References:**

https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8622628&isnumber=8621858

PAPER 2

A COMPLETE PROPOSED FRAMEWORK FOR COASTAL WATER QUALITY MONITORING SYSTEM WITH ALGAE PREDICTIVE MODEL

Publication Year: 2021

- **Author:** N. A. P. Rostam, N. H. A. H. Malim.
- **Journal Name:** Institute of Electrical and Electronics Engineers
- **Summary:**

An end-to-end process to achieve a complete framework methodology for Harmful Algal Bloom (HAB) growth prediction is crucial for water management, especially in implementing robust predictive modelling of HAB to prevent water pollution. This paper takes the initiative to provide a wider coverage on the end-to-end process including the assembly and integration of sensors, data acquisition and predictive modelling using data-driven approaches, for example, machine learning, deep learning and deep time series forecasting algorithm for future algal bloom outbreak mitigation. In the end, this paper presents proof that selecting the right features and utilizing time series with deep learning are much better for tackling the issues of highly non-linear and dynamic algae ecological data. In order to accurately predicting algal growth by the prediction of chlorophyll-a (Chl-a) as a strong indicator of algal presence for coastal.

- **Methodology Used:**

It uses Long Short-term Memory (LSTM) algorithm for prediction method and to outer performed other basic machine learning methods in accurately predicting algal growth.

- **References:**

https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9504580&isnumber=9312710

PAPER 3

DETECTION OF WATER QUALITY USING MACHINE LEARNING

Publication Year: 2011

- **Author:** Manya Kakkar, Vansh Gupta, Dr. Surender Dhiman
- **Journal Name:** International Journal of Engineering Research & Technology
- **Summary:**

In this paper, the study provides a frugal technique for eliminating water pollution in residential overhead tanks. Machine learning algorithms are used to forecast potential water pollution problems. To gather water parameters, the proposed system comprises multiple sensors interfaced with NodeMcU. The user is notified before the water turns contaminated. The implemented solution protects the water from pollution while still being cost efficient. The project's future scope includes detecting illnesses caused by numerous factors and devising the most effective plan to clean the tank.

- **Methodology Used:**

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. This technology is not restricted to home tanks; it may also be applied in water treatment facilities and enterprises.

- **References:**

https://www.ijert.org/detection-of-water-quality-using-machine-learning-and-iot#:~:text=The%20system%20makes%20use%20of,and%20then%20submitted%20for%20analysis.

PAPER 4

EFFICIENT PREDICTION OF WATER QUALITY INDEX(WQI) USING MACHINE LEARNING

Publication Year: 2021

- **Author:** Md. Mehedi Hassan, Laboni Akter
- **Journal Name:** International Journal of Computational Intelligence systems
- **Summary:**

In this StudyThe performance of machine learning techniques such as RF, NN, MLR, SVM, and BTM to predict the water quality components of an Indian water quality dataset was evaluated in this work. The most well-known dataset variables, such as BOD, DO, TC, Nitrate, PH and Temp, were obtained for this purpose. The findings revealed that the applied models performed well in forecasting water quality parameters; however, the greatest performance was linked with the MLR with Accuracy Upper. Further research will be done to build models that combine the proposed method with other techniques and deep learning approaches to improve the efficacy of the selection process.

- **Methodology Used:**

We have implemented five machine learning algorithms to train the dataset: Neural Network, Random Forest, Multinomial Logistic Regression, Support Vector Machine, and Bagged Tree Model, with six distinct color signals indicating the relevance of each feature.

- **References:**

https://www.atlantis -press.com/journals/hcis/125965714/view

PAPER 5

EFFICIENT WATER QUALITY PREDICTION USING SUPERVISED MACHINE LEARNING

Publication Year: 2019

- **Author:** Umair Ahmed, Rafia Mumtaz, Rabia Irfan and Jose Garcia-Nieto
- **Journal Name:** Multidisciplinary Digital Publishing Institute (MDPI)
- **Summary:**

This research explored an alternative method of machine learning to predict water quality using minimal and easily available water quality parameters. The data used to conduct the study were acquired from PCRWR and contained 663 samples from 12 different sources of Rawal Lake, Pakistan. A set of representative supervised machine learning algorithms were employed to estimate WQI. TDS for parameter readings and communicate those readings using an Arduino microcontroller and ZigBee transceiver. It would identify poor quality water before it is released for consumption and alert concerned authorities. It will hopefully result in curtailment of people consuming poor quality water and consequently de-escalate harrowing diseases like typhoid and diarrhea.

- **Methodology Used:**

When it comes to estimating water quality using machine learning algorithms namely, Support Vector Machines (SVM), Neural Networks (NN), Deep Neural Networks (Deep NN) and k Nearest Neighbors (KNN), with the highest accuracy of 93% with Deep NN. The estimated water quality in their work is based on only three parameters: turbidity, temperature and pH, which are tested according to World Health Organization (WHO) standards.

- **References:**

https://www.researchgate.net/publication/336808732_Efficient_Water_Quality_Prediction_Using_Supervised_Machine_Learning/link/5db2f58992851c577ec2d8ad/download

PAPER 6

PREDICTING WATER QUALITY USING WSN AND MACHINE LEARNING

Publication Year: 2020

- **Author:** Md.Saikat Islam Khan, Mostofa kamal Nasir, Sifatul Islam
- **Journal Name:** Research Gate
- **Summary:**

In this work, our goal was to use machine learning to detect whether the water is clean or polluted. Then, principal component regression is implemented to predict the WQI in the Gulshan lake. The water metrics including PH, DO, SS, EC, Turbidity, Chloride, COD, TDS, and Alkalinity are used for WQI prediction. The performance of the proposed PCR model is compared with the several existing support vector regressions and multilayer perceptron-based models.

