ARTIFICIAL INTELLIGENCE FOR SURFACE WATER QUALITY MONITORING

MAJOR PROJECT REPORT

Submitted in partial fulfillment of the requirements for the award of the degree of

BACHELOR OF TECHNOLOGY

IN
CIVIL ENGINEERING
Under the supervision of

Dr. Rishi Rana
Assistant Professor (SG)

By

Ujjwal Sharma [191613]
Manan Sharma [191615]
to

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY WAKNAGHAT SOLAN – 173234

HIMACHAL PRADESH INDIA

MAY, 2023

TABLE OF CONTENTS

STUDENT'S DECLARATION

I hereby declare that the work presented in the Project report entitled **""Artificial Intelligence for Surface Water Quality Monitoring""** submitted for partial fulfillment of the requirements for the degree of Bachelor of Technology in Civil Engineering at **Jaypee University of Information Technology, Waknaghat** is an authentic record of our work carried out under the supervision of **Dr. Rishi Rana**. This work has not been submitted elsewhere for the reward of any other degree/diploma. We are fully responsible for the contents of our project report.

Signature of Student	Signature of Student
Ujjwal Sharma (191613)	Manan Sharma (191615)
Department of Civil Engineering	Department of Civil Engineering
JUIT	JUIT
Waknaghat, India	Waknaghat, India

CERTIFICATE

This is to certify that the work which is being presented in the project report titled **"Artificial Intelligence For Surface Water Quality Monitoring"** in partial fulfillment of the requirements for the award of the degree of Bachelor of Technology in Civil Engineering submitted to the Department of Civil Engineering, **Jaypee University of Information Technology, Waknaghat** is an authentic record of work carried out by **Ujjwal Sharma (191613)** and **Manan Sharma(191615)** under the supervision of **Dr. Rishi Rana and** Department of Civil Engineering, Jaypee University of Information Technology, Waknaghat. The above statement made is correct to the best of our knowledge.

Date: /05/2023

Signature of Supervisor

Dr. Rishi Rana

Assistant Professor (SG)

Department of Civil Engineering

JUIT, Waknaghat

ACKNOWLEDGEMENT

The success and final outcome of this project required a lot of guidance and assistance from many people and we are extremely privileged to have got this all along the completion of our project. All that we have done is only due to such supervision and assistance and we would not forget to thank them.

We would like to express our sincere gratitude to Dr. Rishi Rana (Assistant Professor) for their valuable guidance and kind supervision throughout the project.

# Abstract

This study aims to explore the potential of machine learning algorithms, specifically artificial neural networks (ANNs) and long short-term memory (LSTM) models, for surface water quality monitoring. The study utilizes a dataset with seven critical parameters, and the created models are evaluated based on various metrics. The goal is to categories and properly forecast the water quality index (WQI) using the suggested models. The findings show that the suggested models can accurately assess water quality and forecast WQI with high rates of success. Both the ANN and the LSTM models performed well in predicting WQI, with the ANN and LSTM model achieving an MSE and RMSE value. Temperature, pH, dissolved oxygen, conductivity, total dissolved solids (TDS), turbidity, chlorides 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 ANN and 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 displays their correlation coefficient. The study's findings show how effective machine learning algorithms may be as a tool for monitoring surface water quality. Real-time analysis and forecasting capabilities offered by these models can aid in spotting possible problems with water quality and aid in decision-making. The study highlights the importance of leveraging machine learning techniques in water quality monitoring to ensure the protection and management of water resources. With the advancements in machine learning, artificial intelligence (AI) techniques have emerged as a promising tool for surface water quality monitoring. This study aims 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.

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS & ABBREVIATIONS

AI	Artificial Intelligence
ANFIS	Adaptive Neuro-Fuzzy Inference System
ANN	Artificial Neural Network
BOD	Biochemical Oxygen Demand
BPNN	Back Propagation Neural Network
DO	Dissolved Oxygen
DT	Decision Tree
IoT	Internet of Things
KNN	K-Nearest Neighbors
LSTM	Long Short-Term Memory
ML	Machine Learning
MLP	Multi-Layer Perception
MSE	Mean Square Error
PCA	Principal Component Scrutiny
RMSE	Root Mean Square Error
SVM	Support Vector Machine
TDS	Total Dissolved Solids
WQ	Water Quality
WQI	Water Quality Index
WSN	Wireless Sensor Network

CHAPTER 1

INTRODUCTION

Inspite of the fact that ¾ of the earth surface is secured by water, as it were almost 1% is fresh water. Expending world populace and industrialization is additionally leading to the expansion of toxins in water bodies. In this manner it is essential toceaselessly screen the quantity of water from characteristic sources in common and surface water in specific. Quality of the water observing and appraisal are attainable by the affirmation of particular framework. Water quality can to be checked by organic record or physiochemical parameters. Advancing from the conventional water quality checking and evaluation strategies, Web of Things- enable technologies, and utilize of artificial insights are unused domains being investigated. Artificial intelligence was presented in computer science within 1950s and has undergone substantial changes in enhancement and modernity. AI has made a difference analyst accomplish the plausibility of imitating human behaviour abilities in specific spaces of knowledge. AI instruments include sloppy logic, particle swarm optimization, algorithmic genetics, neural networks artificial, assistance vector machine, adaboost algorithm etc. In this consider a few pertinent questions with respect to the using of neural networks to improve water quality observing and evaluation was interrogated. This is done to assist pickup insight into the reasons for a few patterns within the Enquire about area. The focus of thus consider is the past decade. These kinds of studies are important because they help analysts shape their thoughts. Within the space and direct them to more unconventional arrangements. The significance of the study is justified considering that the accessibility of clean water could be the basic economic advancement objective and neural system in water quality checking and evaluation is a generally youthful investigate region. The scientific commitment of this project is to show a survey on the application of different sorts of neural network in surface quality of water checking centring on the strategies utilized the area of the test the input parameter utilized and the yield meter.

