



## OPEN Reliable water quality prediction and parametric analysis using explainable AI models

M. K. Nallakaruppan<sup>1</sup>, E. Gangadevi<sup>2</sup>, M. Lawanya Shri<sup>1</sup>, Balamurugan Balusamy<sup>3</sup>, Sweta Bhattacharya<sup>1</sup> & Shitharth Selvarajan<sup>4,5</sup>✉

The consumption of water constitutes the physical health of most of the living species and hence management of its purity and quality is extremely essential as contaminated water has to potential to create adverse health and environmental consequences. This creates the dire necessity to measure, control and monitor the quality of water. The primary contaminant present in water is Total Dissolved Solids (TDS), which is hard to filter out. There are various substances apart from mere solids such as potassium, sodium, chlorides, lead, nitrate, cadmium, arsenic and other pollutants. The proposed work aims to provide the automation of water quality estimation through Artificial Intelligence and uses Explainable Artificial Intelligence (XAI) for the explanation of the most significant parameters contributing towards the potability of water and the estimation of the impurities. XAI has the transparency and justifiability as a white-box model since the Machine Learning (ML) model is black-box and unable to describe the reasoning behind the ML classification. The proposed work uses various ML models such as Logistic Regression, Support Vector Machine (SVM), Gaussian Naive Bayes, Decision Tree (DT) and Random Forest (RF) to classify whether the water is drinkable. The various representations of XAI such as force plot, test patch, summary plot, dependency plot and decision plot generated in SHAPELY explainer explain the significant features, prediction score, feature importance and justification behind the water quality estimation. The RF classifier is selected for the explanation and yields optimum Accuracy and F1-Score of 0.9999, with Precision and Re-call of 0.9997 and 0.998 respectively. Thus, the work is an exploratory analysis of the estimation and management of water quality with indicators associated with their significance. This work is an emerging research at present with a vision of addressing the water quality for the future as well.

**Keywords** Water Quality, Potability, TDS, Explainable AI, Water Quality Prediction

The major part of our earth comprises water and it is extremely important for the survival of all humans and animal species. Water makes up over 326 cubic metres of the planet's surface, which is almost 71% of its total area out of which 97% is seawater. Only 0.5 percentage of the drinkable water on earth is accessible, while the remaining 2.5 percentage is either trapped in glaciers, polar ice caps, the atmosphere, on soil, is polluted, or lies beneath the earth's surface far beyond human reach. If the global water supply is 100 L, consequently the amount of drinking water would be only 0.003 L, which is just a teaspoon. Therefore, the management and preservation of drinking water is regarded as a top priority. It is the most critical issue for mankind to address given the extremely limited amount of water that is accessible for use. The quantum of water around the world is represented in Table 1.

Water is a common and crucial resource shared among all humans, animals, and plants and is a necessity for all species. Each one of these species has its own respective needs for water quality. Total Dissolvable Solids (TDS) of soft water for human consumption range from the best quality stated, which is between 50 mg/dL and 150 mg/dL. Between 150 mg/dL and 300 mg/dL is the next level that can be applied to humans. The plants need water that is between 700mg/dL and 800mg/dL. The animals, especially cattle consume water around the quality of 1000 mg/dL. It is thus evident from all these observations that water quality management is essential to ensure

<sup>1</sup>School of Computer Science Engineering and Information Systems, Vellore Institute of Technology, Vellore 632014, India. <sup>2</sup>Department of Computer Science, Loyola College, Chennai, Tamil Nadu 600034, India. <sup>3</sup>Shiv Nadar University, Delhi-NCR 201314, India. <sup>4</sup>School of Built Environment, Engineering and Computing, Leeds Beckett University, Leeds LS13HE, UK. <sup>5</sup>Department of Computer Science, Kebri Dehar University, Kebri Dehar, Ethiopia. ✉email: ShitharthS@kdu.edu.et

Location	Quantity (%)
Oceans	97.2
Ice Caps/Glaciers	2.0
Groundwater	0.62
Freshwater Lakes	0.009
Inland seas/salt lakes	0.008
Atmosphere	0.001
Rivers	0.0001

**Table 1.** Water availability around the globe.

sustainability and a healthy life on Earth. The impact of water quality prediction is crucial at a global level for many reasons. First of all, to get clean and safe water is a basic human necessity and water quality prediction aids to guarantee the availability of potable water for societies worldwide. Water quality is related to public health as polluted water may cause waterborne diseases which could affect millions of humans globally. A sustainable environment is an important aspect of human well-being by preserving ecosystems and biodiversity. The significance of water quality assessment is profound and intricate by various organizations globally. The WHO (World Health Organization), UNEP (United Nations Environment Programme), EPA (United States Environmental Protection Agency), EEA (European Environment Agency), IWA (International Water Association) and WEF (Water Environment Federation) are fanatical for water quality assessment and addressing the mitigation strategies for water quality challenges. Water quality creates impact on public health globally and resulting in dissemination of waterborne diseases like typhoid, dysentery, cholera, dengue and malaria and cause substantial risks worldwide.

The advancement in computing technologies and artificial intelligence have elevated the standards of water quality assessments<sup>1</sup>. Measurements and estimations about the quality of the water have become easier to calculate and accurate, especially with the development of Industry 4.0 standards and Internet of Things (IoT) sensors. With the integration of IoT sensors, AI solely serves as a supporting tool to automate water quality checks<sup>2</sup>. Classification and Regression models based on machine learning help in determining the water quality. Depending on the outcomes, classification results tend to be binary or multi-classified. Real-time sensor data are collected, given feature labels, and then classified based on the importance of the feature labels. Earlier, these measurements used to be carried out with fuzzy-based decision support systems<sup>3</sup> with subjective decision-making models. AI development has made it possible to classify and analyse quality aspects quantitatively. The accuracy of the water quality assessment has been validated using various performance metrics like accuracy, precision, recall, and f1-score. AI models also support such quantitative analysis, classification of water sources, and prediction of drinkable water as well as identifying the mixing of buoyant pollutants in water sources<sup>4</sup>.

Despite its success in automating tasks and making water quality predictions using diverse models, the AI models lack transparency and are considered black-box where the decisions are derived but the reasoning behind such decisions is not revealed. The present generation validation frameworks for water quality management need justifiability, transparency and explainability, which is possible to be rendered by Explainable AI (XAI) based systems. XAI is a technology that is white-box and answers the uncertainty related to the classification and regression problems of AI. XAI applies a model-agnostic approach, where the machine learning models can be treated independently for interpretation. Additionally, XAI discusses how the model is chosen, how it works, and how it performs categorization. Through the assessment of a problem's feature weights, XAI also can determine a feature's relevance. This clarifies how a feature value relates to a certain target class classification. As an example, XAI uses models like Partial Dependency Plots (PDP)<sup>5</sup>, which describes the relationship between the features using lasso functions. This model may identify the linear relationship between two characteristics of water quality data and explain their correlation. In XAI, models like Local Interpretable Model agnostic Explainer (LIME), explain the relationship between a single feature and relevant others in local surrogacy. This infers that, except for the one-row value of the dataset, it is possible to relate a target attribute to the other independent variables. LIME in this regard can be used to explain the target classification for a single row instance about the water quality<sup>6</sup>. In the proposed work, XAI, which employs both local and global surrogates, includes SHAPLEY. The model offers a solution that takes into account the importance of each feature in determining the target as well as the dependency between features, the relationship between features, and the explanation of decisions through a variety of plots, including force plots, summary plots, dependency plots, and decision plots. The framework is very adaptable and capable of giving a thorough explanation of the characteristics of the water quality and how they affect the classification of the water quality.

### Advantages of the proposed model

Explainable AI plays an important role in improving the interpretability of predictions made by machine learning models. More transparent predictions are generated by these models. In the proposed approach, the authors have employed LIME and SHAP to interpret predictions achieved from machine learning, which identifies inputs as an important metric for selecting the features. By applying the XAI approach, the proposed model provides deep insights into the features and allows informed decision-making in water management processes.

## Contributions of the paper

The following points describe the contribution of the proposed work.

- The proposed work offers a comprehensive analysis and white-box description of the classification problem for water quality.
- The framework incorporates extensive pre-processing of the dataset to ensure it fit to be fed into the XAI model.
- Imputation of missing data is carried out to increase the accuracy of the findings.
- The proposed work ensures achievement of most significant features, identification of the feature importance, feature dependencies, and feature weights, that enable optimized classification of water quality dataset.
- The proposed approach employs both model-based and model-agnostic interpretations, using model-based ML implementations and model-agnostic XAI implementations.

## Organization of the paper

Section “[Introduction](#)” of the paper introduces the problem of the research paper with the description of the unique contributions. Section. [Introduction](#)” also describes the literature review of the related problems on water quality, in related works subsection, with an exhaustive survey of the various applications and case studies pertaining to water quality management using AI and machine learning approaches. Section “[System model and architecture](#)” describes the methods applied in the proposed work with the implementation of the mathematical model with the algorithm of the proposed work. Section “[Results](#)” describes the results of various ML and XAI models with relevant tables and graphs. Section “[Discussion](#)” provides the comparative analysis of the results with a discussion of challenges and solutions of the proposed work. Section “[Conclusion](#)” concludes the paper with future directions.

## Related works

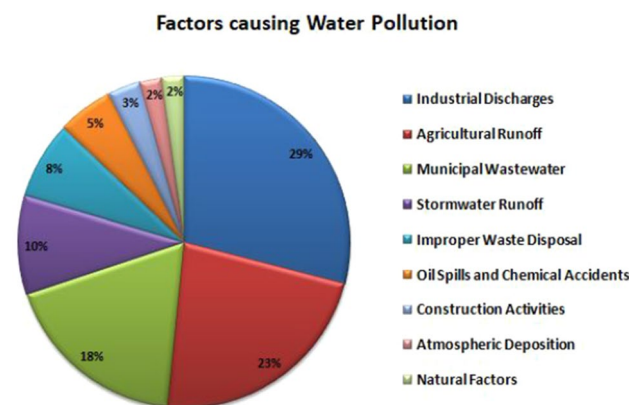
Lu et al.<sup>7</sup> proposed the central environmental protection inspection (CEPI), which was implemented and the causes of transboundary water contamination were investigated. The triple difference technique (DDD) was used to assess how the CEPI affected pollution and the results to determine how significantly water pollution was decreased as well as the significance of CEPI laws for addressing transboundary pollution. Halder et al.<sup>8</sup>, the Turag River’s neighbouring communities are suffering from major health problems as a result of water contamination. For the sustainability of household and aquatic life, the river’s water quality was unsuitable. The study noted that the threshold values for turbidity, total dissolved solids (TDS), chloride (CL-), chemical oxygen demand (COD), carbon dioxide (CO<sub>2</sub>), and biochemical oxygen demand (BOD) are higher than the standard permissible limits, which may result in health problems like respiratory illnesses, diarrhoea, cholera, dengue, malaria, anaemia, and skin problems. A study evaluating metal pollution management and mitigation tactics on soil and water was presented by Wang et al.<sup>9</sup>. In this study, the remediation of metal contamination from water and soil utilising chemical, physical, and biological approaches was discussed. In this study, the current methods for reducing heavy metal pollution of the soil and water are examined. Elehinafe et al.<sup>10</sup> discussed the importance of water contamination and examined the main cause of water scarcity. The proposed work discussed the effect of hazardous chemicals on the water, including pesticides, heavy metals, and micro-pollutants. This study outlined the numerous technologies that are currently available to eliminate hazardous materials and provide sustainable clean water resources. Mu et al.<sup>11</sup> proposed a solution for the investigation into farmers’ readiness to implement Rural Water Pollution Control (RWPC). This study examines farmers’ viewpoints to improve the quality of life for locals who reside in rural regions and avoid water contamination. To analyse the contributions of contaminants, Wang et al.<sup>12</sup> developed a unique contaminant flux variable model for river water quality assessment. The framework effectively identified the sources of pollution and evaluated the efficacy of projects designed to reduce water pollution. Zadeh et al.<sup>13</sup> proposed WQPs for estimating chemical oxygen demand and biochemical oxygen demand using the MKSVR algorithm. PSO algorithm is used for solving optimization problems. The multiple kernel support vector regression (MKSVR) is compared with SVR and Random Forest Regression and achieves a better accuracy level for BOD prediction. Nagaf et al.<sup>14</sup> presented a framework for assessing the WQI values based on the NSF guidelines. This framework uses four data-driven models such as EPR, M5 MT, GEP and MARS for predicting WQI values in the Karun River. The classification uses 12 water quality parameters and missing values were extracted from the image analysis. Zadeh et al.<sup>15</sup> proposed a model that utilizes gene expression programming, evolutionary polynomial regression, and model trees for predicting WQPs. The biochemical oxygen demand, dissolved oxygen and chemical oxygen demand are used for estimation with nine parameters. The gamma test is used for determining important parameters. Najaf et al.<sup>16</sup> proposed a water quality predicting framework for estimating the water quality index in the Hudson River based on Canadian Council of Ministers of the Environment (CCME) guidelines. The four artificial intelligence techniques M5 MT, Multivariate Adaptive Regression Spline, Evolutionary Polynomial Regression and Gene Expression Programming are used with Landsat 8 OLI-TIRS images. The results proved that the MARS technique achieved the best outcome compared to other models.

