





Article

Artificial Intelligence for Surface Water Quality Evaluation, Monitoring and Assessment

Rishi Rana ^{1,*}, Anshul Kalia ², Amardeep Boora ¹, Faisal M. Alfaisal ³, Raied Saad Alharbi ³,
Parveen Berwal ⁴, Shamshad Alam ³, Mohammad Amir Khan ⁴ and Obaid Qamar ⁵

- ¹ Department of Civil Engineering, Jaypee University of Information Technology, Waknaghat 173234, India; amardeep@juitsolan.in
- ² Department of Computer Science, Himachal Pradesh University, Summer Hills, Shimla 171005, India
- ³ Department of Civil Engineering, College of Engineering, King Saud University, Riyadh 11421, Saudi Arabia; rsalharbi@ksu.edu.sa (R.S.A.); salam@ksu.edu.sa (S.A.)
- ⁴ Department of Civil Engineering, Galgotias College of Engineering & Technology, Greater Noida 201310, India; parveenberwal@gmail.com (P.B.); amir.khan@galgotiacollege.edu (M.A.K.)
- ⁵ Department of Environmental Science & Engineering, Yeungnam University, Gyeongsan 38541, Republic of Korea
- * Correspondence: rishi.rana@juitsolan.in

Abstract: The study utilizes a dataset with seven critical constraints and creates models that are estimated based on various metrics. The goal is to categorize and properly predict the water quality index (WQI) using the suggested models. The outcomes show that the implied models can accurately assess water quality and forecast WQI with high rates of success. Temperature, pH, dissolved oxygen (DO), conductivity, total dissolved solids (TDS), turbidity, and chlorides (Cl⁻) are some of the six crucial factors used in the study's dataset. The mean absolute error (MAE), mean squared error (MSE), and coefficient of determination (R²) are some of the metrics used to develop and assess the Artificial Neural Networks (ANN) and Long Short-Term Memory (LSTM) models. The study also makes use of heat maps and correlation graphs to shed further light on the connections between various water quality measures. The color-coded values of the seven parameters, which represent the water quality level of the sample, are displayed on the heat map. The link between the two parameters is shown by the correlation graph between TDS and turbidity, which depicts their correlation coefficient. The study's results show how effective machine learning algorithms may be as a tool for observing surface water quality. Himachal Pradesh is the tourist hub, so with the rapid increase in the volume of surface water contamination, the application of artificial intelligence will give a better view of data analytics and help with prediction and modeling. It was obtained from the study that the mean square error and root mean square error of ANN and LSTM lie between 0.52–6.0 and 0.04–0.21, respectively. However, the LSTM model's accuracy is 95%, which is higher than the ANN model. The study highlights the importance of leveraging machine learning techniques in water quality monitoring to ensure the protection and management of water resources. With advancements in machine learning, artificial intelligence (AI) techniques have emerged as a promising tool for surface water quality monitoring. The major goal of the study is to explore the potential of two types of machine learning algorithms, namely artificial neural networks (ANNs) and long short-term memory (LSTM) models, for surface water quality monitoring.

Keywords: artificial intelligence; surface water; water quality index; neural networks



Citation: Rana, R.; Kalia, A.; Boora, A.; Alfaisal, F.M.; Alharbi, R.S.; Berwal, P.; Alam, S.; Khan, M.A.; Qamar, O. Artificial Intelligence for Surface Water Quality Evaluation, Monitoring and Assessment. *Water* **2023**, *15*, 3919. <https://doi.org/10.3390/w15223919>

Academic Editor: Christos S. Akratos

Received: 17 August 2023

Revised: 18 September 2023

Accepted: 1 November 2023

Published: 9 November 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Even though fresh water covers 34% of the earth's exterior, just 1% of fresh water is available to humans. Global population growth and industrialization are also contributing to the spread of pollutants in water bodies. As a result, it is critical to continuously screen the quantity of water from common and unique sources [1,2]. The quality of the water

tested and evaluated is attainable by the verification of the framework. Water quality can be checked by organic records or physiochemical parameters. Advancing from conventional water quality checking and evaluation strategies, Web of Things-enabled technologies, and the utilization of artificial insights, which are unused domains being investigated, artificial intelligence was presented in computer science in the 1950s and has undergone substantial changes in enhancement and modernity. The scientific commitment of the research is to illustrate a study on the application of different sorts of neural networks to the surface quality of water to finalize the various strategies utilized in the area of the test where the input parameter will be utilized [3–7]. Most water resources, such as rivers, ponds, and tributaries, are subjected to strict purity standards. Various water standards exist for diverse purposes and applications. For instance, irrigation water should not have excessive salinity or hazardous substances that could be transmitted to plants or soil, posing risks to ecosystems. The specific attributes necessary for industrial water quality vary depending on the type of industrial activity. When it comes to drinking water, natural water sources like groundwater and surface water are considered highly preferable. Pollution of such resources can occur because of human or engineering activities as well as other ecological activities [8,9]. As a result, increasing industrial expansion has accelerated the degradation of water quality. Furthermore, infrastructure has a significant impact on drinking water quality due to a shortage of community awareness and less sanitized elements. Undoubtedly, the concerns of polluted water sources are exceptionally detrimental, presenting a significant risk to individual health, the environment, and societal structures [3]. According to data from the United Nations, approximately 1.5 million individuals lose their lives each year due to illnesses resulting from tainted water. In poor nations, it is estimated that 80% of health issues stem from water contamination. Each year, there are five million reported casualties and 2.5 billion instances of sickness related to this issue. These statistics reveal a higher mortality rate compared to deaths resulting from accidents, crimes, or acts of terrorism. Massive population growth, industrial innovation, and the usage of manure and fungicides have all had a negative impact on water quality (WQ) ecosystems [10–12].

As a vital natural resource, surface water is essential for maintaining environmental health, economic activity, and human existence. Any water that is present on the top of the earth's surface, such as rivers, lakes, ponds, wetlands, and seas, is referred to as a surface water source [13–19]. The primary sources of surface water are precipitation and runoff from higher altitudes. Snow melts in the spring when the climate warms, and the water that results rushes into surrounding streams and rivers, making a large contribution to the world's supply of drinking water. Surface water is not always readily available in different areas and at different times of the year, and both human activity and natural processes can have an impact on its quality [1,2,4,5,10,20–46]. Human consumption is one of the key uses of apparent water, especially in places where groundwater supplies are scarce or difficult to reach. Various governmental organizations estimate that 68% of the water that is provided to humans globally originates from surface water. Surface water must be treated to eliminate impurities such as bacteria, viruses, chemicals, and other pollutants before it is safe to drink [15,19–24]. The irrigation of crops, especially in agriculture, is a substantial additional use of surface water. Additionally, surface water is used for leisure pursuits including swimming, boating, and fishing, as well as industrial processes, hydropower production, and cattle irrigation. But there is a growing need for surface water for a variety of purposes, and there are consequences of climate change, such as droughts and floods [25–28]. Evaporation and infiltration, when water seeps into the earth and turns into subterranean water, can also cause surface water levels to drop. Particularly in regions with few surface water supplies, groundwater may be a substantial source of water for human use. However, excessive groundwater pumping can result in pollution and depletion, which can have an impact on the environment and individual fitness. A variety of human activities, such as agriculture, industrial operations, and home wastewater can also have an influence on the superiority of surface water [11,18]. In addition, pollution

from industrial operations and wastewater can introduce chemicals and other pollutants into the water, causing eutrophication and the creation of toxic algal blooms. Nutrient runoff from fertilizers and animal manure can also contribute to this problem. Several measures, including water conservation techniques, wastewater treatment, and rules to regulate pollution and safeguard water quality, have been put in place to guarantee the sustainable use of surface water. A further way to guarantee the fair and effective use of surface water resources is to implement integrated water resource management methods, such as watershed management and water allocation plans. In conclusion, surface water is an important natural resource that supports both ecosystems and people by offering vital functions. For water security, human health, and environmental sustainability, it must be used and managed sustainably [4,10]. It is crucial to put into place efficient policies and practices in order to preserve and safeguard surface water resources for both current and future generations. The quality of surface water can be impacted by a wide range of contaminants. For instance, nutrient contamination is a serious issue in many aquatic environments. When surface waterways obtain an excessive amount of nitrogen and phosphorus from fertilizers, sewage, and other sources, algal blooms can result, which can reduce oxygen levels and endanger aquatic life. Another major issue is sediment contamination, which is brought on by soil erosion into surface waterways, damages aquatic life, and causes siltation and turbidity [22]. Another significant problem for surface water quality is chemical contaminants such as industrial chemicals, medicines, and pesticides. These pollutants, which can endanger aquatic life and human health, can enter streams through spills, runoff, or direct discharges [7,29–31].