The quality of most circulatory water bodies, such as rivers, lakes, and streams, is determined by precise quality standards. Furthermore, there are water criteria for various purposes and uses. Irrigation water, for example, must not be overly saline or include hazardous elements that could be passed to plants or soil, harming ecosystems. Different qualities are required for industrial water quality depending on the various industrial activities. Natural water resources, such as ground and surface water, are some of the cheapest sources of fresh water. Human/industrial activity, as well as other natural processes, can pollute such resources.

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 lack of public awareness and less hygienic elements. Indeed, the effects of adulterated drinking water are quite unhealthy, posing a serious threat to human health, the climate and framework. According to a UN statistic, 1.5 million people die each year as a result of diseases caused by contaminated water. Water contamination is said to be the cause of 80% of health problems in impoverished countries. Annually, there are five million deaths and 2.5 billion illnesses reported. This is a greater decency rate than deaths caused by accidents, crimes, or terrorist attacks.

Massive population growth, the industrial innovation, and the usage of manure and fungicides have all had a negative impact on WQ ecosystems. Having models for predicting the WQ is therefore quite useful for conclusion of water infection.

There are currently two types of models for modeling and assessment of quality of water-mechanism-oriented and non-mechanism-oriented models. The mechanism model is quite sophisticated; it simulates the WQ using headway system infrastructure data, and so it is a multifunctional model that can be applied to any water body. In addition, one of the early WQ simulation models, the Streeter–Phelos (S–P) model, has been frequently employed.

Batur and Maktav conducted another investigation utilizing satellite image fusion and the principal component scrutiny (PCA) approach to forecast the WQ of Lake Galaa (Turkey). Using a decision tree approach, Jaloree et al sought to forecast the WQ of the Narmaada River using five WQ markers. Another study recommended using the deep fractional Stacked Simple Recurrent Unit (Bi-S-SRU) to create a precise WQ forecasting system in smart mariculture.

1.1 SURFACE WATER RESOURCES

A crucial 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 by this term. This kind of water is distinct from subterranean water, which is found below the surface of the earth. 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 during different times of the year, and both human activity and natural processes can have an impact on its quality. Human consumption is one of the main uses of surface water, especially in places where groundwater supplies are few or difficult to reach. Natural preservation organisations 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. The irrigation of crops, especially in agriculture, is a substantial additional use of surface water. Additionally, surface water is utilised 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 droughts and floods. 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 little surface water supplies, groundwater may be a substantial supply of water for human use. However, excessive groundwater pumping can result in pollution and depletion, which can have an impact on the environment and human health. A variety of human activities, such as agriculture, industrial operations, and home wastewater, can also have an impact on the quality of surface water. 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 fertilisers and animal manure can also contribute to this problem. Several measures, including as 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. It is crucial to put into place efficient policies and practises in order to preserve and safeguard surface water resources for both current and future generations.

1.2 NECESSITY OF SURFACE WATER

Surface water, including lakes, ponds, rivers, and streams, plays a crucial role in maintaining the balance of the ecosystem and supporting human life. Despite making up a relatively small portion of the total water on the earth, these bodies of water perform numerous vital tasks for both the natural world and human culture. The ability of surface waters to provide as a habitat for a wide variety of plant and animal species is one of their most important functions. These ecosystems frequently sustain a complex web of life that includes many different kinds of fish, birds, insects, amphibians, and other organisms. These creatures' survival and health are directly impacted by the quality of surface water, which can then have an influence on how well the ecosystem as a whole function. Surface waters not only support biodiversity but also have many uses for people. We depend on these bodies of water for a variety of purposes, including irrigation, industry, and drinking water. Hydropower generation, a crucial source of renewable energy, depends on surface waterways. Many of these applications would not be possible or would be far more difficult to do without surface water. However, different types of pollution may quickly affect the quality of surface water. Surface water quality may be dramatically lowered by the buildup of pollutants, chemicals, oil, rubbish, and other waste on exterior surfaces. For instance, the storm water runoff will be increasingly deteriorated when the usage of chemical compounds surrounding houses and businesses grows.

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 get 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 and damages aquatic life and causes siltation and turbidity. 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 endanger human health, can enter streams by spills, runoff, or direct discharges.

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. Pollutants can often be avoided by using source control measures, such as covering storage places and reducing industrial discharges.

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. Reduce pollution at the source and use various treatment technologies to filter pollutants and enhance water quality in order to safeguard and improve the quality of surface water. We can ensure that surface waters continue to deliver important advantages for future generations by adopting these steps.