Chowdhury et al.<sup>17</sup> emphasized the sources of water contamination which are caused by densely populated industrial areas that are located close to water bodies. The main causes of water contamination are dangerous chemicals and heavy metals. Farmers’ pre-owned pesticides, including different types of carbamate and organophosphorus pesticides, are the main causes of water contamination on agricultural grounds as per the study. Ahivar et al.<sup>18</sup> examined the use of heavy metal pollution indices (HPIs) in soil, water, and sediments. For assessing metal contamination, HPI is considered a crucial instrument. Each method’s pollution index is assessed to interpret the pollution levels. The selection of HPIs based on the parameters and standards for evaluating

the quality of the water and soil is offered. Chen et al.<sup>19</sup> presented a study by used various mathematical and statistical approaches to check the quality of water. The factors indicating the water pollution and the seasonal characteristics are evaluated to reduce the river water pollution. The Principal Component Analysis, Cluster Analysis, Network Analysis and Co-Occurrence Analysis were carried out to find the potential source of river water pollution. Fan et al.<sup>20</sup> examined the quality of water using several mathematical and statistical techniques. To lessen river water pollution, the variables implicating contamination and the seasonal traits are assessed. To identify a likely cause of river water pollution, the Principal Component Analysis, Cluster Analysis, Network Analysis, and Co-Occurrence Analysis were performed. Wang et al.<sup>21</sup> formulated the performance indices for explaining the Water-Energy-Pollution nexus (InWEP) effects of scales. The Nexus Pressure Index (NPI) and Nexus Coupling Index (NCI) were used to represent the pollution pressure and the interacted relations. The factors for InWEP were analysed using the Structural Equation Model (SEM) considering four objects namely enterprises, countries, industrial zones and cities. The performance of InWEP was evaluated for the performance metrics - efficiency, structure and location. To evaluate the quality of groundwater surrounding nearby areas in an industrial metropolis, Asomaku<sup>22</sup> evaluated the water pollution indices. Nine samples from three landfills are used in the analysis of the groundwater's chemical and metal characteristics. The study in Balaram et al.<sup>23</sup> explored many elements that have an impact on water quality, including climate change, industry, aquaculture, mining, and agriculture. For the quantitative and qualitative evaluation of hazardous metals, metal species, isotopes, and other contaminants that are present in water, various ICP-MS techniques are applied. Yuan et al.<sup>24</sup> proposed a water quality monitoring framework using biological sensors for water quality assessment. Borzooei et al.<sup>25</sup> presented a study to estimate the frequency weather events that creates impact on waste water assessment. The Time series data mining approach is used for categorizing the dry and wet weather events. Noori et al.<sup>26</sup> presented a report on decline of groundwater recharge in Iran. The study presents the average amount of ground water recharge is more than the annual runoff<sup>4</sup> utilized WCSPH (A weakly compressible smoothed particle hydrodynamics) model for simulating the near-shore hydrodynamics. The study conducted experimental and numerical evaluation for detecting the causes for mixing the buoyant pollutants in coastal water source. Yeganeh-Bakhtiar<sup>27</sup> presented a framework using MOS (Model Output Statistics) for establishing the statistical relationships among predictor and predicant.

When evaluating water quality using factors like toxicity and pollutants, computer vision and biological sensor systems are utilised in tandem. To retrieve the important data from images taken by a microscope, a microfluidic chip with sensors is utilised. This chip monitors water samples. Figure 1 describes various factors causing water pollution in smart cities including construction activities, atmospheric deposition, natural factors, municipal wastewater, stormwater runoff, incorrect waste disposal, industrial discharges, agricultural runoff, and municipal wastewater. Jaihouni et al.<sup>28</sup> implemented and compared five data mining techniques, including the Ordinary Decision Tree (ODT), Random Forest (RF), Chi-square Automatic Interaction Detector (CHAID), Iterative Dichotomiser 3 (ID3), and Random tree, to identify high-quality water zones. Eight parameters are used in the evaluation process while deriving rules. Compared to the remaining models, the RF performed well, with an accuracy rate of 97.10%. Lee et al.<sup>29</sup> implemented a framework for evaluating the quality of groundwater utilising a Self-Organizing Map (SOM) technique and fuzzy c-means clustering (FCM) was given. The two methods are employed to describe the complex nature of groundwater. SOM employed 91 neurons to categorise 343 groundwater samples, and FCM grouped the water sources into three groups. Agarwal et al.<sup>30</sup> proposed AI based water evaluating technique to predict the water quality index using Particle Swarm Optimization (PSO), Naïve Bayes Classifier (NBC), and Support vector machine (SVM). PSO was used in this regard for optimizing the classifiers wherein the PSO-optimized NBC obtained 92.8% accuracy and PSO-optimized SVM obtained 77.60% accuracy. Table 3 illustrates various existing state-of-art techniques proposed for assessing water quality, its advantages and research gaps.

Figure 1 illustrates the factors causing water pollution. The factors includes Industrial discharges, agricultural runoff, municipal waste water, storm water, improper waste disposal, oil spills and chemical spills, construction wastages, and atmospheric deposition. The factors are very crucial to protect public health and ecosystem , sustainability development, creating public awareness and for pollution prevention.



**Figure 1.** Factors causing water pollution.

Figure 2 depicts the required physical parameters such as Temperature, Turbidity, Conductivity, Odour and Color represented in percentage, for evaluating the quality of water. Examining the physical parameters is essential for identifying the potential hazards that leads to poor water quality and for preventing ecosystem health.

Figure 3 depicts the necessary chemical parameters, such as pH, Dissolved Oxygen (DO), Total Dissolved Solids (TDS), Nutrients (nitrogen and phosphorus), Total Suspended Solids (TSS), Heavy Metals, and Organic Matter (OM), as well as Chemical Oxygen Demand (COD) and Biochemical Oxygen Demand (BOD) with percentages, that must be measured in order to assess the water's quality.

Figure 4 presents various supervised learning models for estimating water quality, including Random Forest, Support Vector Machine (SVM), Decision Trees, Neural Networks, and Gradient Boosting Approaches like XGBoost and AdaBoost.

Figure 5 represents various unsupervised learning models such as Principal Component Analysis, Cluster Analysis and Self-Organizing Maps (SOM) for addressing the quality of the water. PCA is a dimensionality reduction approach mainly utilized for analyzing the high dimensional datasets. Cluster analysis techniques are used primarily for grouping water samples based on similarities. SOM technique is principally used for organizing the water quality data.

Figure 6 highlights the various Hybrid ML models such as ensemble models with Reinforcement Learning (RL) for addressing the evaluation of quality of water. The various machine learning models can be verified based on the applications, parameters in order to determine the quality of the water, dataset size and its quality based on the assessment of the performance metrics.

The motivation for the proposed research, along with the research gap analysis with similar existing research works is discussed as per Table 2. The comparative analysis and research of similar existing works are presented in Table 3. These two discussions provide a comprehensive understanding of the requirements, that are essentially required in the design of the proposed system and implementation.

Table 3 refers to similar literature review of various models of machine learning such as DT,RF,DCF, SVM, and so on. This table also discusses about various deep learning models such as, Artificial Neural Networks (ANN), Probabilistic Neural Network (PNN), Convolution Neural Networks (CNN) and statistical regression models

### Physical Parameters for Water Quality Evaluation

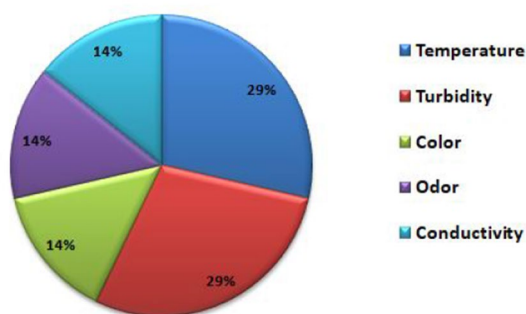


Figure 2. Physical Parameters.

### Chemical Parameters for Water Quality Evaluation

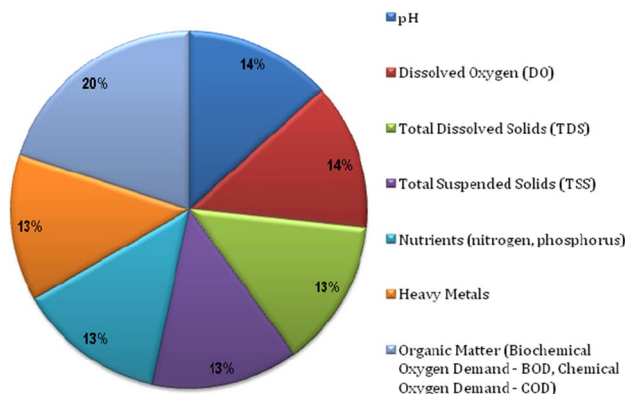


Figure 3. Chemical parameters.



### Supervised Learning Models for Water Quality Evaluation

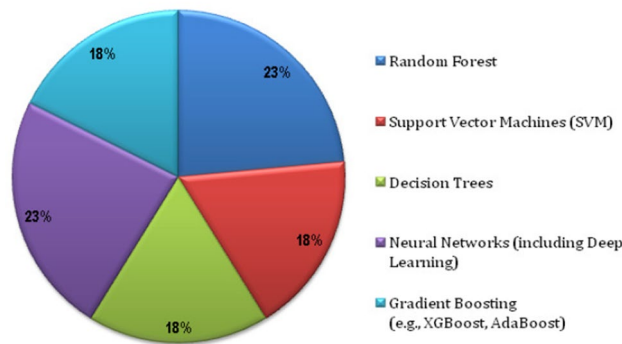


Figure 4. Supervised learning models.

### Unsupervised Learning Models for Water Quality Evaluation

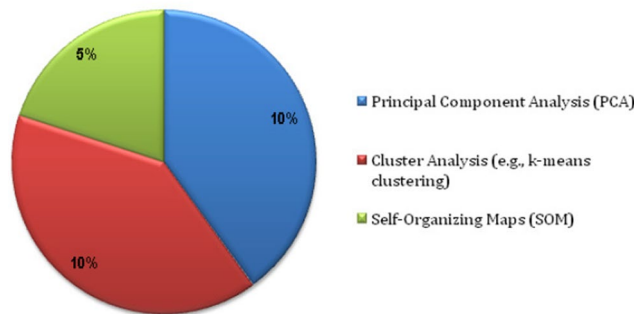


Figure 5. Unsupervised learning models.