There are several management techniques that may be used to preserve and enhance the quality of surface water. Limiting the use of chemicals and other pollutants is one way to decrease pollution at the source. To minimize storm water runoff, for instance, people can use fewer fertilizers, pesticides, and other chemicals surrounding their houses, and companies can utilize green infrastructure techniques like rain gardens and green roofs. Treatment of contaminants before they enter surface waterways is an alternative strategy. This might entail filtering contaminants and enhancing water quality utilizing a variety of methods, including sediment basins, wetlands, and artificial treatment wetlands [32,33]. Pollutants can often be avoided by using source control measures, such as covering storage places and reducing industrial discharges [6].

Therefore, surface waters are essential for preserving natural systems and sustaining human existence. However, contamination from a multitude of sources may swiftly degrade their quality. Reducing pollution at the source and using various treatment technologies to filter pollutants can enhance water quality to safeguard and amend the superiority of surface water. We can ensure that surface waters continue to deliver important advantages for future generations by adopting these steps [16,34].

Artificial intelligence (AI) has made a difference analyst accomplish the plausibility of imitating human behaviour abilities in specific spaces of knowledge [3–5]. AI instruments include sloppy logic, particle swarm optimization, algorithmic genetics, artificial neural networks, assistance vector machine, ad boost algorithm, etc. The sustainability of ecosystems, human health, and economic activities all depend on the quality of surface water. Surface water quality monitoring and evaluation have historically relied on labour-intensive, expensive, and time-consuming laboratory studies and manual sampling [6,7]. This is because AI models are capable of processing large amounts of data quickly and accurately, allowing for the identification of trends and patterns in water quality data that may be difficult to detect using traditional methods. Artificial intelligence (AI) and other emerging technologies provide the potential to fundamentally alter how we measure, track, and evaluate the quality of surface water [8,9]. A promising approach to improving the evaluation, monitoring, and assessment of surface water quality is offered by artificial intelligence. By utilizing its potential, we may better understand our water systems and develop proactive management approaches. To ensure the comprehensive and ethical

management of water resources, AI should be implemented, but with consideration for its limitations and in conjunction with human expertise [10–12].

Having models for predicting the WQ is therefore quite useful for drawing conclusions about water contamination. Currently, there are two primary types of models employed for the purpose of demonstrating and assessing water quality: machinery-oriented models and non-machinery-oriented models. The machinery-oriented model stands out for its advanced nature, as it replicates water quality using data derived from the headway system infrastructure. This versatility allows it to be applied to various water bodies. Among the early simulation models for water quality, the widely utilized Streeter–Phelps (S-P) model deserves mention. Researchers have extensively studied the water quality of Lake Galaa in Turkey, employing techniques such as satellite image fusion and principal component analysis (PCA). In another instance, the water quality of the Narmada River was predicted using a decision tree technique incorporating five water quality indicators. Additionally, a study proposed the utilization of deep fractional stacked simple recurrent units (Bi-S-SRU) for the development of an accurate water quality forecasting system in smart agriculture [13–16].

Keeping in view of the above, this study focuses on reviewing the effectiveness of the AI models in water quality monitoring depending on the range of appropriate key issues, like the quality of the data and the simulation's training. The outcomes of this investigation have meaningful suggestions for legislators and stakeholders involved in water resource management. The use of AI models in surface water quality monitoring can help elaborate actual approaches for water resource control, ensure the availability of safe and clean water for all, and prevent water pollution. The leading area of this consideration was to gather measurable data regarding the physical, chemical, and biological properties of water by conducting water sampling. The aim was to employ machinery knowledge algorithms to analyze the classification of water quality and determine the water quality index. Various artificial intelligence models, including artificial neural networks (ANN) and long short-term memory (LSTM) deep-rooted learning algorithms, were utilized for this purpose. The significance of the study is justified considering that the accessibility of clean water could be the basic economic advancement objective, and the neural system in water quality checking and evaluation is a generally novel area to investigate.

2. AI in Water Quality Monitoring

Artificial intelligence (AI) states the examination and exploration of computer arrangements accomplished to execute the responsibilities stereotypically associated with social intellect, such as graphic awareness, dialog acknowledgement, choice making, and understanding ordinary linguistic [11]. The concept of AI dates to ancient philosophers who were interested in the systematization of reasoning. However, it was not until the development of programmable computers in the 19th century that the focus shifted towards the possibility of creating intelligent machines. AI is a broad and swiftly evolving technology field that encompasses a range of sub-disciplines, including natural language processing, computer vision, robotics, and cognitive computing [17–19]. The technology has to be imminent to transmute numerous activities, including healthcare, economics, manufacturing, and transportation. However, it also enhances ethical, social, and economic questions, such as the effect on occupation and privacy, the capability for preference and perception, and the decent use of AI in decision-making [35].

AI has a rich history that dates to ancient times. The development of programmable computers in the 20th century, coupled with advancements in various fields, led to the creation of topological works that described how machines could be designed to think [3,13,37]. AI is a fast-expanding science that has the potential to disrupt many sectors, but it also poses significant ethical, societal, and economic issues that must be addressed. In surface water quality monitoring, AI Technologies discovered that employing a combination of physical-chemical and biological criteria did not produce excellent results. The fast evolution of the Internet of Things (IoT) in radar, wireless interaction, and trade

IoT is enhancing more presumed to be the next-generation option of control [38]. Many researchers [7,8,10,11,21,33,35,38,39] have used the application of IOT for water quality observing systems and concluded that parameters like pH and TDS for different types of water—salty, mud, drinking, and tap water—showed varied results. The Internet of Things is crucial for increasing trade effectiveness and superiority while cutting trade costs and supplies. Conversely, there have been few openly published real-world IoT project applications thus far. AI for Surface Water Excellence Observing and Estimation demonstrated the various models used for water excellence observing and performed a literature review for previous research [18]. Many studies [11,12,15,17,19,20,38] have also discussed different types of artificial models and the models that can be used to calculate water quality index. Because the predictor parameters of these models can be measured fast, the BOD value may be predicted promptly. Few of the researchers have highlighted the AI approach to predict river water quality and predicted the results using the parameters like BOD, COD, EC, TDS, and turbidity. Studies have shown that the expectation of groundwater level (GWL) using geoelectric properties is one of the trickiest puzzles to solve. It is partly because there is not yet a concrete empirical connection between the amount of groundwater and the geoelectric parameters. This study investigated the ability of advanced artificial neural networks (ANNs) to model nonlinear systems to get around these problems [40].

Water quality monitoring (WQM) parameters like turbidity, temperature, pH, electrical conductivity, oxidation, etc. are essential for depicting the ideal nature of water sources. To find solutions that are physically accurate, it is important to formulate the problem more precisely than has previously been the case in the literature and to represent the underlying processes realistically. It successfully integrates data models, makes wise decisions, does dynamic optimization, and controls. Researchers [5,31] have predicted that using a CNN model detected algae and foam present in water and it was concluded that the model used gave appropriate results. Contaminants are eliminated by the procedure, which then turns them into waste matter that may either be supplied to the water supply or immediately recovered. Studies have shown the comparison of different types of artificial models like ANFIS and ANN [28,34]. It also indicated which models were more accurate at predicting the water quality categorization and index. The modeling methodology also helps in achieving a variety of other parameters like data–model integration, sound decision-making, dynamic optimization, and control, which help in more accurate result description. The Internet of Things (IoT) and smart grid play essential roles in encouraging and guiding information technology and economic growth. IoT applications are now expanding quickly, but some of them have specific criteria that present technology cannot provide. IoT is the focus of a lot of research. Wi-Fi-based wireless sensor networks (WSNs) are capable of non-linear transmission, large-scale data gathering, good cost-effectiveness, and video monitoring, in addition to having high bandwidth and rate [41].

To obtain the most information from the water quality data gathered, the design of a network for monitoring water quality is a difficult process that requires the best configuration. The network design should ideally consider the specific monitoring objectives, representative sampling size, location, and frequency, water quality variable selection, as well as logistical and financial limitations [25]. A workable and simple technique for designing a water quality monitoring network will provide a reliable, effective, and affordable design. Anomalies in water can be detected in real time using multi-sensor systems. While the set of sensors varies depending on the application, the overall principle stays the same. This technology might be used in a wide range of applications, including surface water, urban runoff, food and industrial process water, aquaculture, and several other sectors where water is utilized and reused. The creation and development of AI techniques using ANNs give unique ways in a variety of domain domains; nevertheless, their specific application can provide novel approaches to increasing water quality efficiently and effectively [20,28,36,42].

