LITERATURE SURVEY

S.NO	RESEARCH PAPER
1.	Water Quality Analysis and Prediction using Machine Learning Algorithm
2.	Predictive Modeling Approach for Surface Water Quality: Development and Comparison of Machine Learning Models
3.	Comparative Assessment of Individual and Ensemble Machine Learning Models for Efficient Analysis of River Water Quality

Survey Done By:

Dinesh T

Harikrishnan M

Arvind J

Ashok N

TITLE: Water Quality Analysis and Prediction using Machine Learning Algorithm

AUTHOR: Mr.M.Anbuchezhian, Dr.R.Venkataraman, Mrs.V.Kumuthavalli, M.C.A., M.Phil.,

The analysis of water quality is compound problem due to the various factors influence in it. In particular, this concept is intrinsically tied to the different intended uses of the water .different uses require different criteria. Lot of research works going on the prediction of water quality. Normally water quality must be defined based on a set of physical and chemical variables that are closely related to the water's intended use. For each variable, acceptable and unacceptable values must then be defined. Water whose variables meet the pre-established standards for a given use is considered suitable for that use. If the water fails to meet these standard for a given use is considered suitable for that use. If the water fails to meet these standards. Many physical and chemical characteristics can be used to evaluate water quality or the degree of water pollution. Therefore, it is not possible in practice to clearly define water quality either on a spatial or temporal basis by separately examining the behaviour of every individual variable. The alternative, which is also difficult, consists of integrating the values of a set of physical and chemical variables into a unique value (i.e., an overall or global index). A water quality index is defined as a quality index for any use of water by simply determining the specifications.

These functions were defined using direct measurements of the concentration of a substance or the value of a physical variable obtained through analyses of water samples. The main theme of this paper is to make an analysis of water quality predication using machine learning algorithms with eight kind of parameters such as temperature (Temp), Dissolved Oxygen (DO) (% sat), pH, conductivity, Biochemical oxygen demand (BOD), nitrates (NO3), facal and total coli forms (TC).

The water quality index (WQI) is the most commonly used method to classify and communicate existing water quality. In this method, water quality data collected from large and Complex water quality monitoring programs are converted into single numeric values. The WQI values range between 0 and 100.

LINK: https://www.jetir.org/papers/JETIR1811966

TITLE: Predictive Modeling Approach for Surface Water Quality: Development and Comparison of Machine Learning Models

AUTHOR: Muhammad Izhar Shah, Wesam Salah Alaloul, Abdulaziz Alqahtani, Ali Aldrees, Muhammad Ali Musarat and Muhammad Faisal Javed

Human health, environmental services, and agricultural production are all at risk due to the growing global problem of water pollution. A detailed understanding of the developing concerns about water quality can be gained through the unique characteristics of artificial intelligence (AI) based modelling. The current study examines the accuracy of monthly total dissolved solids (TDS) and specific conductivity (EC) models for the upper Indus River at two outlet stations using gene expression programming (GEP), artificial neural networks (ANN), and linear regression models (LRM). 360 TDS and EC monthly records spanning 30 years of historical water quality data were utilised to train and evaluate the models. The TDS and EC modelling were associated with seven input factors based on a substantial correlation. Various performance measure indicators, error evaluation, and external criteria were all used to analyse the results. When the models' simulated results were compared to actual data, they showed a good correlation, with both TDS and EC showing correlation coefficients above 0.9.

The GEP and ANN models continued to be the effective methods for foretelling TDS and EC. The mathematical equations for the formulated GEP show its originality as compared to ANN and LRM.

The results of the parametric analysis showed that the modelling procedure had taken into account the effects of all the input factors. The generalised outcome and robustness of the suggested approaches were confirmed by the external assessment standards. In conclusion, the findings of this study showed that developing AI- based models for river water quality evaluation, management, and policy making is both affordable and beneficial.

LINK: https://www.mdpi.com/2071-1050/13/14/7515

TITLE: Comparative Assessment of Individual and Ensemble Machine Learning Models for Efficient Analysis of River Water Quality

AUTHOR: Abdulaziz Alqahtani, Muhammad Izhar Shah, Ali Aldrees, and Muhammad Faisal Javed

The outcomes of the parametric analysis demonstrated that the modelling process had accounted for the effects of every input element. The external assessment criteria confirmed the overall result and reliability of the offered approaches. The results of this study demonstrated that creating AI-based models for river water quality assessment, management, and policy making is both feasible and advantageous.

Seven input parameters were chosen for the dataset of the projected models' training and testing on the basis of their significant association. The ensemble RF model was optimised by creating 20 sub-models and selecting the most accurate one. Known statistical measures including the coefficient of determination (R2), mean absolute error (MAE), root mean squared error (RMSE), and Nash-Sutcliffe efficiency were used to evaluate the models' goodness-of-fit (NSE). The R2 value for the GEP, RF, and ANN models, respectively, was found to be 0.96, 0.98, and 0.92, demonstrating a significant correlation between inputs and modelling outputs. The comparative effectiveness of the suggested methodologies demonstrated the RF's relative superiority over GEP and ANN. The most accurate model among the 20 RF sub-models produced R2 values of 0.941 and 0.938, with 70 and 160 numbers of corresponding estimators. On training and testing data, respectively, the ensemble RF model produced the lowest RMSE values of 1.37 and 3.1.

The generalised outcomes of all the aforementioned procedures were ensured by the models' evaluation on outside criteria. As a result of the current study, it was concluded that the RF model with a few essential parameters might be prioritised for water quality assessment and management.

LINK: https://www.mdpi.com/2071-1050/14/3/1183