### Hybrid Models for Water Quality Evaluation

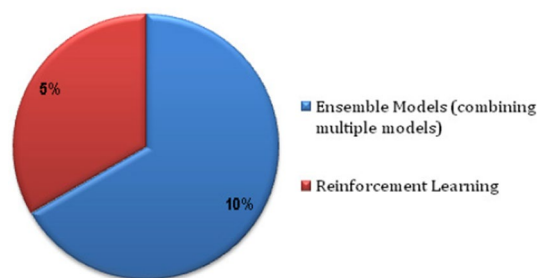


Figure 6. Hybrid ML models.

such as Auto-Regression in Moving Average (ARIMA). This table discusses the research gaps identified and enhanced in the proposed work. These models were mostly numerical evaluations with regression analysis. The proposed model and the system is classifier which deploys XAI framework, to discuss the impact of parameters, that determine the portability of the water with end user perspective. This is towards achieving environmental sustainability on water conservation and harvesting.

#### Statement of objectives

The proposed work offers a comprehensive analysis and white-box description of the classification problem for water quality. The framework incorporates extensive pre-processing of the dataset to ensure it fits into the XAI model. Imputation of missing data is carried out to increase the accuracy of the findings. The proposed

Refs	Title	Advantages	Research gap	Quality parameters
24	A biological sensor system using computer vision for water quality monitoring	• Real-time analysis	• Sensor Sensitivity and Selectivity	• Abstract Fish Behavior
		• Cost Effectiveness	• Robustness and Long term stability	• Movement Velocity
		• Versatility	• Calibration	• Rotation Angle Of The Fish Group
		• Potential for Automation	• Image Analysis and Algorithm Optimization	
31	Monthly water quality forecasting and uncertainty assessment via bootstrapped wavelet neural networks under missing data for Harbin, China	• Accurate Water quality forecasting	• Model Generalization	• Ammonia Nitrogen (NH <sub>4</sub> +N)
		• Handling Missing data	• Comparison with other Models	• Dissolved Oxygen
		• Uncertainty Assessment	• Data Availability and Quality	
		• Temporal Resolution	• Model Optimization	
32	Groundwater quality forecasting using ML algorithms for irrigation purposes	• Spatial distribution mapping	• Spatial and temporal scale	• Total Dissolved Solid (TDS)
		• Effective feature selection	• Data availability and quality	• Potential Salinity (PS)
		• High dimensional data	• Uncertainty estimation	• Sodium Adsorption Ratio (SAR)
		• Non-linearity and flexibility	• Validation and comparison	• Exchangeable Sodium Percentage
			• Transparency	• Magnesium Adsorption Ratio (MAR)
33	Predicting nitrate concentration and its spatial distribution in groundwater resources using Support Vector Machines (SVM) model	• Accurate Water quality forecasting	• Model Generalization	• Water Temperature
		• Handling Missing data	• Comparison with other Models	• Electrical Conductivity
		• Uncertainty Assessment	• Data Availability and Quality	• Groundwater Depth
		• Temporal Resolution	• Model Optimization	• Total Dissolved Solids
		• Site-Specific Application		• Dissolved Oxygen
				• Ph
34	A novel ML-based approach for the risk assessment of nitrate groundwater contamination	• Risk assessment accuracy	• Dataset limitations	Groundwater Vulnerability Map : (DI<80), low (DI=80-120), moderate (DI=120-160), high (DI=160-200), and very high
		• Ability to handle complex datasets	• Can exhibit temporal dynamics	
		• Spatially explicit risk mapping	• Validation	
		• Transferability to different regions	• Uncertainty quantification	
35	Machine Learning predictions of nitrate in groundwater used for drinking supply in the conterminous of the United States			• Comparative analysis
		• Nationwide assessment	• Data quality and availability	• High Precipitation
		• Accuracy and predictive power	• Incorporating temporal dynamics	• Recharge
		• Detection of the high risk areas	• Transferability and regional variability	• Base Flow Index
		• Spatially explicit predictions	• Uncertainty estimation	• Nitrate Concentrations
36	Ensemble modelling framework for groundwater level prediction in urban areas of India		• Comparative analysis	
		• Model training and calibration	• Data quality and availability	• Groundwater Levels
		• Ensemble generation	• Transferability and regional variability	• Rainfall, Temperature
		• Uncertainty estimation	• Validation benchmarking	• NOI
		• Enhanced pre-processing techniques	• Comparative analysis	• SOI
		• NIN		
		• Monthly Population Growth Rate		

**Table 2.** Motivation for the proposed work from the review perspective.

work ensures the achievement of the most significant features, identification of the feature importance, feature dependencies, and feature weights, that enable optimized classification of the water quality dataset. The proposed approach employs both model-based and model-agnostic interpretations, using model-based ML. Donnelly et al.<sup>46</sup> implementations and model-agnostic XAI implementations. The quality of water is greatly challenged by innumerable influencing factors. These factors vary from condition to condition and place to place. For example, Microplastics (MP) are emerging pollutants in the marine environment with potential toxic effects on littoral and coastal ecosystems<sup>47</sup> and as well as identifying the mixing of buoyant pollutants in water sources<sup>4</sup>. The laboratory evaluations show the presence of polyethylene (PE) particles in the waves of the ocean with wave steepness  $Sop$  of 2–5%. The transportation of which could cause severe water pollution on the seashores<sup>48</sup>. These measurements require quantification and feature analysis when it is evaluated with AI. This is where the XAI plays a vital role in measuring the order and degree of the pollutants causing the quantifiable pollution in the water.

### Case studies

**Importance of XAI in Water Quality Assessment:** The following case studies delineate the advent of the potential impact of XAI, with a groundbreaking revolution in water quality assessment.

**Case Study 1: Pollution of Ganges<sup>49</sup>** This case study emphasises the Ganga River pollution issue in India, which has an extremely detrimental impact on humans and the entire ecosystem. The Ganga River is polluted

Reference	Algorithms	Input parameters	Evaluation results
37	PNN	BOD, PO4-P, COD, temperature, NO 3-N, Ca 2+, Cl-, alkalinity, P, Mg 2+, pH, and EC	Interpolation is good performance - R2 : 0.82
31	BWNN, ANN, ARIMA, BANN	Dissolved Oxygen	ARIMA < ANN < WNN < BANN < BWNN
38	LSTM	Dissolved Oxygen	High runoff ratio $\geq 0.45$ • - 74% of sites
39	CCNN	water quality and DO parameters (example: Cl, NO x, pH, TDS, and temperature)	R2 : 0.825 RMSE: 0.550
40	SVM, ANN	TDS, Na+, Mg 2+, Temperature, pH, EC, HCO3 , Cl, and Ca 2+	SVM performs better than ANN
41	SVM, ANN	flow travel time, rainfall, river flow, temperature, DO, TN, and TP	SVM perform better than ANN
42	DT, RF, DCF, and 10 other models	pH, DO, CONMn, and NH 3-N	DCF, DT, and RF performed well
43	SVR	EC, fDOM, turbidity, BGA-PC, chlorophyll-a, DO, and sediments	BGA-PC : (Accuracy: 0.77), chlorophyll-a (Accuracy: 0.78), TSS (Accuracy: 0.81), from (-), turbidity (Accuracy: 0.55)
44	Attention-based neural network	Images of water	Accuracy: Polluted-water= 73.6% Accuracy: Clean-water=71.2%
45	SVM, RF, CNN	Landsat8 images	RF Accuracy: 86.21% SVM Accuracy: 96.89% CNN Accuracy: 97.12%

**Table 3.** Comparative Analysis from the review perspective.

by industrial, animal, and human waste. The main source of pollutants is industrial rubber waste, followed by leather and plastic manufacturers who dump their untreated wastewater into the river. The Ganga Action Plan was developed by the Indian government to combat Ganga pollution. This implies the need for the reinforcement of environmental restrictions to improve river quality.

### Materials and methods

An effective policy for health protection should thus emphasize providing access to safe drinking water regardless of social and economic diversity. In some places, it is evident from previous studies that investments in access to clean water and sanitation yield economic benefits for any country. It is a significant aspect of eco-friendly health and public safety, as it regulates the appropriateness of water for numerous purposes, such as drinking, agriculture, industry, and recreational purposes. The important key indicators related to water quality are its physical, chemical, and biological characteristics and its sources of pollution. The dependent target class is potability. The other independent features are pH value, hardness, solids (Total Dissolved Solids-TDS), Chloramines, sulfate, conductivity, organic carbon, trihalomethanes, and turbidity. Water's potability indicates its purity and safety for ingestion. The parameters used and their WHO limits, the hyper-parametric analysis are listed in Table 4, and the feature description of parameters are listed in Table 5.

XAI framework facilitates transparent and interpretable explanations of the outcome generated by the ML algorithm-based frameworks. XAI can thus be applied in the present context of water quality assessment to ensure accurate decision-making, thereby, enabling trustworthiness, enhancement of transparency and interpretability of the behaviour of the model.

#### *Hydro-climatic application*

XAI framework can be used to solve Hydro-Climatic problems<sup>50</sup> with diverse spatio-temporal scales. XAI is utilized to unveil the nonlinear correlative causes, in which the performance of the model is enhanced. It enables the users to discover new knowledge and further easily understand the rationale behind the decision outcomes.

#### *Groundwater potential predictions*

XAI approach can explain the decisions made by ML models for groundwater potential prediction. The user can easily interpret the outcomes and further comprehend the underlying for an outcome in the realm of water quality evaluation for conservation, and sustainability of water management.

#### *Water quality predictions*

XAI framework can forecast water quality using metrics and factors with interpretable results. Water quality assessment managers can comprehend the variables and parameters used for outcomes. This forces quality managers to mitigate water quality issues.

#### *Flood hazard risk predictions*

Floods can trigger landslides from excessive rainfall. Flooding causes countless casualties and property damage. Disaster warning systems need a flood risk assessment. XAI can forecast rapid water depths and provide timely, interpretable alerts to protect public health and safety.

#### *Environmental impact assessment*

XAI approach can be used for assessing the environmental impact on the water pollution incidents, and provide insight for mitigation and management. It enhances transparency and accountability by providing insights into



Model	Parameters	Values
Logistic regression	• Penalty	• None
	• Dual	• False
	• Tolerance	• Default
	• Regularization Strength	• Default =1.0
	• Fit_Intercept	• True (Boolean)
	• Class_Weight	• True(Boolean)
	• Random_state	• 0 (Default)
	• Solver	• lbfgs(Default)
SVM	• C	• 1,0
	• Kernal	• Linear
	• Degree	• 3
	• Gamma	• Scale
	• Co_ef	• 0
	• Shrinking	• True(Boolean)
	• Tolerance	• False(Boolean)
	• Cache	• Default
	• Class_weight	• 200MB
	• Verbrose	• None
	• Maximum_Iteration	• False
	• decision_function_shape	• 1
	• break_ties	• ovr(one vs rest)
	• random_state	• False
• Random_state	• None	
Decision Tree	• Criterion	• Gini
	• Splitter	• Best
	• Max_Depth	• None
	• Minimum_samples_split	• 2
	• Minimum_samples_leaf	• 1
	• Minimum_weight_fraction_leaf	• 0
	• Max_features	• None
	• Random_state	• None
	• Minimum_impurity_decrease	• 0
	• Maximum_leaf_nodes	• None
	• Random_state	• None
Random Forest	• N-estimators	• 100
	• Criterion	• Gini
	• Max_Depth	• None
	• Minimum_samples_split	• 2
	• Minimum_samples_leaf	• 1
	• Minimum_weight_fraction_leaf	• 0
	• Max_features	• None
	• Random_state	• 0
	• Minimum_impurity_decrease	• 0
	• Maximum_leaf_nodes	• None
	• BootStrap	• True
	• oob_score	• False
	• n_jobs	• 0
	• Verbrose	• None
• Class_weight	• 0	

**Table 4.** Hyper parameter analysis of various ML models.

the factors and parameters influencing environmental conditions. The analysis provided by the XAI model helps the stakeholders to identify the most significant factors contributing towards the environmental outcome.