3. Study Area

The study area chosen for the research was Ashwini Khud. The geographical location of the study area is represented in Figure 1. The study of water samples from the River Ashwini in Sadhupul, Himachal Pradesh, is of utmost significance for assessing the quality of drinking water in the region. The Ashwini River is the primary source of water supply for many villages and towns in the Solan district, and therefore it is essential to determine its suitability for consumption. The substantial and compound attributes of the river water could change depending on the geographical features of the area, and these could have an impact on the water's suitability for drinking purposes. The study includes an analysis of the water samples for various parameters involving pH, total dissolved solids, electrical conductivity, turbidity, and microbial contamination. These parameters could indicate the specter of harmful contaminants such as substantial metals, pesticides, and bacteria, which could make the water unfit for human consumption. The results of this study will help to identify the parameters that indicate the presence of harmful contaminants such as heavy metals and bacteria, which could make the water unfit for human consumption, which in turn can help to identify the potential health risks associated with drinking water from the Ashwini River and to take appropriate measures to improve the water quality if required. It is crucial to ensure the welfare and superiority of drinking water in the region to prevent waterborne diseases and promote the overall health and wellbeing of the local population.

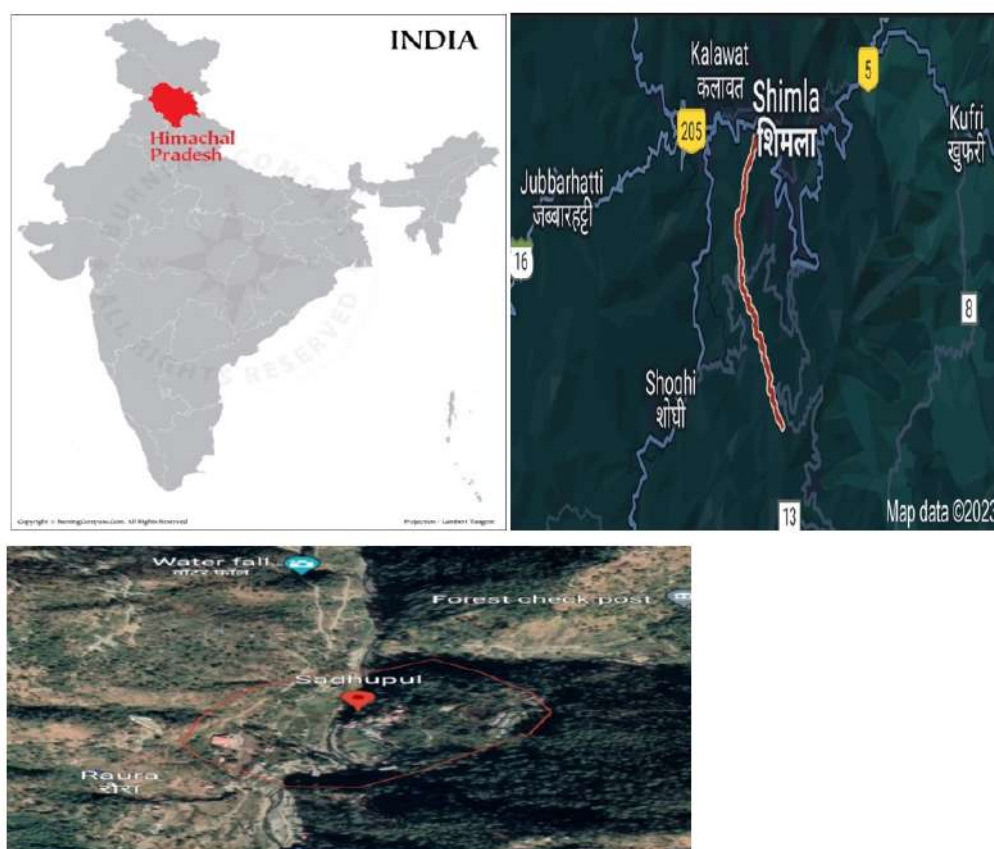


Figure 1. Location of study area showing the surface water source.

Data Collection and Treatment

The quality of water is an essential aspect to consider as it directly affects human health, agriculture, industry, and the environment. The study focused on monitoring twelve water quality measures for six months to determine the suitability of the water for different purposes. Out of the twelve measures tracked, only six variables were chosen for the study based on their significant influence on the traits of water quality. The pH, hardness, total dissolved solids, chlorides, turbidity, and dissolved oxygen are important

indicators of water quality, and their values might alter the water’s usefulness for various purposes. The study used a dataset consisting of 200 water samples to monitor these parameters. Although the dataset is limited, it provides important insights into the water quality of the area and could be useful in identifying trends and patterns in the data as given in Figure 2. The findings of this study might aid in determining the appropriateness of water for various applications, including drinking, irrigation, and industrial usage. The investigation might also assist in identifying potential sources of contamination and establishing measures for protecting and improving the area’s water quality. In conclusion, this study is crucial in determining the quality of water in the area and provides valuable information on the parameters that significantly impact water quality. This information could be used to develop policies and strategies to improve water quality, promote public health, and protect the environment.

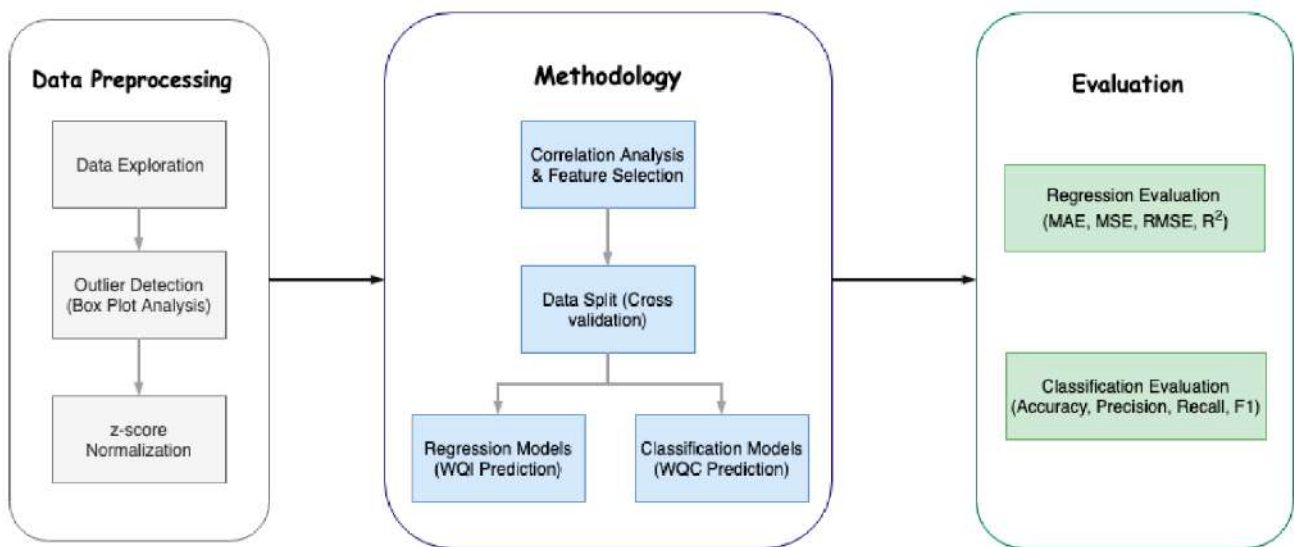


Figure 2. Highlights of the current strategy of suggested approach [46].

4. Artificial Neural Network (ANN) Model

ANN is a type of machine learning algorithm that is capable of learning complex patterns in data, making it useful for identifying trends and patterns in water quality data. After receiving data from a variety of different neurons and mathematically processing it, ANNs are composed of connected, layered neurons that send the results to neurons in the layer below to generate the output as shown in Figure 3 [10,17,32,35,43].

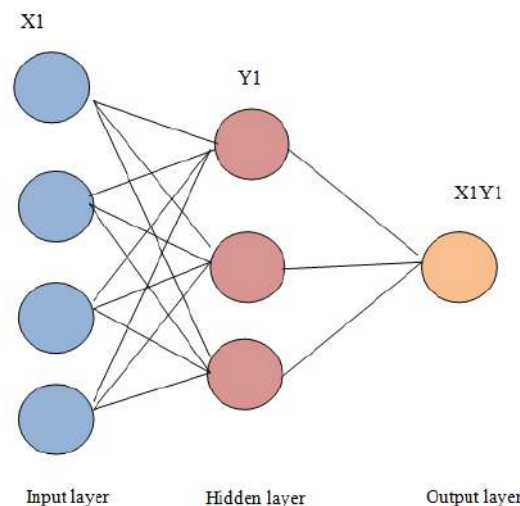


Figure 3. ANN portrayal model [21].