1.3 PRESENT SCENARIO OF SURFACE WATER IN INDIA

Although our earth is known as the 'Blue Planet' as 71% of the earth is protected with the aid of using water, but handiest 2.6% of the world's water is fresh, at the same time as 97% is saline being oceans. Of this small percent of freshwater, handiest 0.4% of this freshwater is to be had from rivers, lakes and reservoirs, 31% from the groundwater, at the same time as the relaxation is saved in remote glaciers, mountainous areas, locations that we will in frequently access. This is the current state of global water distribution, with no action expected. As we're conscious that severe interventions are taking vicinity each day with inside the general environment, and for that reason real freshwater availability with inside the shape of floor water and groundwater assets are a great deal distinctive to what it seems from the worldwide water distribution. The call for of water from diverse water users namely, domestic, municipal, agricultural, horticultural, electricity and commercial sectors are increasing, and this has placed first rate strain at the water assets systems. The developing water pollutants issues for water in addition to groundwater have annoyed the difficulty with water supply. The ideological uncertainty, the improvement sports, array in land, soil, climate, terrain and alongside water scarcity has been exacerbated by temporal unpredictability and socioeconomic aspects. Actual styles complexities, lowering freshwater opportunity and stability of water assets have raised many water troubles at local, nearby and worldwide scales. The moment has come to take a look back at how water is used and misused.

1.4 WATER AVAILABILITY IN INDIAN SCENARIO

The geographical location of India is set 330 million hectares (2.50 % of the earth's land mass) and its populace is 1,030 million primarily based totally at the 2001 census, which is set 15% of that of the world. The renewable sparkling water sources of the u. s. are 1,868 km^3 yr-1, that are handiest approximately 5% of these of the world. As in lots of different countries, water sources of this u. s. aren't frivolously dispensed in area and time. Despite the fact that some water is obtained from upstream nations, precipitation is the principal source WRS of India 17% of water availability. The annual rainfall varies from greater than 9999 mm in components of Meghalaya within side the north-east to much less than 600 mm in semi-arid area of the country. In arid areas, it's miles even much less than a hundred mm. Much of the water is obtained in the course of monsoon season, approximately4 months in and greater vital is that it happens in approximately a hundred hours of the wet days. The availability of in line with capita utilizable water primarily based totally on 1991 populace varies as 183 m3 in Sabarmati, 2,400 m3 in Mahanadi and 3,087 m3 in Narmada basin. The utilizable water to be had in phrases of in line with hectare of cultivable location is 1,245 m3 in Sabarmati, 8,420 m3 in Mahanadi and 7,670 m3 in Narmada basin (MOWR, 2002). These differences are enormous and were the key of countrywide water sources planning. India is largely an agricultural u. s. and more percentage of populace lives in rural regions. Periodically, a few areas of the u. s. face a precarious inadequacy of water sources while others get hold of components in extra of the real requirement of the populace they support. Flood & droughts are the 2contraryfacttouchingone-of-a-kindcomponents of the u.s. simultaneously. With such disparities in water source distribution, the suitable water control with the nearby intention of poverty relief within side the rural regions attract paramount importance. The developing challenge on availability of freshwater sources is elevating many questions concerning now no longer handiest the financial improvement of the nation, however additionally at the socioeconomic improvement in addition to sustainability of lives of mankind and biodiversity. The relentless stress is mounting on water sources because of populace growth, fast unionization, large-scale mechanization, and environmental care, which, in turn, are posing many barriers and colossal obligations for planners, managers, engineers, scientists, and decision-makers of water sources almost everywhere, especially in arid and semi-arid areas and unique regions of acute groundwater great problems, are posing many barriers and colossal obligations for the planners, managers, engineers. The elevated stress is spilling over the groundwater sources as properly due to the hydrological anxiety, developing groundwater infection problems, immoderate and impulsive groundwater mining and shrinking freshwater sources. As the result, availability of in line with freshwater sources is getting worse by the day.