Parameters	WHO limits
PH	6.5–8.5
Hardness	200 mgL%
Solids	1000 ppm%
Chloramines	4 ppm%
Sulfate	1000 mgL%
Conductivity	400 $\mu$ S/cm%
Organic carbon	10ppm%
Trihalomethanes	80 ppm%
Turbidity	5 NTU%

**Table 5.** Feature description.

## System model and architecture

### System model

Worldwide, numerous water bodies are contaminated by a variety of anthropogenic and natural processes, resulting in a variety of health problems for human life. Thus water quality requires rigorous monitoring and management to prevent pollution. In accordance with WHO guidelines, the polluted water must be treated using the proper water treatment techniques before consumption. The quality of water is contaminated by the incessant addition of toxic chemicals and microbes and also by the relentless addition of local and industrial sewage sludge, trash, and extra hazardous waste that are toxic to humans and society. Many uncertainties are required to be quantified for all machine learning models. The uncertainties such as selecting and gathering the training data, absolute and accurate training data, understanding the machine learning models with performance bounds and drawbacks and finally the uncertainties which are based on the operational data. To minimize the challenges, adhoc steps like studying the model variability and sensitivity analysis are applied. In current years, the validation of water quality has taken active momentum because of ever-increasing water pollutants which spoil water that is dedicated for domestic use and irrigation. Water quality indices (WQIs) are used worldwide very efficiently for the assessment of the quality of both groundwater and other relevant water sources. Machine Learning techniques play a substantial role in identifying the quality of water using explainable AI. Figure 7 depicts the overall architecture of the proposed framework of our study. The dataset used in the study is split into the ratio of 70:30 wherein 70% is used for training and 30% is used for testing. The model is trained using a decision tree, random forest, SVM, logistic regression, and Naive Bayes algorithms. XAI model is implemented in the framework wherein LIME and Shapely are used to provide explainability and interpretability to the results generated by the machine learning model.

### Decision tree

The decision tree is stated as a recursive partition of the set of all possible instances<sup>2751</sup>. The goal of a decision tree is to split the data which consequences in maximum information gain<sup>52</sup>. Let  $L$  be a sample for learning,  $L = (v_1, c_1), (v_2, c_2), \dots, (v_n, c_n)$ . Here,  $v_1, v_2, v_3, \dots, v_n$  are represented for measurement vectors, and  $c_1, c_2, c_3, \dots, c_n$  are represented for class labels. The batch conditions are reliant on one of the vector variables denoted as  $s_i$ <sup>53</sup>. Let us assume if the  $e_i$  of an element fits class label  $c_i$ , then  $p_i$  is denoted as per the Eq. (1).

$$p_i = \frac{c_i}{L} \quad (1)$$

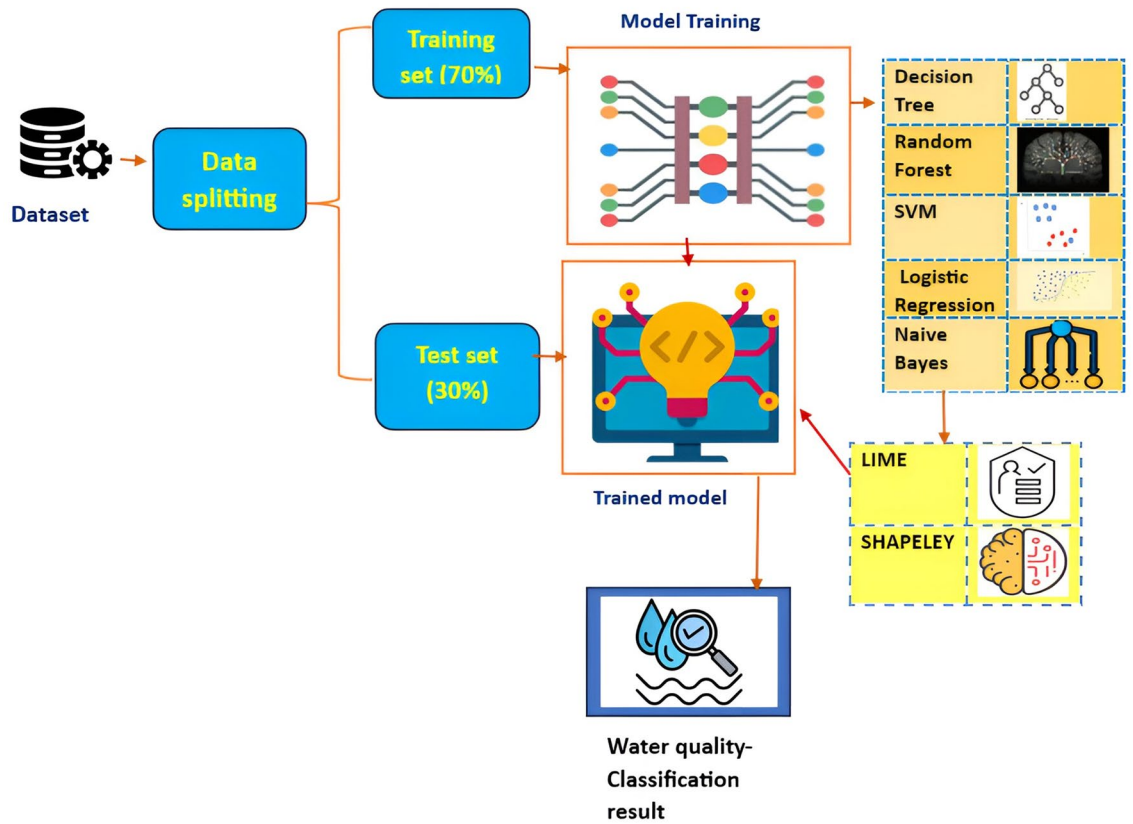
Entropy evaluates the random value from the given samples and the homogeneity of the expected rate of a group of data<sup>54</sup>. To divide the data most optimally, the lowest value of entropy signifies better homogeneity.

$$E(L) = \left(\frac{c_1}{L}\right) \log_2 \left(\frac{c_1}{L}\right) + \left(\frac{c_2}{L}\right) \log_2 \left(\frac{c_2}{L}\right) + \dots + \left(\frac{c_j}{L}\right) \log_2 \left(\frac{c_j}{L}\right) \quad (2)$$

$$E(L) = e_1 \log_2 e_1 + e_2 \log_2 e_2 + \dots + e_n \log_2 e_n \quad (3)$$

$$E(L) = - \sum_{i=1}^j e_i \log_2 e_i \quad (4)$$

$L$  represents the data set evaluated by the entropy, 'i' denotes the classes in the set  $L$ , and  $e_i$  indicates the number of data labels that fit class 'i'<sup>55</sup>. The least value of entropy is used for choosing the best feature. Information gain enumerates the amount of information provided by a particular characteristic about the target variable to minimize the uncertainty present in the data set. It is calculated by comparing the weighted average of entropy to the original data set after the splitting process. Let us assume that  $R$  is the rate for the features  $f_i$  [ $L^R$ ] denotes the subset of  $L$  so that  $\text{bf} = R$ <sup>56</sup>. After splitting  $L$  on the feature, information gain is given as follows.



**Figure 7.** Interfacing ML algorithms with XAI.

$$IG(L, bf) = Entropy(L) - \sum_{R=1}^R \frac{|L^R|}{|LS|} Entropy(L^R) \tag{5}$$

The Gini index evaluates the heterogeneity of a selected node in the decision tree. It counts the probability of wrongly identifying data in the node. The Gini index begins from the value 0 to 1, where 0 indicates a pure node and 1 denotes a node that is distributed equally. The Gini index is represented as

$$Gini(L) = 1 - \sum_{i=1}^j e_i^2 \tag{6}$$

Here,  $e_i$  represents the quantity of data labels. When the data is divided on class  $d$  as  $L_1$  and  $L_2$  with sizes  $s_1$  and  $s_2$ , Gini is evaluated as

$$Gini_d(L) = \frac{s_1}{s} Gini(L_1) + \frac{s_2}{s} Gini(L_2) \tag{7}$$

Due to its comprehensible nature, decision trees can manage both numerical and categorical data with automatic feature selection.

*Random forest*

Random forest is an ensemble method that groups the results of multiple decision trees to compute predictions with enhanced accuracy. Every decision tree is improved on a random subset of labels from the dataset, to achieve diversity between the trees. When the data in the training label is  $t$ , then with replacement ‘ $n$ ’ data are verified as bootstrap data<sup>57</sup>. This is done to produce the decision tree with training data. When there are ‘ $m$ ’ labels,  $a < m$  is selected so that ‘ $a$ ’ values are considered at random from ‘ $m$ ’. The value ‘ $a$ ’ is constant when the tree is growing to the highest level. The highest vote is noted as a new instance. ( $GE^*$ ) is the generalization error for the random forest and is denoted as

$$GE^* = P_{x,y}(f(X, Y)) < 0 \tag{8}$$

Here,  $f(X, Y)$  is a margin function to count the average number of votes from  $(X, Y)$ .  $X$  denotes the prediction value and  $Y$  denotes the classification problem. The margin function is represented as

$$f(X, Y) = av_k F(h_k(X) = Y) - \max_{j \neq y} av_k F(h_k(X) = j) \tag{9}$$

where 'F' is for the indicator function. The value for the margin function is indicated as

$$R = E_{X,Y}(f(X, Y)) \tag{10}$$

The average value of a random forest and the mean correlation of the classifiers are combined as generalization errors. The  $\rho$  denotes the mean of the correlation. The generalization error for the upper bound is

$$GE^* \leq \rho(1 - s^2)/s^2 \tag{11}$$

Random forest reduces the over-fitting problem compared to a single decision tree. It can effectively manage high-dimensional data.

*Support vector machine (SVM)*

Let us consider a binary classification problem 1 or -1 to represent the sample variables<sup>58</sup>. When i elements of the sample variable is -1, it is a positive class. When the i variables of the samples is 1, it is a negative set. Let  $V_i = X_1, X_2, \dots, X_n, Y_i, i = 1, 2, \dots, n, Y_i \in -1, 1, S_i$  indicates i item from the samples.  $Y_i$  is the i item of the tests performed<sup>59</sup>. To split the samples into two parts, the function  $f(X) = Z^T X + b$  is used, where Z is the coefficient vector to normalize the hyperplane. The optimal margin is given as

$$\underbrace{\text{MIN}}_{w, b, \xi} \left( \frac{1}{2} Z^T Z + C \sum_{i=1}^n \varepsilon_i \right)$$

subject to:

$$Y_i (Z^T X_i + b) \geq 1 - \varepsilon_i, \varepsilon_i \geq 0 \tag{12}$$

The Lagrangian equation is given as

$$\underbrace{\text{MAX}}_{\alpha} \left( \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j Y_i Y_j X_i X_j \right)$$

subject to:

$$0 \leq \alpha_i \leq C, i = 1, 2, \dots, n, \sum_{i=1}^n \alpha_i Y_i = 0 \tag{13}$$

The Lagrangian equation with the maximum value with  $\alpha_i$  a positive multiplier for the equation  $\sum_{i=1}^n \alpha_i Y_i = 0$  and  $\alpha_i \geq 0$  to change the optimal hyperplane<sup>60</sup> is presented. The optimal equation is given as

$$f(X, \alpha^*, b^*) = \sum_{i=1}^n Y_i \alpha_i^* < X_i, X_j > + b^* \tag{14}$$

$$= \sum_{i \in sv} Y_i \alpha_i^* < X_i, X_j > + b^* \tag{15}$$

In the above equation  $\alpha_i = 0$  of the Lagrangian multiplier is nearest to the margin of the optimal hyperplane denoted as a support vector. This data is linearly separable by the kernel to evaluate the expected result from the instance<sup>61</sup>. The kernel function is denoted as

$$K(X_i, X_j) = \varphi(X_i)^T \cdot \varphi(X_j) \tag{16}$$

The generalized linear equation is changed to represent the non-linear dual Lagrangian  $La(\alpha)$ .

$$Lag(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j Y_i Y_j K(X_i, X_j)$$

Subject to:

$$0 \leq \alpha_i \leq C, i = 1, 2, \dots, n, \sum_{i=1}^n \alpha_i Y_i = 0 \tag{17}$$

The Lagrangian equation can be used for the separable case as

$$f(X, \alpha^*, b^*) = \sum_{i=1}^n Y_i \alpha_i^* K(X_i, X_j) + b^* \tag{18}$$

The SVM algorithm is very effective when the quantity of features is higher than the number of samples<sup>62</sup>.