- **Methodology Used:**

For this study, the WQI is calculated using the "weighted arithmetic index method and then a PCR model is constructed to estimate the WQI. PCR's basic concept is that PCA is applied to the dataset to minimize the dimension, and at the same time, regression algorithm is applied to the PCA output.

- **References:**

https://www.researchgate.net/publication/356496741

PAPER 7

A REAL TIME WATER QUALITY MONITORING USING MACHINE LEARNING ALGORITHM

Publication Year: 2020

- **Author:** S.Angel Vergina, Dr. S. Kayalvizhi, Dr. R.M.Bhavadharini.
- **Journal Name:** European Journal of Molecular &Clinical Medicines
- **Summary:**

In this paper, machine learning algorithm is used for predicting the quality of water. The predicted water quality data are stored in Cloud server for future access. The predicted data is sent to the water controller unit for further action. This has brought about complete computerized Water Quality Monitoring framework utilizing IoT and AI Technologies by which the gadgets impart among themselves in anticipating the Water Quality for private country region. Subsequently, the water quality can be watched consequently with no human interference. The proposed system can be extended further by water retreatment mechanism.

- **Methodology Used:**

IoT, Machine Learning and Cloud Computing technique are used.

- **References:**

https://ejmcm.com/article_4532_f6ab95e8dbd3dead7b98da14d1376c90.pdf

PAPER 8

GROUND WATER QUALITY PREDICTION USING MACHINE LEARNING ALGORITMS

Publication Year: 2019

- **Author:** S.Vijay & Dr.K.Kamaraj
- **Journal Name:** International Journal of Research and Analytical Reviews
- **Summary:**

The ground water plays a prime role in a country like India. In this paper, they proposed three classification algorithms like C5.0, Naive Bayes and Random Forest with data analytics tool to generate effective predictive model which predicts whether water is "High" of "Low" for drinking purpose based on water quality parameters. Naïve Bayes and Random forest produced better result with accuracy and classification error. In future we intend to use more classification algorithm with extended dataset to analyze the ground water quality Hence proper water treatment is required in terms of community health.

- **Methodology Used:**
  - Statistics, Data Analysis and Machine Learning.
  - It makes statistical computing easy and the programming effort is reduced.
- **Reference:**

http://ijrar.com/ijrar_issue_20543244.pdf

PAPER 9

PREDICTING AND ANALYZING WATER QUALITY USING MACHINE LEARNING

Publication Year: 2016

- **Author:** Yafra Khan, Chai Soo See
- **Journal Name:** Institute of Electrical and Electronics Engineers
- **Summary:**

In this paper survey of Different methodologies have been proposed and applied for analysis and monitoring of water quality as well as time series analysis. The methodologies range from statistical techniques, visual modeling, analysis algorithms and prediction algorithms and decision making. Ensure the quality of the water.

- **Methodology Used:**
  - The methodology used in this study comprises of Machine learning with training
  - Testing data from USGS online data repository
  - Artificial Neural Networks, Neural Network for Time Series.

- **Reference:**

https://ieeexplore.ieee.org/document/7494106/

PAPER 10

PREDICTIVE MODELS FOR RIVER WATER QUALITY USING MACHINE LEARNING AND BIG DATA TECHNIQUES

Publication Year: 2021

- **Author:** Jitha P Nair, M S Vijaya
- **Journal Name:** Institute of Electrical and Electronics Engineers IEEE
- **Summary:**

In this paper, due to contamination of water due to various factors can be discussed and the possible solution and analyses various prediction models developed using machine learning and big data techniques and their experimental results of water prediction and evaluation. Various challenges and issues are reviewed and possible solutions to some research issues are proposed.

- **Methodology Used:**
  - Big Data Analytics
  - Deep Learning and Machine Learning
  - Time Series, Water Quality Evaluation and Prediction
  - Water Quality Index.

- **Reference:**

https://ieeexplore.ieee.org/document/9016825

PAPER 11

PREDICTING WATER QUALITY PARAMETERS USING MACHINE LEARNING

Publication Year: 2019

- **Author:** G Manju, Ravishankar Holla, Nikhil M Ragi
- **Journal Name:** Institute of Electrical and Electronics Engineers
- **Summary:**

In this survey paper, instead of using traditional method of testing the water we use the brief methodology to predict unknown parameters such as Alkalinity, Chloride, Sulphate values using known parameters such as pH, Electrical Conductivity, TDC, etc. using Levenberg-Marquardt algorithm, which helps in further classification of water bodies for different application. Results gave accuracy of 83.94%, 87.9%, 81.736%, 79.48% in predicting chloride, total-hardness, sulphate, total alkalinity respectively.

- **Methodology Used:**
  - Artificial Neural Network (ANN) is used to solve this problem.
  - This method eliminates chemical method of evaluating water quality parameters.
  - It is also cost effective.
- **Reference:**

https://ieeexplore.ieee.org/document/9395832/

PAPER 12

EFFICIENT WATER QUALITY PREDICTION FOR INDIAN RIVERS USING MACHINE LEARNING

Publication Year: 2021

- **Author:** Yogalakshmi S, Mahalakshmi A
- **Journal Name:** Asian Journal of Applied Science and Technology (AJAST)
- **Summary:**

In this survay paper is mainly for conserve or alert for the future, because the world is polluting rapidly due to various pollution A series of representative supervised prediction calculations were tried on the dataset worked here. The total procedure is proposed with regards to water quality mathematical investigation.

- **Methodology Used:**
  - C machine learning techniques
  - Determining water quality with minimal parameters
  - regression algorithms.

- **References:**

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3829041