The strength of these connections between neurons can be adjusted based on the data that the network is trained on, allowing the network to learn and improve its performance over time. The basic unit of an ANN is the artificial neuron, also known as a perception. A perception takes in one or more inputs, multiplies each input by a weight, adds them together, and applies an activation function to produce an output [44]. The activation function is usually non-linear, allowing the network to learn complicated correlations between inputs and outputs. There are several types of ANNs, including feed forward networks, recurrent networks, and convolution networks. Feed forward networks are the simplest type of network, and they are used for tasks like classification and prediction. Recurrent networks are designed to process sequences of data, and they are commonly used in tasks like speech recognition and natural language processing. Convolution networks are used for tasks like image and audio recognition, and they are designed to detect patterns in data that are spatially or temporally localized [25,31,37]. Training an ANN involves adjusting the weights between neurons so that the network produces the desired output for a given input. This is typically carried out using a process called back propagation, which involves computing the error between the network's output and the desired output and then using that error to adjust the weights in the network. This process is repeated over many iterations, and the network's performance gradually improves as the weights are adjusted to minimize the error.

Image and audio identification, natural language processing, and financial modelling are just a few of the uses for ANNs. They have been particularly successful in tasks like object recognition and speech recognition, where they have achieved human-level performance in some cases [11,14,33,40,44]. However, training ANNs may be computationally costly and requires a huge quantity of data to attain decent performance. Additionally, they can be difficult to interpret, which can make it challenging to understand how the network is making decisions. In conclusion, ANNs are a powerful type of machine learning model that are inspired by the way the human brain works. They have been successful in a wide range of applications, but they can be computationally expensive to train and difficult to interpret [26]. An artificial neural network (ANN) may be used to forecast and monitor water quality. The following are the steps for developing an ANN model for measuring water quality.

Data collection: To gain facts on the properties of water characteristics, data can be collected from various sources such as rivers, lakes, and wells. The data can be gathered through manual sampling or automated monitoring systems. Parameters that can be measured include temperature, pH, dissolved oxygen content, and pollutants. For example, temperature can be measured using thermometers or temperature probes, pH can be measured using pH meters, dissolved oxygen can be measured using oxygen sensors, and pollutants can be measured using analytical instruments such as spectrophotometers or gas chromatographs. Data can also be collected from government agencies or research organizations that monitor water quality, such as the Environmental Protection Agency or the US Geological Survey [25,26]. Collecting comprehensive and accurate data on water characteristic is essential for certifying the welfare of consumption water, protecting aquatic ecosystems, and monitoring the impacts of human activities on water resources.

Data preprocessing: Data preprocessing is a crucial step in preparing data for artificial neural network (ANN) analysis. It involves cleaning the data to remove errors and inconsistencies, and normalizing it to establish a standardized format suitable for analysis. This includes identifying and removing missing values, outliers, and irrelevant data points. Normalization techniques such as scaling or standardization are used to ensure that all features are on a similar scale, allowing the ANN to learn the patterns in the data more effectively [41]. A well-preprocessed dataset is essential for accurate and effective ANN analysis.

Data splitting: Once data preprocessing is completed, the dataset is split into three sets: training, validation, and testing. The training set is used to train the ANN model, while the validation set is used to adjust model parameters and prevent over fitting [25]. Finally,

the testing set is used to evaluate the performance of the trained model on unseen data [9]. This approach ensures that the performance of the ANN model is not overly influenced by the training data and can generalize to new, unseen data.

ANN structural design: The structural design of an artificial neural network (ANN) involves creating an input layer, one or more hidden layers, and an output layer [25]. The number of nodes in each layer and the activation functions used can be optimized using techniques such as grid search and cross-validation. Grid search involves systematically testing different combinations of hyper parameters to identify the optimal configuration, while cross-validation involves evaluating the performance of the model on different subsets of the data to prevent over fitting.

Model justification: The performance of an artificial neural network (ANN) is justified by evaluating its accuracy, precision, recall, and other metrics on a separate validation set. If necessary, the ANN's settings can be modified to improve its performance, such as adjusting the number of hidden layers or nodes, changing the learning rate, or using different activation functions [38,45]. The goal is to optimize the ANN to achieve the highest possible accuracy on unseen data while avoiding over fitting.

Model testing: To analyze the performance of an artificial neural network (ANN) model, a testing set can be used to evaluate its F1-score, accuracy, precision, and recall. Additionally, other machine learning models like k-nearest neighbor (KNN) and decision tree (DT) can be used to compare their performance with that of the ANN. For both classification and prediction problems, KNN and DT models can provide insights into the relationships among variables and may be used to identify the most important features. By comparing the performance of these models, it is possible to identify the most accurate and effective approach to solving the given problem statement [8,9,13]. Using the steps mentioned above, an artificial neural network (ANN) model can be developed and deployed to regulate and monitor water quality, ensuring the security and sustainability of water supplies. The ANN can be trained on data collected from various sources, preprocessed, and validated using testing and validation sets. Finally, the ANN's performance can be analyzed and compared to other machine learning models to identify the most accurate and effective approach for water quality monitoring and regulation.

The equation of the simulation is exhibited as Equations (1) and (2)

$$R^2 = 1 - \frac{\sum(x - y)^2}{\sum y^2 - \frac{y^2}{n}} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n} \sum n(x - y)^2} \quad (2)$$

where x represents the detected data, y is the expected data and n is the number of observations [1,2].

The network architecture of the model is designed to facilitate a structured flow of information. Input signals, representing independent variables, are directed to the hidden layer for processing and are then transmitted to the output layer through a network of weighted connections.

LSTM (Long Short-Term Memory)

The primary objective behind the development of recurrent neural networks (RNNs) incorporating long short-term memory (LSTM) is to overcome the problem of vanishing gradients encountered in traditional RNNs. LSTM is a type of neural network that is particularly useful for processing sequential data, making it well suited for time-series analysis of water quality data. In traditional RNNs, when the error gradient in the backpropagation process diminishes significantly, it becomes challenging for the network to learn long-term relationships [17,19]. To tackle this issue, LSTM models incorporate a memory cell that can selectively retain, or input information based on the input data. The architecture of an LSTM includes an input layer, an output layer, and one or more LSTM layers with memory

cells as shown in Figure 4 [14]. Each memory cell consists of three gates: an input gate to regulate the flow of new input data, a forget gate to determine which data to retain or discard, and an output gate to control the flow of output data. The gates in LSTM models are regulated by sigmoid activation functions, which produce values ranging from 0 to 1. A value of 0 indicates “forget” or “closed,” while a value of 1 signifies “input” or “open” [14,35].

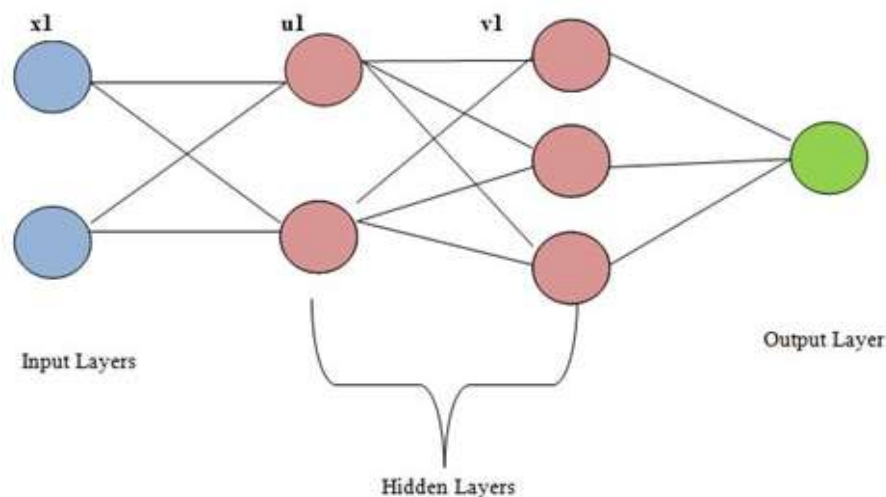


Figure 4. Schematic presentation of descriptive LSTM model [14].

LSTMs have demonstrated superior performance compared to traditional RNNs and other machine learning techniques across various applications. This has made them a popular choice for tasks involving time series analysis and prediction. The following steps outline the implementation of an Artificial Neural Network (ANN) model for monitoring water quality.

Collection of data: To obtain details on the properties of water quality, data can be collected from various sources, such as rivers, lakes, and wells. The data can be gathered through manual sampling or automated monitoring systems. Parameters that can be measured include temperature, pH, dissolved oxygen content, and pollutants as shown in Table 1. For example, temperature can be measured using thermometers or temperature probes; pH can be measured using pH meters; dissolved oxygen can be measured using oxygen sensors; and pollutants can be measured using analytical instruments such as spectrophotometers or gas chromatographs.

Table 1. Standard values of parameters according to WHO [46].

Parameter	Range
pH	6.5–8.5
Hardness	300 mg/L
TDS	500 mg/L
Chlorides	10 mg/L
Turbidity	Below 1 NTU
Dissolved oxygen	6.5–8 mg/L

Water quality index and classification: WQI may be used to evaluate the water’s quality by using measured values for various parameters that impact it. The experiment involved

measuring the nine previously indicated factors, which were then utilized to calculate the WQI (Equation (4)).

$$WQI = \frac{\sum_{i=1}^N qi \times xi}{\sum_{i=1}^N xi} \quad (3)$$

In the given expression, N represents the total number of parameters, qi signifies the quality rating scale assigned to each parameter, and xi denotes the corresponding unit weight assigned to each parameter [20].