1.5 HISTORY OF AI

Artificial Intelligence (AI) is a field of study and research concerned with the development of computer systems that can perform tasks that typically require human intelligence, such as visual perception, speech recognition, decision-making, and natural language understanding. The concept of AI dates back to the 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. The early 20th century saw significant advancements in various fields such as mathematics, psychology, engineering, economics, and neurology. These advancements, coupled with the development of large code machines during World War II, led to the creation of topological works that described how machines could be designed to "think." Alan Turing's work on the "imitation game" in 1950 laid the foundation for the development of the first AI systems. The term "artificial intelligence" was coined in 1956 at the Dartmouth Summer Research Project on AI, where researchers from different fields came together to explore the possibilities of creating intelligent machines. Since then, the field of AI has gone through several waves of rapid development, followed by periods of cooling in funding and interest. These periods are often referred to as "AI winters." The first wave of AI development focused on the development of expert systems that could perform specific tasks using predefined rules and knowledge bases. In the 1980s, machine learning emerged as a new approach to AI that allowed computers to learn from data instead of being explicitly programmed. The second wave of AI development, often referred to as the "AI renaissance," began in the mid-2010s with the emergence of deep learning, a form of machine learning that uses neural networks with multiple layers to learn from vast amounts of data. AI is a broad and rapidly evolving field that encompasses a range of sub-disciplines, including natural language processing, computer vision, robotics, and cognitive computing. The technology has the potential to transform various industries, including healthcare, finance, manufacturing, and transportation. However, it also raises ethical, social, and economic questions, such as the impact on employment and privacy, the potential for bias and discrimination, and the ethical use of AI in decision-making. In conclusion, AI has a rich history that dates back to ancient times. The development of programmable computers in the 19th century, coupled with advancements in various fields, led to the creation of topological works that described how machines could be designed to "think." Since the term "artificial intelligence" was coined in 1956, the field has gone through several waves of rapid development, followed by periods of cooling. Today, AI is a rapidly evolving field with the potential to transform various industries, but it also raises important ethical, social, and economic questions that need to be addressed. The cause of this report, we consciousness on what the United States DARPA, one of the important public investors of AI studies with inside the beyond decade, have lately classified the 2nd one and 0.33 waves of AI (DARPA, 2016). In advance levels of the sphere that normally targeted on a high-stage “symbolic” presentation of troubles. Operation evolved early days had been suitable at interpretation however restricted capacity to study. The 2nd wave of Artificial Intelligence took off on a flip of the 21st century, with arithmetical fashions that had been educated with “massive data”. Those tactics permit for category and information talents however there is no contextual functionality & best minimum reasoning capacity. To make It greater concrete: 2nd wave algorithms are thoroughly capable of apprehend a cat from a photograph, however they're now no longer capable of provide an explanation for why it's far a cat. These tactics are ruled via way of means of statistical mastering, depending closely on device mastering, including, deep mastering and growing strategies. ML lets us in computerized development via experiences (Jordann and Mitchel, 2015). This specializes in the improvement of logical fashions that pc structures use to carry out a selected undertaking without the usage of express instructions, counting onstyles.ML data construct a mathematical version primarily based on the pattern data, recognized as “schooling data”. A unique magnificence of ML strategies entails deep mastering, which refer to data that use more than one layer to steadily withdraw high-stage functions; it makes them especially appropriate for mastering from unformed and unlabeled data (Dang and Yu, 2013). Growing computation circle of relatives for international optimization stimulated by way of means of organic growth. A unique case of growing computation worries genetic computation that is predicated on bio-stimulated operators consisting of mutation and choice. Mimicking organic evolution, a set of rules generates a preliminary answer to a hassle. This set of answers is iteratively updated, in which every technology is introduced via way of means of casting off answers that carry out much less on the selected health feature of the set of rules, and introducing changes to the organic methods of herbal choice. Resulting, the health of the populace, that is the fine of the answers to the to start with described hassle, will regularly growth. One of the factors of disapproval raised towards those 2nd wave AI strategies is that the workings of those strategies stay black-bins and that they are closely depending on fine of schooling facts. In reaction to this hassle of opacity, DARPA 0.33 wave of Artificial Intelligence strategies specializes in explain ability and the improvement of so-called “white container Artificial Intelligence” (2018). Those AI structures vary from 2ndone wave of capacity for adjustment. 3rdwave AI strategies apprehend meaning & capable of adjust accordingly. To use of the instance of the photograph of cat again, strategies within side the 0.33 waves will now no longer best be capable of apprehend the cat, however they'll additionally be capable of provide an explanation for why it's far cat & the way they have arrived at conclusion. It could a smaller number of troubles related to the black container-nature of modern-day ML strategies. Additionally, 0.33 wave AI strategies are much less depending on big units of schooling details. At time of writing the assessment in 2016, DARPA placed0.33 wave of Artificial Intelligence prospectively within side the 2020, so whether or not & the way this 0.33 wave will without a doubt question. DARPA 2nd & 0.33 wave of AI strategies are constant with definitions in current pc science, in which the capacity to study and reply to converting surroundings is regularly visible as key.

CHAPTER 2

LITERATURE REVIEW

Paul D. Robillard A search was undertaken for studies that used neural network models to investigate the quality of surface water for this investigation (both modeling and prediction). Only papers indexed in Web of Science (WoS) and/or SCOPUS were evaluated. The terms 'WQ ' and “AI” were used in the search. Only studies on surface water (rather than groundwater or precipitation) were considered in the analysis. In addition, the time period considered was the previous decade (2011–2020). The review's literature was compiled using these criteria. Table shows how the required information from the article was carefully gathered and summarized. This table served as the foundation for creating the appropriate plots to obtain answers to the queries.

Table 2.1 Summary of application of different types of neural networks in Water Quality Monitoring

YEAR	LOCATION	MODEL	SIZE OF INPUT DATA	INPUT PARAMETERS	HIGHEST R^2	LOWEST RMSC
2020	Surbaya River	ANN	12year (monthly)	BOD, COD, DO, pH, temperature	-	-
	Aiye river China	ANN	unclear	12	-	-
	Huaxi river China	MLP	2 years (monthly)	Temperature, pH, PI, EC, TP, NH3-N, TN and COD	0.496	1.165
2019	2 lakes university India	ANN	4 months (every 5 days)	TSS, BOD	-	-
	Yamuna river,India	ANFIS	14 years	DO	-	0.1515
	Yamuna river,India	BPNN	14years (monthly)	DO	-	0.0936
	Yamuna river,India	SVM	14 years(monthly)	DO	-	0.0948