*Logistic regression*

Logistic regression is used for binary classification problems to forecast the probability of an occurrence matching to a particular class. If the dependent value is binary, a regression analysis is used. The idea in logistic regression(logreg) is the logarithm 'logn' of odds of X, and odds are the ratios of probabilities 'pb' of X<sup>63</sup>. The

rate of the independent value is termed odds because logistic regression measures the probability of an act that happens over the likelihood of an occurrence that does not happen.

$$\text{logreg}(Y) = \text{logn}(\text{odds}) = \text{logn}\left(\frac{pb}{1-pb}\right) = a + \beta x \quad (19)$$

where  $p$  is the probability of a positive output and  $x$  is the variable. The  $\alpha$  and  $\beta$ , are the logistic regression parameters<sup>64</sup>. The above equation is used for finding the number of occurrences as

$$p = \text{probability}(Y = \text{positive outcome} | X = x, \text{a specific value}) \\ = \frac{e^{\alpha + \beta x}}{1 + e^{\alpha + \beta x}} \quad (20)$$

For multiple predictors, a logic regression equation can be written as

$$\text{logreg}(X) = \text{logn}(\text{odds}) = \text{logn}\left(\frac{pb}{1-pb}\right) = a + \beta_1 x_1 + \dots + \beta_k x_k \quad (21)$$

$$p = \text{probability}(Y = \text{positive outcome} | X_1 = x_1, \dots, x_k) \\ = \frac{e^{\alpha + \beta_{x_1} + \dots + \beta_{x_k}}}{1 + e^{\alpha + \beta_{x_1} + \dots + \beta_{x_k}}} = \frac{1}{1 + e^{\alpha + \beta_{x_1} + \dots + \beta_{x_k}}} \quad (22)$$

Here,  $pb$  refers to the probability of the positive occurrence of the event, the Y-intercept is  $\alpha$ , the regression coefficient is  $\beta$ , and  $e$  is 2.71828. Logistic regression is applied in various domains like finance, healthcare, social sciences, and many more for predicting diseases, credit default, etc.

### Naive Bayesian classification

Gaussian Naive Bayes is a probabilistic classification algorithm developed based on Bayes theorem. It refers to the features which represent a normal distribution<sup>65</sup>. It classifies the samples as most likely classified as

$$P((C_i | Y) = \text{Max}\{P(C_1 | Y), P(C_2 | Y), \dots, P(C_n | Y)\} \quad (23)$$

If the sample  $Y_j$  is a vector,  $x_j$  is the  $j^{\text{th}}$  value which contains different values of  $y_j$ . The attributes used are dependent and it is shown as

$$P(Y | C_i) = \prod_{j=1}^k P(A_j = (y_j | C_i)) \quad (24)$$

Substituting the above equation into Bayes classification, we get

$$P((C_i | Y) = \frac{\prod_{j=1}^k P(A_j = (y_j | C_i) P(C_i))}{P(Y)} \quad (25)$$

$$\text{If } \frac{1}{P(Y)} = \alpha (> 0), \text{ then} \quad (26)$$

$$P(C_i | Y) = \alpha \prod_{j=1}^k P(A_j = (y_j | C_i) P(C_i). \quad (27)$$

The Gaussian Naive Bayes algorithm is mainly applied for spam filtering, sentiment analysis, and text classification problems where the features must be continuous and follow the Gaussian distribution<sup>66</sup>.

### LIME (Local interpretable model-agnostic explanations)

LIME explains the predictions of any kind of classifier by approximating locally along with an interpretable system. It changes the data sample by altering the values of features and monitors the impact of the result. It explains the predictions from every sample<sup>67</sup>. To receive the labels for the current data, alter the samples  $z$ 's into the unique form  $z \in \mathbb{R}^d$ . Since the samples  $x'$  are generated randomly,  $x$  samples closer to the unique instance  $z$  for weighing are considered. The weight is evaluated as  $\Pi_z(x)$  for measuring the intimacy between the data  $z$  to  $x$ . The currently weighted data  $X$  and the samples formed by  $f(x)$ , are trained as  $g \in G$ , where  $G$  is a model. The interpretable model  $\xi(x)$  of the current data  $g$  for explaining  $f(x)$  as

$$\xi(z) = g \in \text{Garg min}(L(f, g, \Pi_z) + \Omega(g)) \quad (28)$$

$L$  is the loss function to measure whether  $g$  is following the state of  $f$  in the nearest neighborhood of  $z$ . If the loss function is reduced, the behaviour of  $g$  takes the behaviour of  $f$  as  $\Pi_z$ . The complexity of the model  $\Omega(g)$  should be low. When  $g(x')$  is considered as a linear function,  $g(x') = \varphi^T x' + \varphi_0$ , changes the equation into a linear regression task to evaluate  $\varphi$  and  $\varphi_0$ .



$$L(f, \varphi_0, \Pi_x) = \sum_{z, z' \in Z} \Pi_x(z) \left( f(z) - (\varphi_0 + \varphi^T z') \right)^2 \quad (29)$$

SHAP (SHAPELY Additive exPlanations)

SHAP values determine the status of each feature for the prediction of a specific class<sup>68</sup>. The prediction  $f(y)$ , using  $s(y')$ , a model for the binary elements  $x' \in \{0, 1\}^M$  with the sets  $\emptyset_i \in \mathbb{R}$ , is given as

$$s(x') = \emptyset_0 + \sum_{i=1}^M \emptyset_i x'_i \quad (30)$$

$M$  refers to the explanation variable.

$$\Phi_i(f, z) = \sum_{x' \subseteq z'} \frac{(|x'|!(M - |x'| - 1)!)}{M!} [f_y(x') - f_y(x' \setminus i)] \quad (31)$$

where  $f$  is the model of the SHAP,  $z$  refers to the variable, and  $z'$  are the variables chosen. The value  $f_y(x') - f_y(x' \setminus i)$  indicates all the predictions.

### Algorithm

In this section two algorithms are discussed: one for the algorithm-based evaluation of water quality **1** and another for the algorithm-based explanation of water quality **2**. These two algorithms provide a holistic analysis and explanation of water quality management.

---

```

1 Input:  $x = [\sum_{i=0}^n I_n;$ 
2  $x_i r$  form = label.encoder( $x$ );
3  $y \leftarrow y_{train}, y_{test};$ 
4  $x \leftarrow x_{train}, x_{test};$ 
5  $n \leftarrow samples, o, f, images;$ 
6  $TRUE_P \leftarrow TruePositive;$ 
7  $TRUE_N \leftarrow TrueNegative;$ 
8  $FALSE_P \leftarrow FalsePositive;$ 
9  $FALSE_N \leftarrow FalseNegative;$ 
10  $Features \leftarrow a, b;$ 
11 Accuracy:  $\frac{TRUE_P + TRUE_N}{TRUE_P + TRUE_N + FALSE_P + FALSE_N};$ 
12 Precision:  $\frac{TRUE_P}{TRUE_P + FALSE_P};$ 
13 Recall:  $\frac{TRUE_P}{TRUE_P + FALSE_N};$ 
14 F1-Score:  $\frac{2 * TRUE_P}{2 * TRUE_P + FALSE_P + FALSE_N};$ 
15 Activation:  $\max[Accu, Prec, ReCa, F1 - Sco, Sensi, Speci];$ 
16 while  $y_{is} \neq 0$  do
17   if  $x_{test}$  is Potable then
18      $accu \leftarrow \frac{TRUE_P + TRUE_N}{TRUE_P + TRUE_N + FALSE_P + FALSE_N};$ 
19      $Preci \leftarrow \frac{TRUE_P}{TRUE_P + FALSE_P};$ 
20      $reca \leftarrow \frac{TRUE_P}{TRUE_P + FALSE_N};$ 
21      $f1 - sco \leftarrow \frac{2 * TRUE_P}{2 * TRUE_P + FALSE_P + FALSE_N};$ 
22      $Sensi \leftarrow \frac{TRUE_P}{TRUE_P + FALSE_N};$ 
23      $Speci \leftarrow \frac{FALSE_P}{FALSE_P + TRUE_N};$ 
24 else
25    $x_{test}$  is Not Potable
26  $accu \leftarrow \frac{TRUE_P + TRUE_N}{TRUE_P + TRUE_N + FALSE_P + FALSE_N};$ 
27  $Preci \leftarrow \frac{TRUE_P}{TRUE_P + FALSE_P};$ 
28  $reca \leftarrow \frac{TRUE_P}{TRUE_P + FALSE_N};$ 
29  $f1 - sco \leftarrow \frac{2 * TRUE_P}{2 * TRUE_P + FALSE_P + FALSE_N};$ 
30  $Sensi \leftarrow \frac{TRUE_P}{TRUE_P + FALSE_N};$ 
31  $Speci \leftarrow \frac{FALSE_P}{FALSE_P + TRUE_N};$ 

```

---

**Algorithm 1.** Algorithm for water quality classification

---

```

1 Input:  $x = [\sum_{i=1}^{c_n} x_n(1 - x_n)]$ ;
2  $y \leftarrow y_{train}, y_{test}$ ;
3  $x \leftarrow x_{train}, x_{test}$ ;
4  $k \leftarrow No\_of\_samples$ ;
5  $cf \leftarrow Complexityfunction$ ;
6  $r \leftarrow local\_surrogate\_regressor$ ;
7  $L \leftarrow Loss\_function$ ;
8  $d \leftarrow Permutations$ ;
9  $d \cup D$ ;
10  $E \leftarrow Number\_of\_players$ ;
11  $v \leftarrow value\_function\_of\_the\_players$ ;
12 while  $Y \neq Local$  do
13   if Predictproba is local then
14      $L(x) \leftarrow Loss\_function$ ;
15      $exp \leftarrow \theta(k) = L(r, cf, \pi_k) + \omega(cf)$ ;
16     ▷ Decision Explained with Local Surrogates (LIME)
17   else
18     Predictproba is global
19      $\theta(y) \leftarrow Costfunction$ ;
20      $exp \leftarrow \theta(v) = \sum_{d \subset E} |d|!(E - |d| - 1)!/E! * (v(E \cup i - V(p)))$ ;
21     ▷ Decision Explained with Global Surrogates (SHAPLEY)

```

---

**Algorithm 2.** Algorithm for water quality Explanation

## Results

The water quality is assessed in the proposed work based on nine parameters such as pH value, Hardness (Total Dissolved Solids), Sulphate, Chloramines, Trihalomethanes, Conductivity, Organic carbon, and Turbidity. The target class for this dataset is Potability which is binary where 0 indicates that the water is not potable and 1 reflects its potability.

The dataset consisted of high missing values on sulphate and lower missing values on Chloramines and Trihalomethanes. The missing value imputation is hence performed and all the attributes are imputed for the missing values. The target class is converted into a numeric array for the processing of XAI models. This is done with the label encoder application of Python. The dataset is split with a ratio of 80:20 for training and testing.