The following equations can be used to calculate qi and xi (Equations (4) and (5)):

$$qi = 100 \times \frac{(Pi - Pideal)}{(Si - Pideal)} \quad (4)$$

$$xi = \frac{K}{Si} \quad (5)$$

In the context provided, Pi represents the measured values of parameters, P ideal represents the ideal values of parameters, and Si represents the standard values of parameters [14,31].

Preprocessing method: Data normalization is a crucial step in data preparation for machine learning. The objective of normalization is to rescale input values and output variables to a standardized scale, enabling consistent and comparable comparisons. One of the most used normalization methods is min-max normalization, which scales input variables to an average, with the range containing only ones and zeros. To perform min-max normalization, the lowest and greatest values of each variable are identified, and the values are rescaled to lie between 0 and 1. This is carried out by deducting the least value from every value and splitting by the differentiation concerning the greatest and lowest values [4,30]. This results in a new set of values that are all within the range of 0 to 1. Overall, data normalization is critical for machine learning because it ensures that each variable is given equal weight during model training. Without normalization, variables with large ranges may dominate the training process, resulting in suboptimal model performance Equation (6).

$$x = \frac{x - xmin}{xmax - xmin} \quad (6)$$

Performance Measurement: The study of artwork involves the use of various metrics, including mean square error (MSE), root mean square error (RMSE), mean absolute error, and correlation coefficient. These metrics are used to evaluate the performance of machine learning models that analyze artwork, such as those used for image classification or style transfer. They help assess the accuracy of the models and identify areas for improvement.

Mean Square Error: Equation (7).

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_{iobserved} - y_{iestimated})^2 \quad (7)$$

In this context, yi represents the observed value and the estimated value.

Rootmean square error: Equation (8).

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_{iobserved} - y_{iestimated})^2}{n}} \quad (8)$$

Coefficient of Correlation: Equation (9).

$$R = \frac{n(\sum_{i=1}^n y_{iob} \times y_{iest}) - (\sum_{i=1}^n y_{iob})(\sum_{i=1}^n y_{iest})}{\sqrt{n(\sum_{i=1}^n y_{iob})^2 - (\sum_{i=1}^n y_{iest})^2}} \quad (9)$$

5. Results and Discussion

It was investigated if artificial intelligence algorithms could replace more traditional techniques for estimating and forecasting water quality. Because of the demonstration and prediction of water attributes, the time and resources needed for laboratory analysis have greatly and crucially decreased. The SES preprocessing approach and updated LSTM and ANN simulations were used to predict water superiority and anticipate the features of water quality in surface water. In this study, we have compared two distinct models, i.e., the artificial neural network model and the long short-term memory model. The ANN model presents the data in the form of histograms that show us the correlation between different parameters. But, in the case of the LSTM model, it tells us about the water quality index, MSE, and RMSE. We can also test the model's accuracy by using two different classifications: the KNN (k-nearest neighbor) and the DT (decision tree) models.

With the use of a potent artificial intelligence model, the main goal is to initiate a real-time approach and test a fresh strategy for accurately anticipating and classifying water quality [30]. The study proposes merging the discussed artificial intelligence methods to precisely duplicate water levels and quality. The dataset had a total of six parameters. The study concluded that categorization and forecasting of water quality may be performed using LSTM and ANN models. The principle of this study was to show how the LSTM and ANN models may be used to forecast the quality of surface water.

Heat Map: Monitoring water quality is a crucial part of maintaining and defending our water resources. Data on many aspects of water quality, including pH, temperature, dissolved oxygen, turbidity, and nutrient concentrations, are gathered and analyzed during monitoring. The heat map is a helpful tool for visualizing and examining data on water quality. In a heat map, values are represented graphically by colors, with greater values denoted by warmer hues like red and lower values denoted by cooler hues like blue [11,26]. Heat maps can be used in water quality monitoring to show the geographical and temporal fluctuations in water quality parameters. Finding problem regions or hotspots is one of the main uses of heat maps in water quality monitoring. The heat map may display regions with high or low values for each parameter by showing water quality data on a geographic map. This makes it simple to pinpoint places where water quality may be impaired and where more research or intervention may be required.

Heat maps may also be used to evaluate the success of water quality control plans. The efficiency of various management techniques may be assessed by contrasting heat maps from various time periods, and changes in water quality can be connected to particular treatments. Additionally, the management and protection of water quality can be prioritized using heat maps. Resources can be directed towards implementing tactics to enhance water quality in areas with low water quality by identifying these places. The color code referred to in the statement is likely a color-coded representation of water quality parameters in a histogram or similar visual display. The range of values for this color code is -0.2 to 1.0 , with darker colors indicating negative effects on the corresponding parameter. The statement notes that most of the colors in the histogram are light, which suggests that the quality of surface water in that area is good. This could indicate that the water quality parameters being measured are within acceptable ranges and that there are no significant negative impacts on the water quality. Overall, color-coded visual displays of water quality data can be a useful tool for quickly and easily identifying areas of concern or areas where water quality is good. They can aid in decision-making for water management and protection and help to ensure that our water resources remain safe and healthy for both human use and the environment.

Figure 5 discusses the correlation of two different parameters. One can take the example of TDS and turbidity. An illustration of the link between these two indicators of water quality is a correlation graph between TDS (Total Dissolved Solids) and turbidity (Figure 6). Turbidity is the cloudiness or haziness of the water brought on by suspended particles, whereas TDS is the quantity of dissolved solids in the water. An outlier or other anomaly in the data can be found using a correlation graph. It may be a sign that there are

additional variables influencing the link between TDS and turbidity, such as the presence of pollutants or other contaminants, if, for instance, most of the data points on the graph follow a distinct pattern but a small number of points fall outside of this pattern.

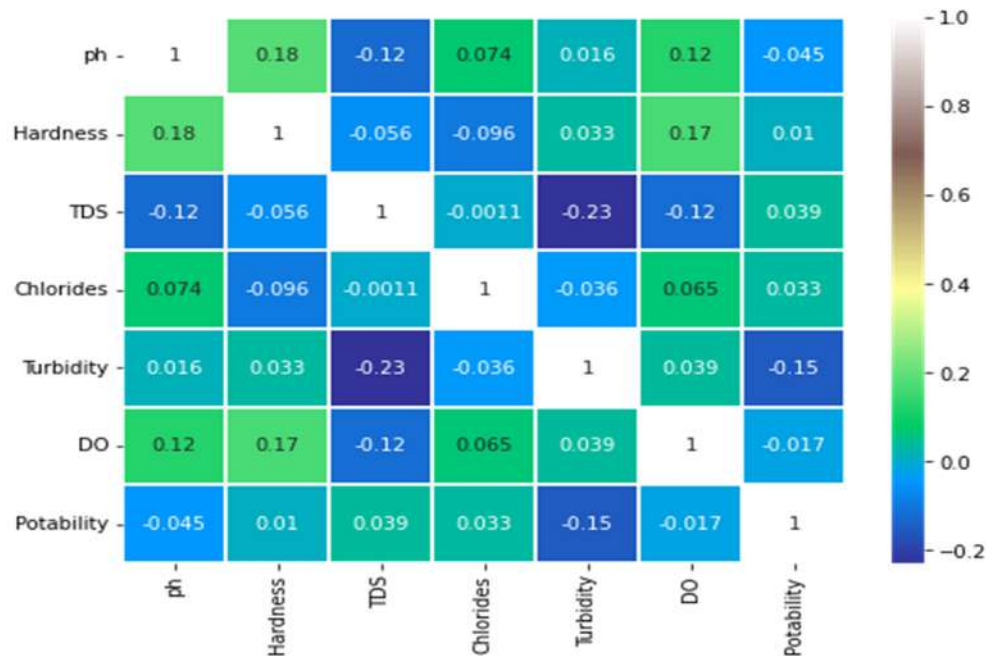


Figure 5. Schematic view of heat map generated with ANN model [16].