2018	Gorganrood basin Iran	ANFIS	18 years (monthly)	EC	0.98	58.50mu/cm
	Gorganrood basin Iran	ANFIS	18 years	SAR	0.97	0.13
	Gorganrood basin Iran	ANFIS	18 years	TH	0.95	18.45 mg/liter
2017	Hooghly river India	NN-CS	-	Ph, chlorides, TH, TA, Turbidity	-	-
	Hooghly river India	NN-GA	-	Ph, chlorides, TH, TA, Turbidity	-	-
	Hooghly river India	NN-PSO	-	Ph, chlorides, TH, TA, Turbidity	-	-
	Saint john River Canada	BPNN	-	Turbidity	-	0.57 NTU
	Saint john River Canada	BPNN	-	TSS	-	0.66 mg/l
2016	AJi- Chay river Iran	ANN	28 years (monthly)	Salinity (EC)	0.9936	3.97 E-5 mus/cm
	AJi- Chay river Iran	ANFIS	28 years (monthly)	Salinity (EC)	0.9954	3.76 E-5 mus/cm
	AJi- Chay river Iran	W-ANN	28 years (monthly)	Salinity (EC)	0.9960	3.47 E-5 mus/cm

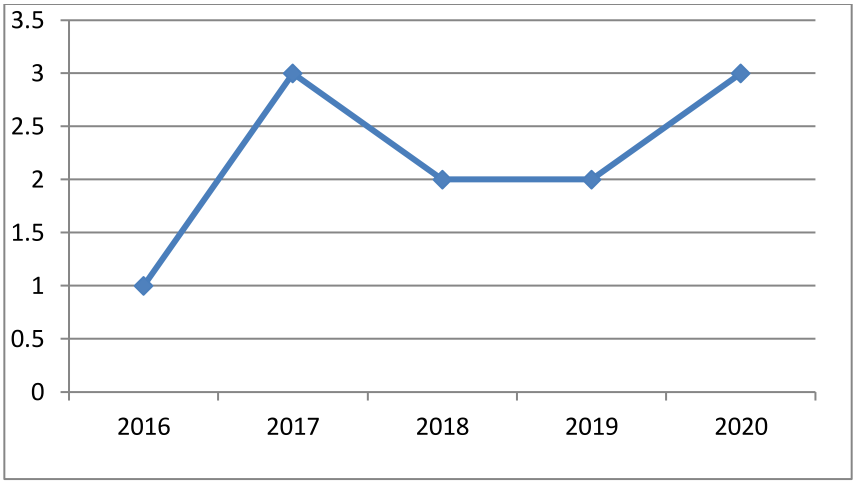


Figure 2.1 Number of studies for every year (Source: Joshua et al. 2021)

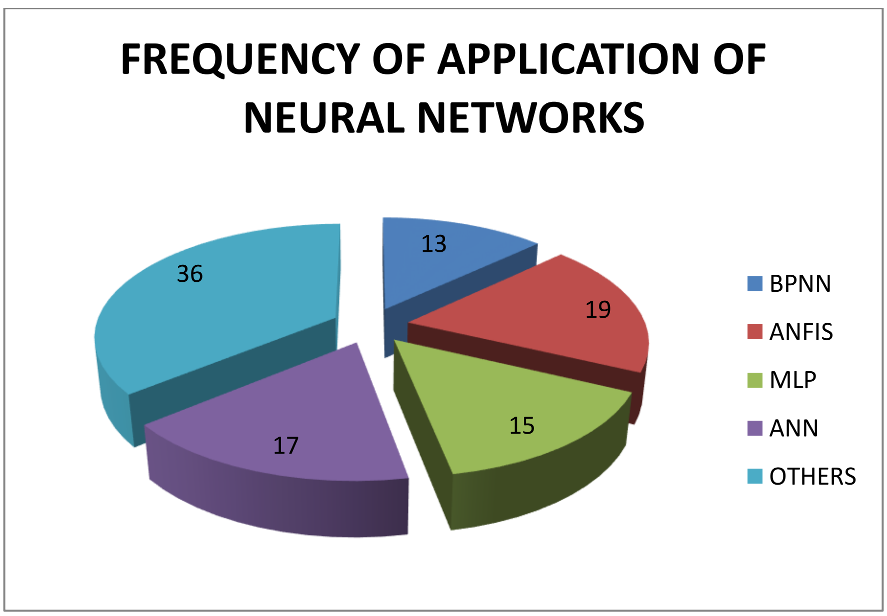


Figure 2.2 Frequency of application of Neural Networks **(**Source: Joshua et al... 2021)

Robert O. StrobI and Paul D. Robillard in AI Technologies in Surface Water Quality Monitoring concluded that using a combined set of physical-chemical and biological parameters did not yield good results.The IoT's rapid innovation in sensor, wireless communication, and retail IoT is becoming more and more expected to be the next-generation choice of control.

Karthick.T and Suresh Sankaranarayanan (2017) in Intelligent IOT Based Water Quality Monitoring System Calculated pH and TDS for different types of water – salty, mud, drinking, tap water and then showed the graph for pH v/s TDS for same. The Internet of Things (IoT) is critical for improving industrial efficiency and quality, as well as lowering industrial costs and resources. However, there have been few freely reported real IoT project uses so far.

Adewale George Adeniyi, Josha O. Ighalo (2020) in AI for Surface water quality monitoring and assessment showed the different models used for water quality monitoring and conducted a literature analysis for the past studies. They have discussed about different types of artificial models and the model the model that can be used to calculate water quality index. The BOD value can be predicted immediately since the predictor parameters of these models can be measured quickly.