The correlation analysis is performed on the dataset. The attribute Hardness has a high correlation of 0.34 with the target attribute potability. The next best correlation value is 0.24, which is rendered by the attribute Chloramines, followed by 0.21 produced by the Trihalomethanes attribute. Turbidity is the next better parameter with a correlation value of 0.16. The correlation heat map between the attributes of interest and the target attribute is presented in Fig. 8.

The trained dataset is applied with SVM, LR, DT, RF and Gaussian Naive Bayes machine learning models. The SVM did not provide the desired classification and failed to converge for the portable data. The other models generated the results within the desired range and are presented in Table 6.

The sensitivity and specificity measurements for the Machine learning models are presented in Table 7. Considering the performance metrics, the results reveal the superiority of the RF model which generates a better outcome in comparison to the other models and thus it has been selected to be fed into the XAI model to provide enhanced interpretability, justifiability and transparency.

The XAI model implementation is performed considering SHAPLEY values in the pandas' application. This application focuses on the value of each feature in determining the target attribute which is potability. The significance of every feature is assessed through the various applications of SHAPLEY. The first XAI model generated is the force plot, which provides the minimum and maximum prediction score of the target attribute in a dataset. The blue colored contour shows that a low score is measured and the red color shows a high score. The values at the separation boundary have the highest priority attribute. The force plot is presented in the Figs. 9 and 10.

The Global surrogate version of the force plot is presented in Fig. 11. The blue regions indicate no potability and the red-coloured regions indicate potability. The border areas of the intersection show the attributes which have higher significance for the feature selection. The Sulphate value of 444 at the point of intersection indicates its significance in explaining this test patch for the entire dataset.

The next XAI application of SHAPLEY is the summary plot. This plot describes the features in determining binary classification problems. This predicts the scale of low to high for two significant results. The blue contour indicates lower significance towards the prediction and red indicates higher significance. The summary plot is shown in Fig. 12. The Solids, pH, Sulfate, and Hardness show higher significance in determining the output.

The dependency plot shows the relationship between two features in the dataset. It provides the output in granular form with a variable-like result rather than simply a graph-like result of a Partial Dependency Plot (PDP). The relationship between the Sulphate and Potability is depicted in Fig. 13. The mid-range of the dataset provides

more granular output, which shows that the Sulphate parameter values are more significant in determining the values of potability in the mid-range of the dataset.

The decision plot, which displays how the values of the features affect the goal, is the final model of XAI. This plot is a local surrogate plot, which would only explain a certain data instance, in which what values of the attributes influence the decision to be 1 or 0 as the decision of the model. The decision plot for the potability as 1 is illustrated in Fig. 14. The potability 0 is illustrated in Fig. 15.

## Discussion

The results of the experiment reveal the superiority of the RF model which generates an accuracy of 0.999 followed by DT, generating an accuracy of 0.998. The lowest accuracy is generated by the SVM model of 0.63. The RF is thus chosen for the implementation of the XAI model using SHAPLEY. The comparative analysis of the aforementioned various models is depicted in Fig. 16, considering evaluation metrics accuracy, precision, recall, and f1-score. In the case of all the performance metrics, the RF model outperforms the other models. Figure 17 shows the comparison of the sensitivity and specificity measures. The RF model stands superior in these considerations as well. Thus, the discussion offers a visual representation and justification of the reasoning behind the choice of RF to be included in the XAI framework to offer explainability.

Apart from the selection of the RF model, SHAPLEY provided five different models to explain the feature importance and relationships. The proposed work presented the force plot, summary plot, test patch, dependency plot, and decision plot. The Final decision plot explained how the classification is carried out using the corresponding values of the independent variables. Thus the black-box classification is explained in the white-box context of XAI. The following section describes the challenges and opportunities of the proposed work with an emphasis on future directions.

## Challenges

The proposed work may be influenced by the following challenges which are described in detail as follows,

### *Global unity*

For the successful implementation of the system, a unanimously accepted implementation is essential. Unfortunately, water quality estimation and related research are limited to consideration of specific datasets acquired for a particular region, wherein the generated results may differ with the changes in geographic location. Thus the generated results can never be considered suitable on a global scale. The parameters that influence the water quality may also vary across the world, and hence the proposed work can never be considered as a universal solution.

### Training and re-training

The qualifying attributes that determine the quality of water vary across the globe and hence the proposed model needs to be re-trained<sup>69</sup> when applied to a new environment of study. This would allow the model to unlearn and re-learn new environments. On the contrary, the complexity of the model would also increase. The accuracy and other performance metrics which are measured in the proposed work may drastically decrease as well in a different environment of study. Thus applying this model to versatile environments is complex and would be a challenging task.

### Subjective or quantitative

The trade-off from subjective analysis (which was done through fuzzy-based methods in the form of the Analytical Hierarchy Process (AHP) and The Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS)) has improved the performance and ability to classify the models with better accuracy. However, the involvement of a subject matter expert is a missing point in the current research. Despite all the implementation and analysis from an engineering perspective, the involvement of an environmental scientist in any aspect of water research would contribute towards the enhancement of research quality.

### Confusing solids

The proposed work identifies Solids as the primary influencing factor that affects potability. In real-world applications, solids can be of any form. For example, in sewage water treatment plants it can be either mud, Fat-Oil-Grease(FOG), or any other substances. Every solid wastage has its way of filtration and impact on water quality, which makes the recordings unstable from time to time. The attributes of research are too complex to handle in real-life scenarios, which acts as an inevitable yet detrimental impact.

### Environmental challenges

Water resources are under serious threat due to water scarcity, water contamination, water conflicts and climate changes. Chemical and the municipal wastewater contaminates the water and endangering the life of the aquatic organisms and affect their ability to reproduce. This also makes them an easier prey to their predators. The food cycle and livelihood of the human is also greatly affected by the water contamination. Chemical substances make the water hard to recycle and consume by reducing the regeneration ratios.

### Water quality and industrial sustainability

The era of Industry 5.0 focuses on the consumer centric industrial evolution with the idea of environmental sustainability. The futuristic technologies evolve with the improvement of technical viability, with the mission



Figure 8. Correlation analysis for water quality attributes.

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.77	0.74	0.75	0.76
SVM	0.63	0.39	0.63	0.48
Gaussian Naive Bayes	0.996	0.994	0.992	0.997
Decision Tree	0.998	0.996	0.997	0.998
Random Forest	0.999	0.997	0.998	0.999

Table 6. Comparison of Metrics of Machine Learning Models on Water Quality.

Model	Sensitivity	Specificity
Logistic Regression	0.76	0.75
SVM	0.55	0.54
Gaussian Naive Bayes	0.995	0.994
Decision Tree	0.996	0.997
Random Forest	0.998	0.999

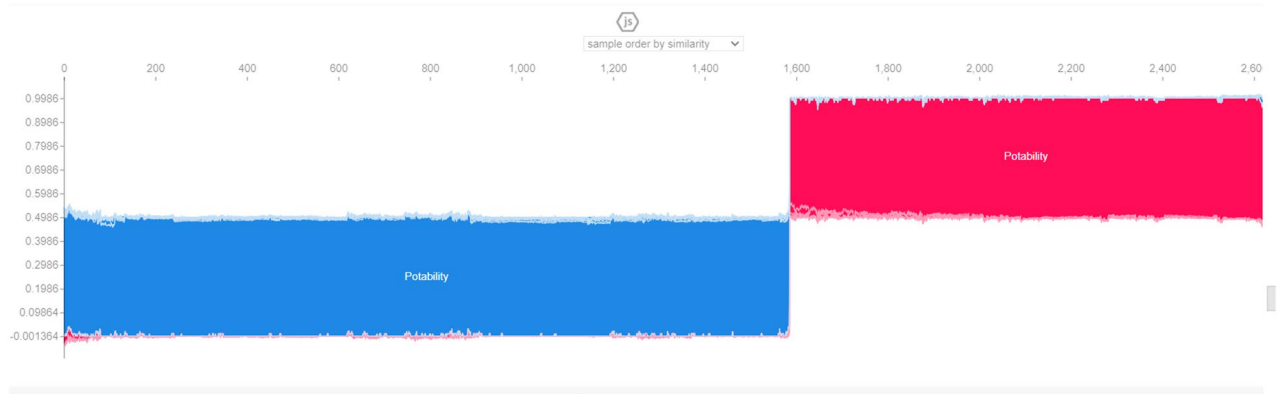
Table 7. Comparison of sensitivity and specificity for the machine learning models.



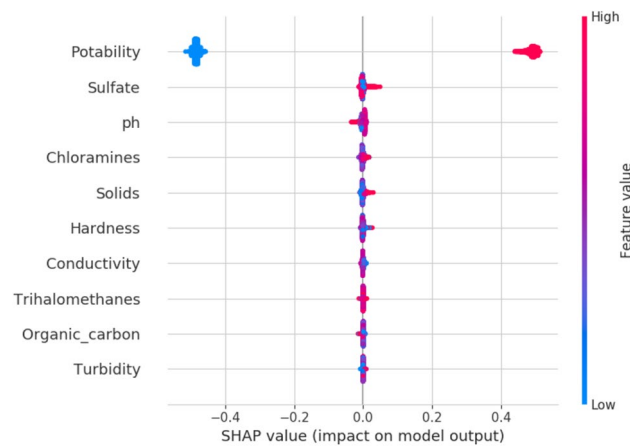
Figure 9. Force plot for water quality.



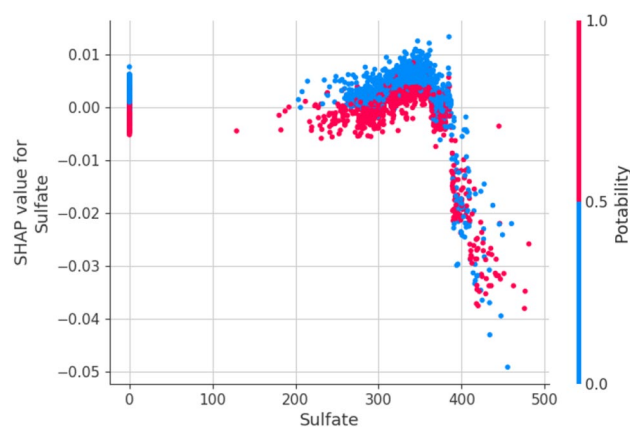
**Figure 10.** Force plot for potability.



**Figure 11.** Test patch for potability.



**Figure 12.** Summary plot for potability.



**Figure 13.** Dependency plot for potability.

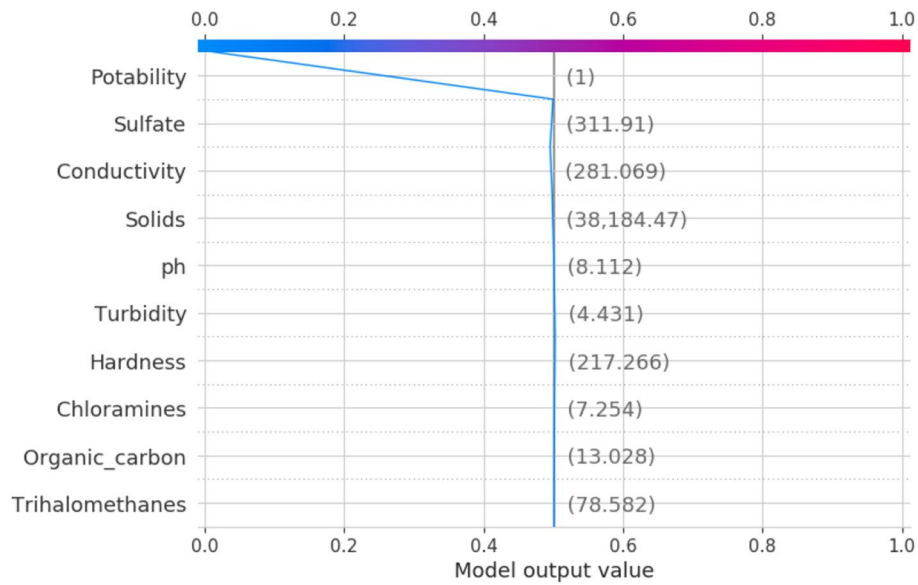


of sustainable development in the environmental aspects. Since the water is an irreplaceable and finite, the demand of the water is increasing with the industrial evolution and the water requirements on manufacturing and production industries would be very much essential as ever. The challenge is enhancement of the water harvesting, recycling and conservation. For all the above said processes quality of the water is the common essential requirement. Thus the quality of the water is more critical in all futuristic technological developments.