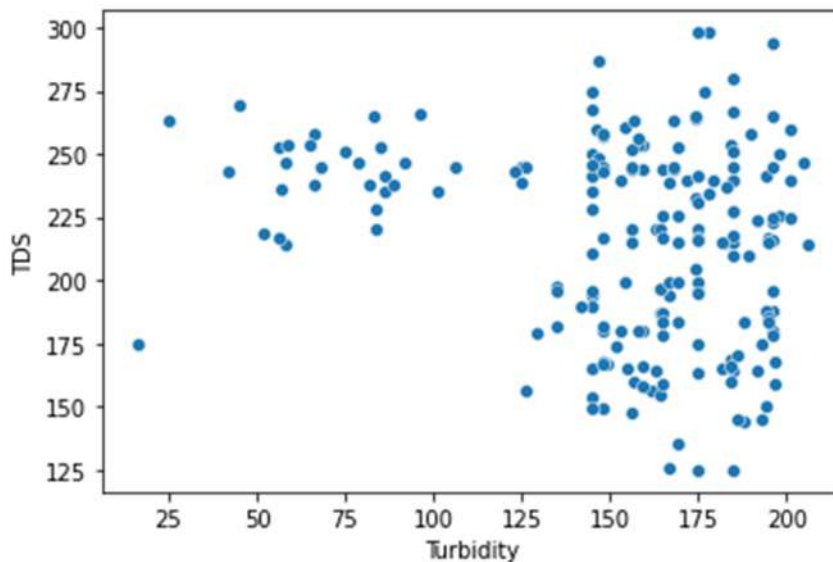


Figure 6. Observed correlation between TDS and turbidity of surface water samples [21].

Distplot graph: The distplot displays the data distribution of a single variable in comparison to the density distribution as shown in Figure 7.

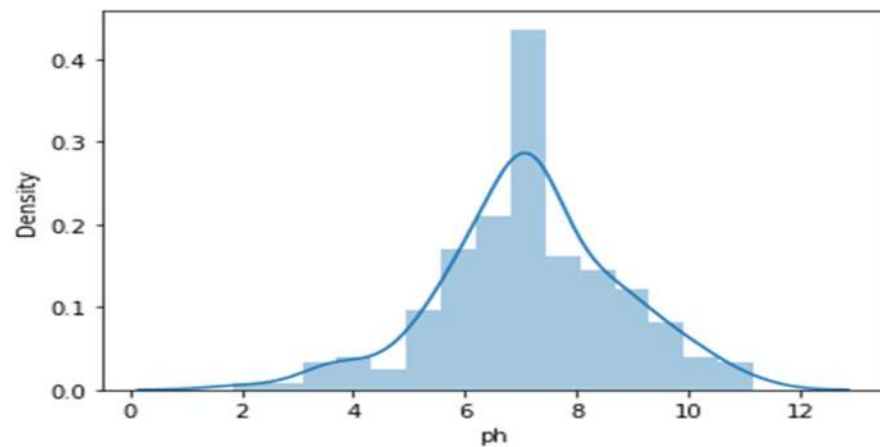


Figure 7. Comparison analysis between density and pH value using distplot graph [26].

Boxplot graph: Boxplots are employed to assess the distribution of data within a dataset and determine their level of dispersion. They depict key statistical measures such as the minimum, maximum, median, first quartile, and third quartile of the dataset, creating three distinct quartiles as shown in Figure 8.

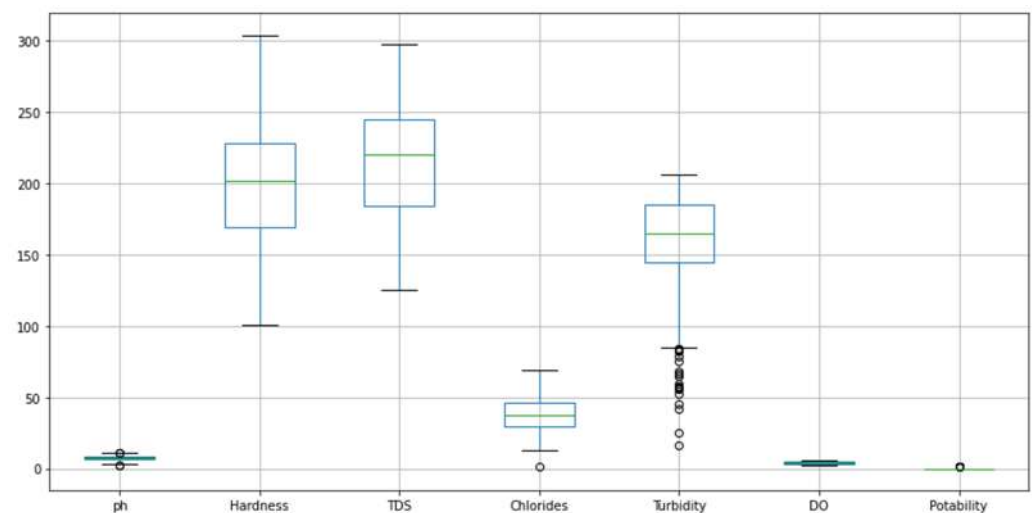


Figure 8. Measurement of data distribution using Boxplot graph [22].

The aim of this study is to find the accuracy of models. Two classifiers were used to find the model accuracy of the ANN model.

Decision Tree classifier: It is a type of machine learning algorithm that is used for classification tasks. It functions by creating a tree-like representation of decisions and potential outcomes. The tree is made up of leaf nodes, which represent the output class or category, and interior nodes, which reflect judgments depending on the values of one or more input attributes. Beginning at the root node, the decision tree classifier determines a course of action depending on the value of a single input characteristic (Figure 9). After that, it descends the tree to the following node and bases its judgment on a different characteristic. The projected class or category is represented by a leaf node, which is reached by continuing this procedure. For classification problems in machine learning, decision tree classifiers are an all-around effective and flexible tool. They may offer important insights into complicated datasets and are applicable to a wide range of tasks, such as forecasting consumer behavior and identifying medical disorders. The Figure 10 shows the accuracy of an ANN model using a DT classifier.

```

DT
[59] from sklearn.tree import DecisionTreeClassifier
      from sklearn.metrics import accuracy_score, confusion_matrix, precision_score
      dt=DecisionTreeClassifier(criterion='gini', min_samples_split= 10, splitter='best')
      dt.fit(X_train,Y_train)

DecisionTreeClassifier
DecisionTreeClassifier(min_samples_split=10)

[60] prediction=dt.predict(X_test)
      accuracy_dt=accuracy_score(Y_test,prediction)*100
      accuracy_dt

87.5
    
```

Figure 9. Decision tree predicted value for decision and potential outcomes [27].

```

KNN
[ ] from sklearn.neighbors import KNeighborsClassifier

[ ] knn=KNeighborsClassifier(metric='manhattan', n_neighbors=22)
      knn.fit(X_train,Y_train)

KNeighborsClassifier
KNeighborsClassifier(metric='manhattan', n_neighbors=22)

[ ] prediction_knn=knn.predict(X_test)
      accuracy_knn=accuracy_score(Y_test,prediction_knn)*100
      print('accuracy_score score : ',accuracy_score(Y_test,prediction_knn)*100,'%')

accuracy_score score : 92.5 %
    
```

Figure 10. Accuracy of predicted data distribution using KNN model [22].

KNN classifier: It is a non-parametric lazy learning method, which means it does not assume anything about how the data are distributed and does not need a training phase. In KNN classification, a new data point’s class is predicted using the training data’s k-nearest neighbor’s classes. A user-defined hyperparameter called k controls how many neighbors are considered. The algorithm determines the distances between each new data point and every other data point in the training set to categorize it. Then, it chooses the k-nearest data points and determines the new data point’s class based on the dominant class of the chosen neighbor. The following figure shows the accuracy of the ANN model using the KNN classifier. But in the case of the LSTM model, we have found the accuracy of the model using the DT classifier as shown in Figure 11. The accuracy of the model using DT comes out to be 95%, which is more than the ANN model.

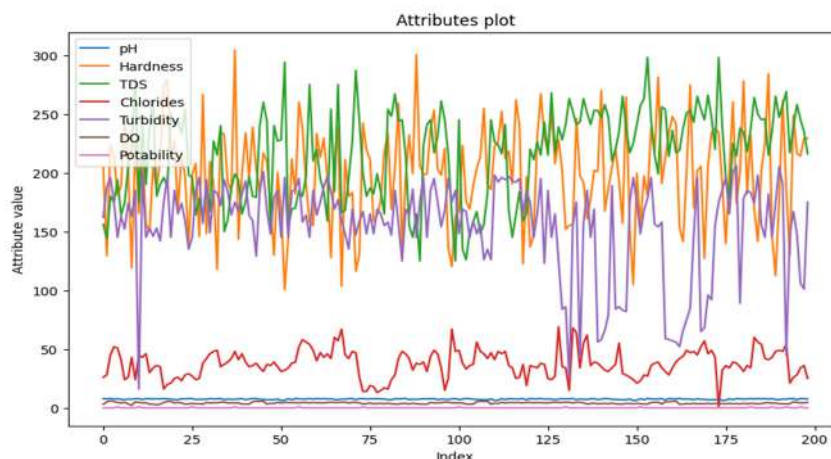


Figure 11. Prediction of future water quality parameters using LSTM Model [14].

Mean square error and Root mean square error.