Arini Dwi Astuti, Azmi Aris, Samila Azman (2020) in AI Approach to Predict River Water Quality predicted BOD, COD, EC, TDS and Turbidity. The prediction of groundwater level (GWL) using geoelectric properties is one of the trickiest puzzles to solve. It is partly because there isn't yet a concrete empirical connection between the amount of groundwater and the geoelectric parameters. This study looked into the ability of advanced artificial neural networks (ANNs) to model nonlinear systems in an effort to get around these problems.

Farmanullah Jan, Nasro Min-Allah and Dilek Dustegor (2021) gives us most recommend WQM parameters are Turbidity, Temperature, pH, Electrical conductivity, Oxidation reduction potential. To get solutions that are physically accurate, it is important to formulate the problem more precisely than has previously been done in the literature and to represent the underlying processes realistically. It successfully integrates data models, makes wise decisions, does dynamic optimization, and controls.

C.A. Biraghi, M. lotfian and D. Carrion, M.A. Brovelli (2021) predicted that Using CNN model detected algae and foam present in water. Model works well with the algae. Contaminants are eliminated by the procedure, which then turns them into effluents that may either be supplied to the water supply or immediately recovered.

Mohd Razman Salim (2021) It has been shown that a significant indication of water quality is the biological oxygen demand (BOD), which assesses the amount of biodegradable organic material in water. Long-term incubation of water samples is required for the determination of biological oxygen demand. It requires a lot of time and energy as a result. Since the predictor parameters of these models can be quickly measured, the BOD value can be predicted right away.

Mohammed Al-Yaari and Ali Khalaf Ahmed (2022) evaluated that there is total eight parameters in dataset. Further he says that minimum of four parameters must be calculated accurately. He categorizes the quality of water samples. It can eliminate impurities and transform them into effluents that may be recovered immediately after water treatment or supplied to the water supply.

Theyazn H.H Aldhyani and Mohammed Al-Yaari (2020) it compares different types of artificial models like ANFIS, ANN. It also indicates which models are more accurate at predicting the water quality categorization and index. Effective data-model integration, wise decision-making, dynamic optimisation, and control are all accomplished.

F. R. Islam, H. Haqva (2011) The Internet of Things (IoT) and Smart Grid play critical roles in promoting and guiding information technology and economic development. 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 study. 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.

Rafia Mumtaz, Umer Ahmed (2020), to get the most information out of 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 take into account the specific monitoring objectives, representative sampling size, location, and frequency, water quality variable selection, as well as logistical and financial limitations. A workable and simple to use technique for designing a water quality monitoring network will provide a reliable, effective, and affordable design.

A.N.Prasad, K. A. Mamun, F. R. Islam, H. Haqva demonstrates a smart water quality monitoring system. For a 12-hour period, hourly tests were conducted on four different water sources to confirm the system's measurement accuracy. Furthermore, given the explicit management and information requirements, it is crucial to identify an optimal design because water quality monitoring networks can be extremely expensive.

Josha O. Ighalo (2012) How to protect fresh water sources is one of the most important concerns facing human society, with industrial effluent being the main source of pollution. We found that classic WSN nods have the issue of short-range communication among nods, which restricts their application in monitoring industrial effluent, despite the development of several WSN-based water quality monitoring systems. In this paper, we provide a novel monitoring system that offers long-distance communication and real-time monitoring between nodes by using the ANN network and dynamic routing. This system has a broad overlay area, is generic, extensible, self-organizing, and adaptable, and has a greater compatibility to forecast various AI models.

Samila Azman (2019) the problem in this project was to create an ANN (Artificial Neural Network) and automation system. With excellent water quality and cleaning effectiveness, the automated station will operate continuously. With a 10-year guarantee against mechanical issues, this station requires less maintenance than other predictive models.

M.A. Brovelli (2019). Due to the complexity of managing and running wastewater treatment in rural areas. Internet of things is projected to be the control of choice for the next generation because to its quick improvement in sensor, wireless connection, and retail. A control system is present in a wastewater treatment facility. IoT technology is presently developing quickly.

Dilek Dustegor (2021) examined how the urgent need to address the issue of water shortage has alarmed the modern society. In order to satisfy the enormous demand for drinking water among the world's continuously expanding population, it is sensible to treat wastewater. The creative application of Internet of Things (IoT) technology has benefitted several smart cities. This study suggests a real-time, Internet of Things-based system that routinely monitors specific important parameters at a water treatment facility and notifies the operator if anything goes wrong.

M. lotfian (2014) IoT-based monitoring system were built to increase the efficiency of treatment systems and the immature operation management experience. The system is composed of an application layer, a network layer, and a sensor layer. Numerous studies now include such techniques in their research as a result of ML and computer vision advancements, particularly. In recent years, interest in the integration of ML approaches with several fields has increased.

Azmi Aris (2017) Water Quality monitoring is very much needed as it is consumed by residents. Traditional water Quality monitoring and some of the technology-based Water Quality got lot of challenges. In addition, there is no intelligence in existing water Quality Monitoring for analysis and prediction.