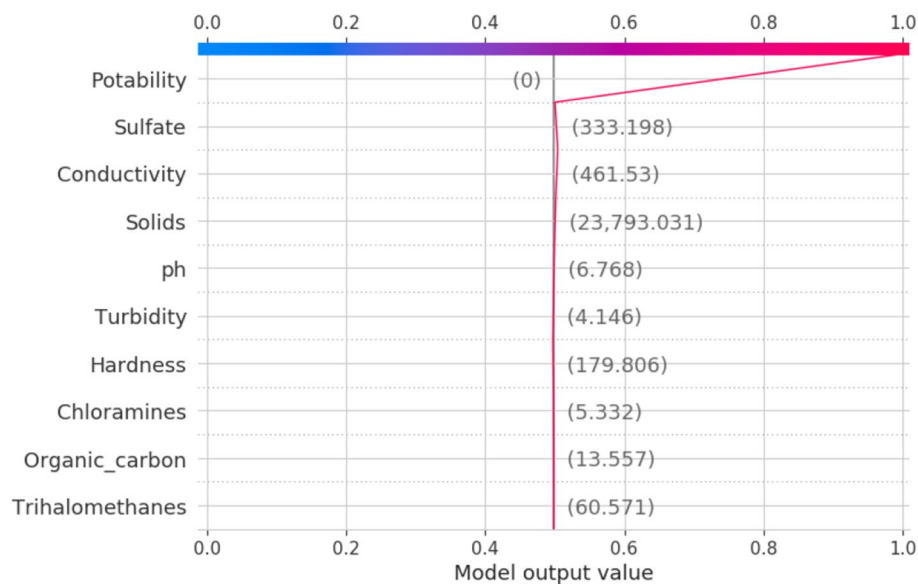
**Research finding of the proposed work**

The following items are presented as the findings are outcomes of the proposed work

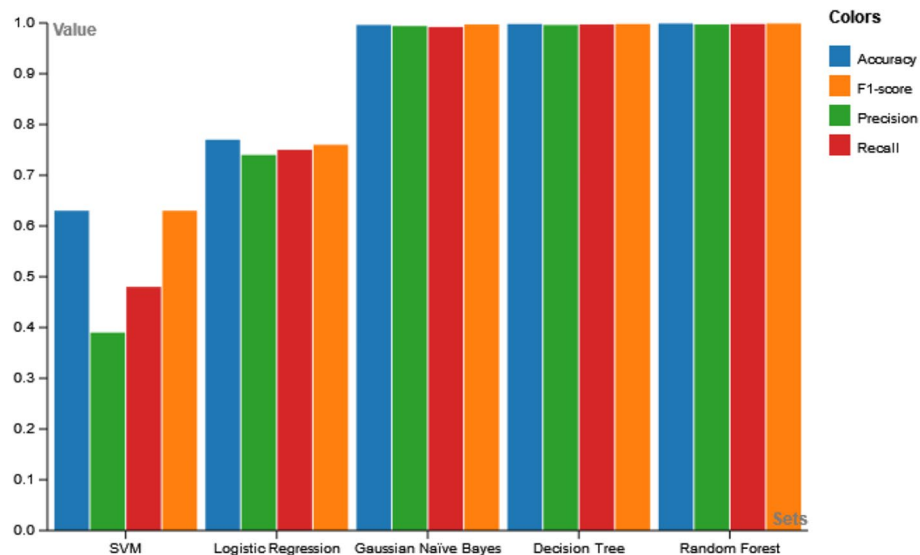
- The proposed work performs an exploratory analysis with XAI implementation providing an ability to improve the reliability of machine learning models providing explanation and transparency to the classification process.
- The proposed work acquires data from a single dataset, where the performance of classification yields optimized results. This result may vary if the model is subjected to a different dataset constituting different features and instances.



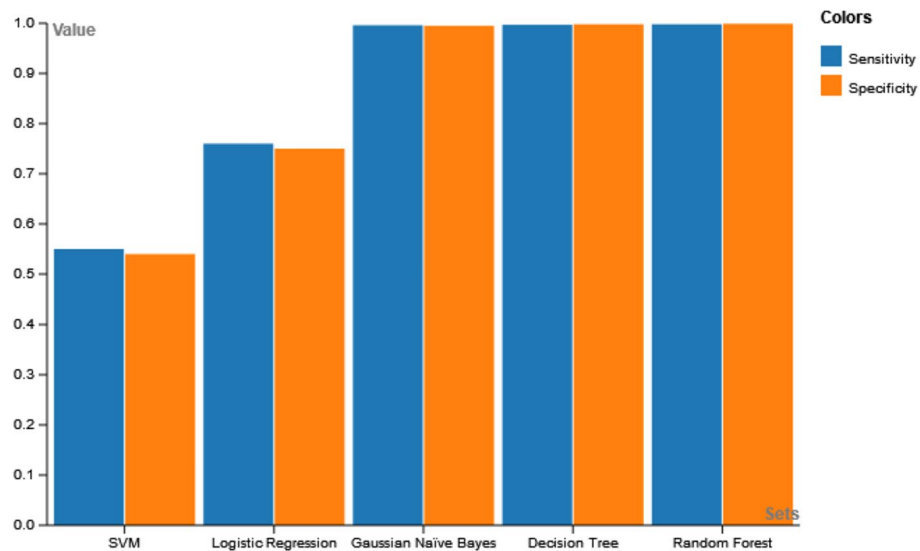
**Figure 14.** Decision plot for potability.



**Figure 15.** Decision plot for potability.



**Figure 16.** Comparative analysis of machine learning models used.



**Figure 17.** Comparative analysis of sensitivity and specificity.

- The XAI reveals the most significant features contributing towards classification results and also explains the same.
- The best fitting machine learning model is chosen for the explanation through an exhaustive analysis and evaluation of all the models considering the essential performance metrics. Thus the results produced by SHAPLEY can be considered as the most reliable and acceptable.
- The proposed work also suggests the importance of the subject matter expert, which can extend the usability of the proposed model at the universal level.
- The predictions of the proposed work with the support of an explainer, helps end users and consumers to understand the quality of the water they use.
- The features related to the classification and explanation, can be further controlled to diminish the levels of chemicals and pollutants in water recycling.
- Total dissolvable solids quantification and the feature weights for the same determine the levels of filtration and carbon purification required in the recycling plants.
- The proposed work brings insights of pollutants on the seashore and how the explainability can support the impurity estimations for such conditions also.

## Conclusion

Water quality management impacts almost all aspects of life on earth and clean water is a basic necessity. The proposed work is extremely relevant in this regard wherein an exploratory analysis conducted to analyze and control the factors that deteriorate the quality of the water. The impact of these factors is explained using XAI models. The contribution of the XAI model lies in its ability to explain the role of the underlying parameters towards the classification of water being potable or not, based on their relative importance and unique properties. The XAI model uses SHAPLEY considering the probabilistic prediction generated from the Random Forest classifier. This RF model in this regard is chosen as it yields the highest accuracy of 0.999 with sensitivity and specificity of 0.999 and 0.998, which is found to be superior in comparison to the other state-of-the-art models considered in the study. This justifies the reason for the RF to be selected for XAI implementation. The proposed model identifies the parameter “solid” as the most significant in terms of its impact on the potability of water. The proposed model yields optimized and explainable results considering the dataset used in the study. Future work may involve more complex and heterogeneous datasets to generate predictions. In such scenarios, the metric evaluations may differ. The usage of deep learning algorithms could further enhance the examination the solid sediments and generate classification results based on their mass, dimensions, and shape. The use of XAI in such a model would ensure a better explanation of factors relevant to the solid sedimentation in water.

## Data availability

The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Received: 21 October 2023; Accepted: 11 March 2024

Published online: 29 March 2024

## References

- Zhu, M. *et al.* A review of the application of machine learning in water quality evaluation. *Eco-Environ. Health* **1**, 107–116. <https://doi.org/10.1016/j.eehl.2022.06.001> (2022).
- Miller, M., Kisiel, A., Cembrowska-Lech, D., Durlik, I. & Miller, T. Iot in water quality monitoring are we really here?. *Sensors* **23**, 960. <https://doi.org/10.3390/s23020960> (2023).
- Akhtar, N. *et al.* Modification of the water quality index (wqi) process for simple calculation using the multi-criteria decision-making (mcdm) method: A review. *Water* **13**, 905. <https://doi.org/10.3390/w13070905> (2021).
- Abolfathi, S. & Pearson, J. Application of smoothed particle hydrodynamics (sph) in nearshore mixing: A comparison to laboratory data. *Coastal Eng. Proc.* **35**, 1–13 (2017).
- Hájek, M. *et al.* A European map of groundwater ph and calcium. *Earth Syst. Sci. Data* **13**, 1089–1105. <https://doi.org/10.5194/essd-13-1089-2021> (2021).
- Li, L. *et al.* Interpretable tree-based ensemble model for predicting beach water quality. *Water Res.* **211**, 118078. <https://doi.org/10.1016/j.watres.2022.118078> (2022).
- Lu, J. Can the central environmental protection inspection reduce transboundary pollution? Evidence from river water quality data in china. *J. Clean. Prod.* **332**, 130030 (2022).
- Halder, J. N. & Islam, M. N. Water pollution and its impact on the human health. *J. Environ. Hum.* **2**, 36–46 (2015).
- Wang, Z. *et al.* Overview assessment of risk evaluation and treatment technologies for heavy metal pollution of water and soil. *J. Clean. Prod.* **379**, 134043 (2022).
- Elehinafe, F. B., Agboola, O., Vershima, A. D. & Bamigboye, G. O. Insights on the advanced separation processes in water pollution analyses and wastewater treatment: A review. *S. Afr. J. Chem. Eng.* **48**, 188–200 (2022).
- Mu, L., Mou, M., Tang, H. & Gao, S. Exploring preference and willingness for rural water pollution control: A choice experiment approach incorporating extended theory of planned behaviour. *J. Environ. Manag.* **332**, 117408 (2023).
- Wang, Y., Ding, X., Chen, Y., Zeng, W. & Zhao, Y. Pollution source identification and abatement for water quality sections in Huangshui River Basin, China. *J. Environ. Manag.* **344**, 118326 (2023).
- Najafzadeh, M. & Niazmardi, S. A novel multiple-kernel support vector regression algorithm for estimation of water quality parameters. *Nat. Resour. Res.* **30**, 3761–3775 (2021).
- Najafzadeh, M., Homaei, F. & Farhadi, H. Reliability assessment of water quality index based on guidelines of national sanitation foundation in natural streams: Integration of remote sensing and data-driven models. *Artif. Intell. Rev.* **54**, 4619–4651 (2021).
- Najafzadeh, M., Ghaemi, A. & Emamgholizadeh, S. Prediction of water quality parameters using evolutionary computing-based formulations. *Int. J. Environ. Sci. Technol.* **16**, 6377–6396 (2019).
- Najafzadeh, M. & Basirian, S. Evaluation of river water quality index using remote sensing and artificial intelligence models. *Remote Sens.* **15**, 2359 (2023).
- Chowdhury, M. A. Z. *et al.* Organophosphorus and carbamate pesticide residues detected in water samples collected from paddy and vegetable fields of the Savar and Dhamrai Upazilas in Bangladesh. *Int. J. Environ. Res. Public Health* **9**, 3318–3329 (2012).
- Ahirvar, B. P., Das, P., Srivastava, V. & Kumar, M. Perspectives of heavy metal pollution indices for soil, sediment, and water pollution evaluation: An insight. *Total Environ. Res. Themes* **6**, 100039 (2023).
- Chen, K., Liu, Q.-M., Peng, W.-H., Liu, Y. & Wang, Z.-T. Source apportionment of river water pollution in a typical agricultural city of Anhui province, Eastern China using multivariate statistical techniques with apcs-mlr. *Water Sci. Eng.* **16**, 165–174 (2023).
- Fan, S. *et al.* Improved multi-criteria decision making method integrating machine learning for patent competitive potential evaluation: A case study in water pollution abatement technology. *J. Clean. Prod.* **403**, 136896 (2023).
- Wang, Z., Wang, C. & Liu, Y. Evaluation for the nexus of industrial water-energy-pollution: Performance indexes, scale effect, and policy implications. *Environ. Sci. Policy* **144**, 88–98 (2023).
- Asomaku, S. O. Quality assessment of groundwater sourced from nearby abandoned landfills from industrial city in Nigeria: Water pollution indices approach. *HydroResearch* **6**, 130–137 (2023).
- Balaram, V., Copia, L., Kumar, U. S., Miller, J. & Chidambaram, S. Pollution of water resources and application of icp-ms techniques for monitoring and management: A comprehensive review. *Geosyst. Geoenviron.* **2**, 100210 (2023).
- Yuan, F., Huang, Y., Chen, X. & Cheng, E. A biological sensor system using computer vision for water quality monitoring. *Ieee Access* **6**, 61535–61546 (2018).
- Borzooei, S. *et al.* Impact evaluation of wet-weather events on influent flow and loadings of a water resource recovery facility. In *New Trends in Urban Drainage Modelling: UDM 2018* 11 706–711 (Springer, 2019).
- Noori, R. *et al.* Decline in Iran’s groundwater recharge. *Nat. Commun.* **14**, 6674 (2023).