Based on previous measurements, we provide predictions for future water quality levels in this analysis. The LSTM model would thus be a solid option for this investigation. If the research requires figuring out detailed relationships between several water quality indicators, ANN could be a better option. Because the MSE of the LSTM model is less than 1, it can be assumed that model predictions are, on average, relatively close to actual values. A smaller mean squared error (MSE) indicates that the model performs better in predicting the output values as given in Table 2. This metric quantifies the average squared difference between the expected values and the actual values.

Table 2. MSE and RMSE values for ANN and LSTM model.

MODELS	MSE	RMSE
ANN model	0.52	0.60
LSTM model	0.04	0.21

However, the MSE value in the ANN model is also less than 1, slightly higher than the LSTM model value. The second element is determined by the model's accuracy rating. Using KNN and DT classifiers, the ANN model's accuracy score is calculated to be 87.5% and 92.5%, respectively. However, the LSTM model's accuracy is 95%, which is higher than the ANN model. This demonstrates that, for limited datasets, the LSTM model outperforms the ANN model in terms of predicting water quality analysis.

6. Conclusions

In recent years, the use of artificial intelligence (AI) models in monitoring and evaluating water quality has become increasingly popular. This is because AI models are capable of processing large amounts of data quickly and accurately, allowing for the identification of trends and patterns in water quality data that may be difficult to detect using traditional methods. In this study, several research questions were raised regarding the use of AI models in surface water quality monitoring and evaluation, and the findings shed light on the most commonly used models, input parameters, and output measures, which are as follows:

- One of the major findings of the study was that long short-term memory (LSTM) and artificial neural networks (ANN) were the most commonly used AI models for water quality monitoring and evaluation in the past decade.
- The study also found that Iran and Southeast Asia account for most of the research on neural networks for surface water quality monitoring and evaluation. This suggests that these regions may be particularly interested in using AI models to improve water quality monitoring and evaluation.
- Another important finding of the study was that the most accurate models for predicting surface water quality were LSTM models for small datasets. This suggests that LSTM models may be particularly useful for analyzing small datasets, such as those that may be collected in rural or remote areas where water quality monitoring resources may be limited. Interestingly, the study found that there was no clear relationship between the size of the dataset and the R^2 value at the testing stage. This suggests that even small datasets can be used to train accurate AI models for water quality monitoring and evaluation.

Overall, the findings of this study suggest that AI models, particularly LSTM and ANN models, are a promising tool for improving surface water quality monitoring and evaluation. By analyzing large amounts of data quickly and accurately, these models can help identify trends and patterns in water quality data that may be difficult to detect using traditional methods. However, further research is needed to determine the most effective ways to implement these models in real-world water quality monitoring and evaluation programs. It was depicted from the heat map generation of the study that the color code

reference for water quality parameters falls in the range of values for this color code of -0.2 to 1.0 , with darker colors indicating negative effects on the corresponding parameter. The study models gave the correlation between pH and density, indicating the distribution of variables. It was obtained from the study that the mean square error and root mean square error of ANN and LSTM lie between 0.52 – 6.0 and 0.04 – 0.21 , respectively. This indicates the model performs better in predicting the output values. The study also indicated that, using KNN and DT classifiers, the ANN model's accuracy score is calculated to be 87.5% and 92.5% , respectively. However, the LSTM model's accuracy is 95% , which is higher than the ANN model.

It is important to note that there are still several issues that need to be addressed to improve the accuracy and applicability of these models. These issues could serve as a platform for future research in this area. One of the main issues that needs to be addressed is the need for a wider variety of neural network topologies to be examined in surface water quality prediction studies. Future studies could explore the use of convolutional neural networks (CNNs), recurrent neural networks (RNNs), and deep belief networks (DBNs), among others. Another important issue that needs to be addressed is the lack of research on neural network models in certain regions.

While the study found that Iran and Southeast Asia have been the most active regions in terms of research on neural networks in surface water quality monitoring and evaluation, there are still many regions where research in this area is lacking. It is imperative for American researchers to take up the challenge and take advantage of the numerous prospects for using neural networks in WQA. With the potential for new neural network topologies and the continued development of ensemble models, the accuracy of water quality prediction could be pushed even higher.

Author Contributions: R.R.: drafting—data collection and preparation of the manuscript, writing—review and editing; A.K.: drafting—preparation of the manuscript, revision, and correction; A.B.: composing—reviewing and modifying; F.M.A.: composing—reviewing and modifying; R.S.A.: reviewing; P.B.: composing—reviewing and modifying; S.A.: composing—reviewing and modifying; M.A.K.: composing—reviewing and modifying, O.Q.: composing—reviewing and modifying. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Researchers Supporting Project Number RSP2023R310, King Saud University, Riyadh, Saudi Arabia.

Data Availability Statement: Data are contained within the article.

Acknowledgments: The authors would like to acknowledge the support provided by the Researchers Supporting Project Number RSP2023R310, King Saud University, Riyadh, Saudi Arabia.

Conflicts of Interest: Authors have no conflict of interest.

References

1. Babu, C.N.; Reddy, B.E. A moving-average filter-based hybrid ARIMA-ANN model for forecasting time series data. *Appl. Soft Comput.* **2014**, *23*, 27–38. [[CrossRef](#)]
2. Lai, Y.C.; Yang, C.P.; Hsieh, C.Y.; Wu, C.Y.; Kao, C.M. Evaluation of non-point source pollution and river water quality using a multimedia two-model system. *J. Hydrol.* **2011**, *409*, 583–595. [[CrossRef](#)]
3. Zeilhofer, P.; Zeilhofer, L.V.A.C.; Haridoim, E.L.; Lima, Z.M.; Oliveira, C.S. GIS applications for mapping and spatial modeling of urban-use water quality: A case study in District of Cuiabá, Mato Grosso, Brazil. *Cad. Saúde Pública* **2007**, *23*, 875–884. [[CrossRef](#)] [[PubMed](#)]
4. Farrell-Poe, K.; Payne, W.; Emanuel, R. *Water Quality & Monitoring*; University of Arizona Repository: Tucson, AZ, USA, 2000; Available online: <http://hdl.handle.net/10150/146901> (accessed on 22 July 2018).
5. Pinto, M.M.S.C.; Marinho-Reis, A.P.; Almeida, A.; Ordens, C.M.; Silva, M.M.V.G.; Freitas, S.; Simões, M.R.; Moreira, P.I.; Diniz, P.A.; Diniz, M.L.; et al. Human predisposition to cognitive impairment and its relation with environmental exposure to potentially toxic elements. *Environ. Geochem. Health* **2018**, *40*, 1767–1784. [[CrossRef](#)] [[PubMed](#)]
6. Ahmed, A.A.M.; Shah, S.M.A. Application of adaptive neuro-fuzzy inference system (ANFIS) to estimate the biochemical oxygen demand (BOD) of Surma River. *J. King Saud Univ. Eng. Sci.* **2017**, *29*, 237–243. [[CrossRef](#)]
7. Aldhyani, T.H.H.; Alrasheedi, M.; Alqarni, A.A.; Alzahrani, M.Y.; Bamhdi, A.M. Intelligent hybrid model to enhance time series models for predicting network traffic. *IEEE Access* **2020**, *8*, 130431–130451. [[CrossRef](#)]