Hunter etal (2018). Multi-sensor systems can be used to detect anomalies in water in real-time. Although optical sensors were used in this study to develop an AI for a drinking water application, other sensors could be used to broaden the application area. While the exact combination of sensors may vary by application, the overarching technique remains the same. This method could be used in a variety of areas including surface-water, urban runoff, food and industrial process water, aquaculture, and numerous areas where water is used and reused.

Salmiati (2020) concluded that current models applied to water quality prediction are not user-friendly enough and mostly their implementation is subjected to substantial restrictions. Assessing water reclamation potential by neural network model. Engineering Applications of Artificial Intelligence numerical model is a challenging task for novice application users. The incorporation of current heuristic understanding of model manipulation and the intellectual manipulation of calibration parameters are therefore instrumental. The latest developments in AI technology provide a way of filling the gap between the designer and the model professional. This article examined the state-of-the-art models proposed for water quality prediction and the progress made in integrating AI into these models. Several plans for the future are investigated and submitted for further advancement and their potential. More progress in function modeling in this direction is expected to be promising with the ever-increasing potential of AI technologies

Azman (2019) the problem in this project was to create an ANN (Artificial Neural Network) and automation system. The automated station will run nonstop, with high water quality and cleaning

efficiency. This station comes with a 10 years warranty against mechanical problems, resulting in lower maintenance costs than different predicting models.

Azmi Aris (2020) in AI Approach to Predict River Water Quality predicted BOD, COD, EC, TDS and Turbidity. One of the most difficult problems to solve is the prediction of groundwater level (GWL) using geo electric characteristics. It is due in part to the fact that an empirical relationship between the level of groundwater and the geo electric parameters has not been established yet. In this study, an effort was made to circumvent these obstacles by investigating the capacity of advance artificial neural networks (ANNs) to simulate nonlinear systems.

A smart system for monitoring water quality is demonstrated by A.N. Prasad, K. A. Mamun, F. R. Islam, and H. Haqva. For a 12-hour period, hourly tests were conducted on four distinct water sources to confirm the system's measurement accuracy. Furthermore, given the specified management and information requirements, it is crucial to identify an appropriate design since water quality monitoring networks may be quite expensive.

F. R. Islam and H. Haqva (2014) The Internet of Things (IoT) and Smart Grid are crucial for advancing and directing the growth of information technology and the economy. 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 study. WiFi-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.

Azman (2019) Making an ANN (Artificial Neural Network) and automation system was the project's challenge. With excellent water quality and cleaning effectiveness, the automated station will operate continuously. With a 10-year guarantee against mechanical issues, this station requires less maintenance than other predictive models.

Hunter etal (2019) Real-time detection of water abnormalities is possible with the use of multi sensor systems. In spite of the fact that optical sensors were utilized in this study to create an AI for a drinking water application, additional sensors may be employed to expand the application. Although the precise set of sensors used may vary depending on the application, the general methodology is the same. Surface water, urban runoff, food and industrial process water, aquaculture, and many other sectors where water is utilized and reused might all benefit from the adoption of this technique.

In their 2017 study, Intelligent IOT Based Water Quality Monitoring System, Karthick.T and Suresh Sankaranarayanan calculated pH and TDS for several types of water, including salty, mud, drinking, and tap water, and then displayed the graph for pH v/s TDS for the same.

CHAPTER 3

PROJECT DESCRIPTION

Surface water is one of the major sources of drinking water in our nation wherein this cutting-edge time, where we devour Surface water & the water collected by Rain Collecting and utilizing it after sifting it, there are still the individuals who depend on the groundwater (water from WELL) basically agriculturists. So, to be able to create it a great source as well as a crisis choice for us presently, we ought to get genuine that how can it be spared by ceasing its wastage and keep up its quality & what best way instead of observing it utilizing our advanced innovations i.e., Sensors may it be.

The importance of surface water quality monitoring cannot be overstated, as it is essential to ensure the availability of safe and clean water for human consumption and other purposes. The traditional methods of water quality monitoring have several limitations, such as the requirement for manual sampling and laboratory analysis, which are time-consuming, expensive, and often cannot provide real-time data. Therefore, the use of artificial intelligence (AI) models in surface water quality monitoring has gained significant attention in recent years. This project aimed to investigate the application of various AI models in surface water quality monitoring.

The techniques used in the experiments were diverse, including machine learning, deep learning, and neural networks. The geographical area of the experiments was also varied, with studies conducted in different countries across the world. Additionally, the input parameters used in the models were unique, ranging from physical, chemical, and biological parameters to meteorological and hydrological data. The results of the investigation indicated that AI models can improve surface water quality monitoring and management by providing accurate and real-time data.

The study also revealed that the effectiveness of AI models in water quality monitoring depends on the selection of appropriate input parameters, the quality of data, and the model's training. The findings of this investigation have significant implications for policymakers and stakeholders involved in water resource management. The use of AI models in surface water quality monitoring can help develop effective strategies for water resource management, ensure the availability of safe and clean water for all, and prevent water pollution. Therefore, policymakers and stakeholders should consider the use of AI models in water quality monitoring and management to address the challenges of traditional monitoring methods. In conclusion, the investigation demonstrated the potential of AI models in surface water quality monitoring and management. Further research and development in this area can lead to the creation of more advanced and accurate AI models, which can revolutionize the way we monitor and manage our water resources.