27. Yeganeh-Bakhtiary, A., Eyvazoghli, H., Shabakhty, N., Kamranzad, B. & Abolfathi, S. Machine learning as a downscaling approach for prediction of wind characteristics under future climate change scenarios. *Complexity* **2022**, 8451812 (2022).
28. Jeihouni, M., Toomanian, A. & Mansourian, A. Decision tree-based data mining and rule induction for identifying high quality groundwater zones to water supply management: a novel hybrid use of data mining and gis. *Water Resour. Manag.* **34**, 139–154 (2020).
29. Lee, K.-J. *et al.* The combined use of self-organizing map technique and fuzzy c-means clustering to evaluate urban groundwater quality in Seoul Metropolitan City, South Korea. *J. Hydrol.* **569**, 685–697 (2019).
30. Agrawal, P. *et al.* Exploring artificial intelligence techniques for groundwater quality assessment. *Water* **13**, 1172 (2021).
31. Wang, Y. *et al.* Monthly water quality forecasting and uncertainty assessment via bootstrapped wavelet neural networks under missing data for Harbin, China. *Environ. Sci. Pollut. Res.* **20**, 8909–8923 (2013).
32. El Bilali, A., Taleb, A. & Brouziyne, Y. Groundwater quality forecasting using machine learning algorithms for irrigation purposes. *Agric. Water Manag.* **245**, 106625 (2021).
33. Arabgol, R., Sartaj, M. & Asghari, K. Predicting nitrate concentration and its spatial distribution in groundwater resources using support vector machines (svms) model. *Environ. Model. Assess.* **21**, 71–82 (2016).
34. Sajedi-Hosseini, F. *et al.* A novel machine learning-based approach for the risk assessment of nitrate groundwater contamination. *Sci. Total Environ.* **644**, 954–962 (2018).
35. Ransom, K. M., Nolan, B. T., Stackelberg, P., Belitz, K. & Fram, M. S. Machine learning predictions of nitrate in groundwater used for drinking supply in the conterminous united states. *Sci. Total Environ.* **807**, 151065 (2022).
36. Yadav, B., Gupta, P. K., Patidar, N. & Himanshu, S. K. Ensemble modelling framework for groundwater level prediction in urban areas of India. *Sci. Total Environ.* **712**, 135539 (2020).
37. Tomić, A. Š., Antanasijević, D., Ristić, M., Perić-Grujić, A. & Pocajt, V. A linear and non-linear polynomial neural network modeling of dissolved oxygen content in surface water: Inter- and extrapolation performance with inputs' significance analysis. *Sci. Total Environ.* **610**, 1038–1046 (2018).
38. Zhi, W. *et al.* From hydrometeorology to river water quality: Can a deep learning model predict dissolved oxygen at the continental scale?. *Environ. Sci. Technol.* **55**, 2357–2368 (2021).
39. Srinivas, R., Bhakar, P. & Singh, A. P. Groundwater quality assessment in some selected area of Rajasthan, India using fuzzy multi-criteria decision making tool. *Aquat. Procedia* **4**, 1023–1030 (2015).
40. Haghbi, A. H., Nasrolahi, A. H. & Parsaie, A. Water quality prediction using machine learning. *J. Water Qual. Res.* **53**, 3–13 (2018).
41. Liu, M. & Lu, J. Support vector machine-an alternative to artificial neuron network for water quality forecasting in an agricultural nonpoint source polluted river?. *Environ. Sci. Pollut. Res.* **21**, 11036–11053 (2014).
42. Chen, K. *et al.* Comparative analysis of surface water quality prediction performance and identification of key water parameters using different machine learning models based on big data. *Water Res.* **171**, 115454 (2020).
43. Sagan, V. *et al.* Monitoring inland water quality using remote sensing: Potential and limitations of spectral indices, bio-optical simulations, machine learning, and cloud computing. *Earth-Sci. Rev.* **205**, 103187 (2020).
44. Wu, Y., Zhang, X., Xiao, Y. & Feng, J. Attention neural network for water image classification under iot environment. *Appl. Sci.* **10**, 909 (2020).
45. Pu, F., Ding, C., Chao, Z., Yu, Y. & Xu, X. Water-quality classification of inland lakes using landsat8 images by convolutional neural networks. *Remote Sens.* **11**, 1674 (2019).
46. Donnelly, J., Daneshkhal, A. & Abolfathi, S. Forecasting global climate drivers using gaussian processes and convolutional autoencoders. *Eng. Appl. Artif. Intell.* **128**, 107536 (2024).
47. Abolfathi, S., Cook, S., Yeganeh-Bakhtiary, A., Borzooei, S. & Pearson, J. Microplastics transport and mixing mechanisms in the nearshore region. *Coast. Eng. Proc.* <https://doi.org/10.9753/icce.v36v.papers.63> (2021).
48. Stride, B., Abolfathi, S., Odara, M. G. N., Bending, G. D. & Pearson, J. Modeling microplastic and solute transport in vegetated flows. *Water Resour. Res.* **59**, e2023WR034653. <https://doi.org/10.1029/2023WR034653> (2023).
49. Unacademy (2022).
50. Başağaoğlu, H. *et al.* A review on interpretable and explainable artificial intelligence in hydroclimatic applications. *Water* **14**, 1230 (2022).
51. Habib, M., O'Sullivan, J., Abolfathi, S. & Salauddin, M. Enhanced wave overtopping simulation at vertical breakwaters using machine learning algorithms. *PLoS ONE* **18**, e0289318 (2023).
52. Mpia, H., Mburu, L. & Mwendia, S. Applying data mining in graduates' employability: A systematic literature review. *Int. J. Eng. Pedag.* **13**, 86–108. <https://doi.org/10.3991/ijep.v13i2.33643> (2023).
53. Raileanu, L. E. & Stoffel, K. Theoretical comparison between the gini index and information gain criteria. *Ann. Math. Artif. Intell.* **41**, 77–93. <https://doi.org/10.1023/b:amai.0000018580.96245.c6> (2004).
54. Gulati, P., Sharma, A. & Gupta, M. Theoretical study of decision tree algorithms to identify pivotal factors for performance improvement: A review. *Int. J. Comput. Appl.* **141**, 19–25. <https://doi.org/10.5120/ijca2016909926> (2016).
55. Tangirala, S. Evaluating the impact of GINI index and information gain on classification using decision tree classifier algorithm. *Int. J. Adv. Comput. Sci. Appl.* **11**, 110277. <https://doi.org/10.14569/ijcasca.2020.01110277> (2020).
56. Xu, P. Review on studies of machine learning algorithms. *J. Phys.* **1187**, 052103. <https://doi.org/10.1088/1742-6596/1187/5/052103> (2019).
57. Purwanto, A. D., Wikantika, K., Deliar, A. & Darmawan, S. Decision tree and random forest classification algorithms for mangrove forest mapping in Sembilang National Park, Indonesia. *Remote Sens.* **15**, 16. <https://doi.org/10.3390/rs15010016> (2022).
58. Huang, H. *et al.* A new fruit fly optimization algorithm enhanced support vector machine for diagnosis of breast cancer based on high-level features. *BMC Bioinform.* <https://doi.org/10.1186/s12859-019-2771-z> (2019).
59. Ji, Y. & Sun, S. Multitask multiclass support vector machines: Model and experiments. *Pattern Recogn.* **46**, 914–924. <https://doi.org/10.1016/j.patcog.2012.08.010> (2013).
60. Übeyli, E. D. ECG beats classification using multiclass support vector machines with error correcting output codes. *Dig. Signal Process.* **17**, 675–684. <https://doi.org/10.1016/j.dsp.2006.11.009> (2007).
61. Cortes, C. & Vapnik, V. Support-vector networks. *Mach. Learn.* **20**, 273–297. <https://doi.org/10.1007/bf00994018> (1995).
62. Ye, F., Lou, X. Y. & Sun, L. F. An improved chaotic fruit fly optimization based on a mutation strategy for simultaneous feature selection and parameter optimization for SVM and its applications. *PLoS ONE* **12**, e0173516. <https://doi.org/10.1371/journal.pone.0173516> (2017).
63. Peng, C.-Y.J., Lee, K. L. & Ingersoll, G. M. An introduction to logistic regression analysis and reporting. *J. Educ. Res.* **96**, 3–14. <https://doi.org/10.1080/00220670209598786> (2002).
64. Park, H.-A. An introduction to logistic regression: From basic concepts to interpretation with particular attention to nursing domain. *J. Korean Acad. Nurs.* **43**, 154. <https://doi.org/10.4040/jkan.2013.43.2.154> (2013).
65. Chen, H., Hu, S., Hua, R. & Zhao, X. Improved Naive Bayes classification algorithm for traffic risk management. *EURASIP J. Adv. Signal Process.* <https://doi.org/10.1186/s13634-021-00742-6> (2021).
66. Shen, J. & Fang, H. Human activity recognition using gaussian Naive Bayes algorithm in smart home. *J. Phys.* **1631**, 012059. <https://doi.org/10.1088/1742-6596/1631/1/012059> (2020).
67. Gramegna, A. & Giudici, P. SHAP and LIME: An evaluation of discriminative power in credit risk. *Front. Artif. Intell.* <https://doi.org/10.3389/frai.2021.752558> (2021).

68. Zaremba, L., Zaremba, C. S. & Suchenek, M. Modification of shapley value and its implementation in decision making. *Found. Manag.* **9**, 257–272. <https://doi.org/10.1515/fman-2017-0020> (2017).
69. Krishnan, S. R. *et al.* Smart water resource management using artificial intelligence;a review. *Sustainability*<https://doi.org/10.3390/su142013384> (2022).

### Acknowledgements

Acknowledgements should be brief, and should not include thanks to anonymous referees and editors, or effusive comments. Grant or contribution numbers may be acknowledged.

### Author contributions

All authors contributed equally in this research work.

### Competing interests

The authors declare no competing interests.

### Additional information

**Correspondence** and requests for materials should be addressed to S.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024