8. Yan, J.; Xu, Z.; Yu, Y.; Xu, H.; Gao, K. Application of a hybrid optimized BP network model to estimate water quality parameters of Beihai Lake in Beijing. *Appl. Sci.* **2019**, *9*, 1863. [[CrossRef](#)]
9. Ranković, V.; Radulović, J.; Radojević, I.; Ostojić, A.; Čomić, L. Neural network modeling of dissolved oxygen in the Gruža reservoir, Serbia. *Ecol. Model.* **2010**, *221*, 1239–1244. [[CrossRef](#)]
10. Zhang, X.; Hu, N.; Cheng, Z.; Zhong, H. Vibration data recovery based on compressed sensing. *Acta Phys.* **2014**, *63*, 119–128.
11. Solanki, A.; Agrawal, H.; Khare, K. Predictive analysis of water quality parameters using deep learning. *Int. J. Comput. Appl.* **2015**, *125*, 29–34. [[CrossRef](#)]
12. Kangabam, R.D.; Bhoominathan, S.D.; Kanagaraj, S.; Govindaraju, M. Development of a water quality index (WQI) for the Loktak Lake in India. *Appl. Water Sci.* **2017**, *7*, 2907–2918. [[CrossRef](#)]
13. Ahmad, Z.; Rahim, N.A.; Bahadori, A.; Zhang, J. Improving water quality index prediction in Perak River basin Malaysia through a combination of multiple neural networks. *Int. J. River Basin Manag.* **2016**, *15*, 79–87. [[CrossRef](#)]
14. Abyaneh, H.Z. Evaluation of multivariate linear regression and artificial neural networks in prediction of water quality parameters. *J. Environ. Health Sci. Eng.* **2014**, *12*, 40. [[CrossRef](#)]
15. Verma, D.; Berwal, P.; Khan, M.A.; Alharbi, R.S.; Alfaisal, F.M.; Rathnayake, U. Design for the Prediction of Peak Outflow of Embankment Breaching due to Overtopping by Regression Technique and Modelling. *Water* **2023**, *15*, 1224. [[CrossRef](#)]
16. Yesilnacar, M.I.; Sahinkaya, E.; Naz, M.; Ozkaya, B. Neural network prediction of nitrate in groundwater of Harran Plain, Turkey. *Environ. Earth Sci.* **2008**, *56*, 19–25. [[CrossRef](#)]
17. Min, C. An improved recurrent support vector regression algorithm for water quality prediction. *J. Comput. Inf.* **2011**, *12*, 4455–4462.
18. Jaloree, S.; Rajput, A.; Sanjeev, G. Decision tree approach to build a model for water quality. *Bin. J. Data Min. Netw.* **2014**, *4*, 25–28.
19. Yan, L.; Qian, M. AP-LSSVM modeling for water quality prediction. In Proceedings of the 31st Chinese Control Conference, Hefei, China, 25–27 July 2012; pp. 6928–6932.
20. Gazzaz, N.M.; Yusoff, M.K.; Aris, A.Z.; Juahir, H.; Ramli, M.F. Artificial neural network modeling of the water quality index for Kinta River (Malaysia) using water quality variables as predictors. *Mar. Pollut. Bull.* **2012**, *64*, 2409–2420. [[CrossRef](#)] [[PubMed](#)]
21. Sakizadeh, M. Artificial intelligence for the prediction of water quality index in groundwater systems. *Model. Earth Syst. Environ.* **2016**, *2*, 8. [[CrossRef](#)]
22. Tyagi, S.; Sharma, B.; Singh, P.; Dobhal, R. Water quality assessment in terms of water quality index. *Am. J. Water Resour.* **2013**, *1*, 34–38. [[CrossRef](#)]
23. Tang, G.; Li, J.; Zhu, Z.; Li, Z.; Nerry, F. Two-dimensional water environment numerical simulation research based on EFDC in Mudan River, Northeast China. In Proceedings of the 2015 IEEE European Modelling Symposium (EMS), Madrid, Spain, 6–8 October 2015; pp. 238–243.
24. Liao, H.; Sun, W. Forecasting and evaluating water quality of Chao Lake based on an improved decision tree method. *Procedia Environ. Sci.* **2010**, *2*, 970–979. [[CrossRef](#)]
25. Shafi, U.; Mumtaz, R.; Anwar, H.; Qamar, A.M.; Khurshid, H. Surface water pollution detection using internet of things. In Proceedings of the 2018 15th International Conference on Smart Cities: Improving Quality of Life Using ICT & IoT (HONET-ICT), Islamabad, Pakistan, 8–10 October 2018; pp. 92–96.
26. Al-Othman, A.A. Evaluation of the suitability of surface water from Riyadh Mainstream Saudi Arabia for a variety of uses. *Arab. J. Chem.* **2019**, *12*, 2104–2110. [[CrossRef](#)]
27. Kahlowan, M.A.; Tahir, M.A.; Rasheed, H. *National Water Quality Monitoring Programme; Fifth Monitoring Report (2005–2006)*; Pakistan Council of Research in Water Resources Islamabad: Islamabad, Pakistan, 2007.
28. Lee, S.; Lee, D. Improved prediction of harmful algal blooms in four major South Korea's rivers using deep learning models. *Int. J. Environ. Res. Public Health* **2018**, *15*, 1322. [[CrossRef](#)] [[PubMed](#)]
29. Fadlullah, Z.M.; Tang, F.; Mao, B.; Liu, J.; Kato, N. On intelligent traffic control for large-scale heterogeneous networks: A value matrix-based deep learning approach. *IEEE Commun. Lett.* **2018**, *22*, 2479–2482. [[CrossRef](#)]
30. Srivastava, G.; Kumar, P. Water quality index with missing parameters. *Int. J. Res. Eng. Technol.* **2013**, *2*, 609–614.
31. Bouamar, M.; Ladjal, M. A comparative study of RBF neural network and SVM classification techniques performed on real data for drinking water quality. In Proceedings of the 2008 5th International Multi-Conference on Systems, Signals and Devices, Amman, Jordan, 20–22 July 2008; pp. 1–5.
32. Hu, L.; Zhang, C.; Hu, C.; Jiang, G. Use of grey system for assessment of drinking water quality: A case study of Jiaozuo city, China. In Proceedings of the 2009 IEEE International Conference on Grey Systems and Intelligent Services (GSIS 2009), Nanjing, China, 10–12 November 2009; pp. 803–808.
33. Verma, I.; Berwal, P.; Setia, S.; Goel, R. Analysis on the Behaviour of Stiffened and Unstiffened Steel Plate Shear Walls with Enhanced Performance. In *IOP Conference Series: Materials Science and Engineering, Proceedings of the International Symposium on Fusion of Science and Technology (ISFT 2020), New Delhi, India, 6–10 January 2020*; IOP Publishing: Bristol, UK, 2020; Volume 804, p. 012035. [[CrossRef](#)]
34. Hayes, D.F.; Labadie, J.W.; Sanders, T.G.; Brown, J.K. Enhancing water quality in hydropower system operations. *Water Resour. Res.* **1998**, *34*, 471–483. [[CrossRef](#)]
35. Marir, N.; Wang, H.; Feng, G.; Li, B.; Jia, M. Distributed abnormal behavior detection approach based on deep belief network and ensemble SVM using spark. *IEEE Access* **2018**, *6*, 59657–59671. [[CrossRef](#)]

36. Batur, E.; Maktav, D. Assessment of surface water quality by using satellite images fusion based on PCA method in the Lake Gala, Turkey. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 2983–2989. [[CrossRef](#)]
37. Huang, J.; Liu, N.; Wang, M.; Yan, K. Application WASP model on validation of reservoir-drinking water source protection areas delineation. In Proceedings of the 2010 3rd International Conference on Biomedical Engineering and Informatics, Yantai, China, 16–18 October 2010; pp. 3031–3035.
38. Maiti, S.; Tiwari, R.K. A comparative study of artificial neural networks, Bayesian neural networks and adaptive neuro-fuzzy inference system in groundwater level prediction. *Environ. Earth Sci.* **2014**, *71*, 3147–3160. [[CrossRef](#)]
39. Taskaya-Temizel, T.; Casey, M.C. A comparative study of autoregressive neural network hybrids. *Neural Netw.* **2005**, *18*, 781–789. [[CrossRef](#)]
40. Khan, Y.; See, C.S. Predicting and analyzing water quality using Machine Learning: A comprehensive model. In Proceedings of the 2016 IEEE Long Island Systems, Applications and Technology Conference (LISAT), Farmingdale, NY, USA, 29 April 2016; pp. 1–6.
41. Li, X.; Song, J. A new ANN-Markov chain methodology for water quality prediction. In Proceedings of the 2015 International Joint Conference on Neural Networks (IJCNN), Killarney, Ireland, 12–16 July 2015; pp. 1–6.
42. Pinto, M.M.S.C.; Ordens, C.M.; de Melo, M.T.C.; Inácio, M.; Almeida, A.; Pinto, E.; da Silva, E.A.F. An inter-disciplinary approach to evaluate human health risks due to long-term exposure to contaminated groundwater near a chemical complex. *Expo. Health* **2020**, *12*, 199–214. [[CrossRef](#)]
43. Warren, I.R.; Bach, H.K. MIKE 21: A modelling system for estuaries, coastal waters and seas. *Environ. Softw.* **1992**, *7*, 229–240. [[CrossRef](#)]
44. Najah, A.; Teo, F.Y.; Chow, M.F.; Huang, Y.F.; Latif, S.D.; Abdullah, S.; Ismail, M.; El-Shafie, A. Surface water quality status and prediction during movement control operation order under COVID-19 pandemic: Case studies in Malaysia. *Int. J. Environ. Sci. Technol.* **2021**, *18*, 1009–1018. [[CrossRef](#)] [[PubMed](#)]
45. Latif, S.D.; Azmi, M.S.B.N.; Ahmed, A.N.; Fai, C.M.; El-Shafie, A. Application of Artificial Neural Network for Forecasting Nitrate Concentration as a Water Quality Parameter: A Case Study of Feitsui Reservoir, Taiwan. *Int. J. Des. Nat. Ecodyn.* **2020**, *15*, 647–652. [[CrossRef](#)]
46. Latif, S.D.; Azmi, A.H.; Ahmed, A.N.; Hatem, D.M.; Al-Ansari, N.; Fai, C.M.; El-Shafie, A. Development of prediction model for phosphate in reservoir water system based machine learning algorithms. *Ain Shams Eng. J.* **2022**, *13*, 101523. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.