This study aimed to determine quantitative information about the physical, chemical, and biological characteristics of water through water sampling. The ultimate objective was to analyze the classification of water quality using machine learning algorithms and determine the water quality index using different AI models such as Artificial Neural Networks (ANN) and Long Short-Term Memory (LSTM) profound learning calculations. The dataset used for this study contained 7 critical parameters, and the models developed were assessed depend on several parameters. The results revealed that the proposed models accurately predicted the water quality index and classified the water quality. Machine learning algorithms such as ANN and LSTM are useful in analyzing complex datasets and predicting outcomes. The models developed in this study used these algorithms to accurately predict the water quality index and classify the water quality based on the dataset's parameters. The use of machine learning algorithms such as ANN and LSTM can be an effective tool for predicting and classifying water quality based on various parameters. The findings of this study can have important implications for water management and public health.

Objectives:

- To determine the quantitative information on the physical, chemical, and biological characteristics of water via water sampling.
- To determine the correlation between different parameters of water quality using ANN model and find the accuracy of the model using DT and KNN model.
- To determine the water quality index and water quality classification using LSTM (Long Short-Term Model).
- To determine performance of ANN and LSTM model and comparing results for small dataset.

CHAPTER 4

MATERIALS AND METHODS

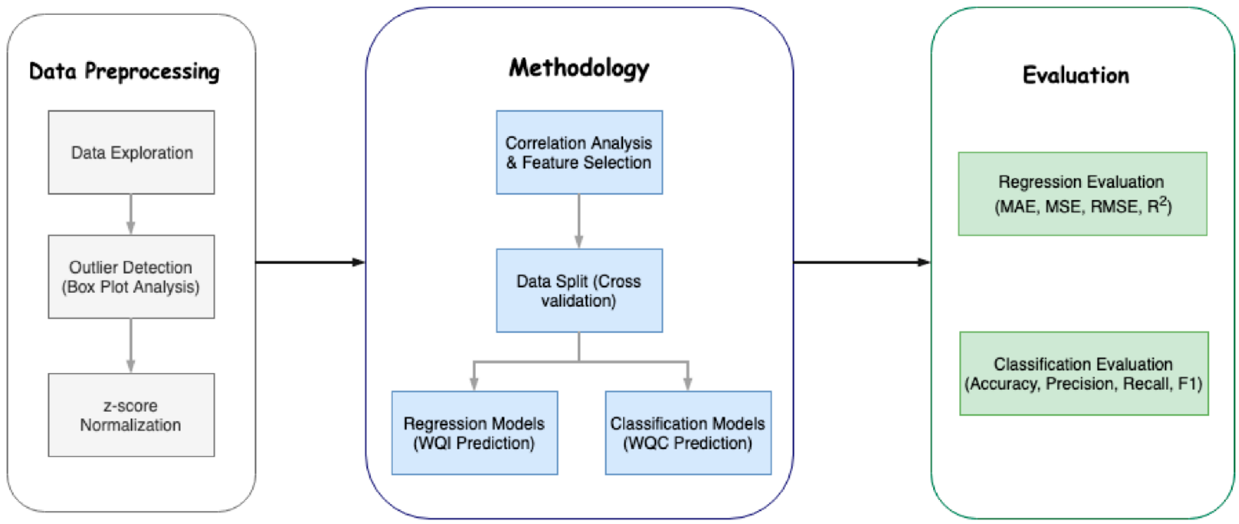


Figure 4.1 Highlights the current strategy of suggested approach.

4.1 Methodology

4.1.1 Description of the study area

The study of water samples from the River Ashwini in Sadhupul, Himachal Pradesh is of utmost importance for assessing the quality of drinking water in the region. The Ashwini River is the primary source of water for many villages and towns in the Solan district, and therefore it is essential to determine its suitability for consumption. The physical and chemical characteristics 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 should include an analysis of the water samples for various parameters, including pH, total dissolved solids, electrical conductivity, turbidity, and microbial contamination. These parameters could indicate the presence of harmful contaminants such as heavy metals, pesticides, and bacteria, which could make the water unfit for human consumption. The results of this study will 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 safety and quality of drinking water in the region to prevent waterborne diseases and promote the overall health and wellbeing of the local population.

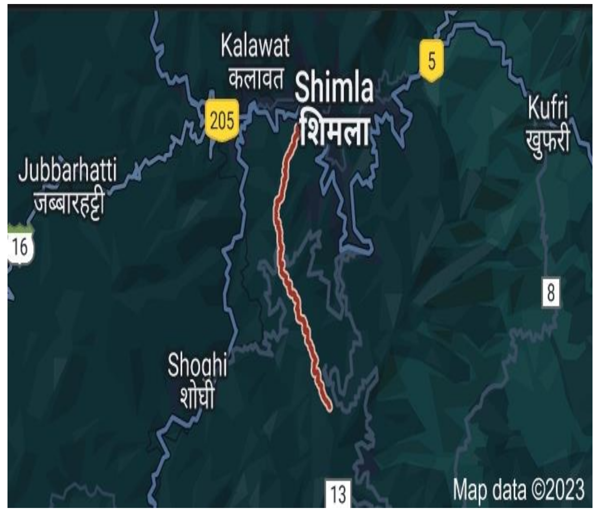


Figure 4.2 Location of Himachal Pradesh in India

Figure 4.3 Location of Ashwini River.

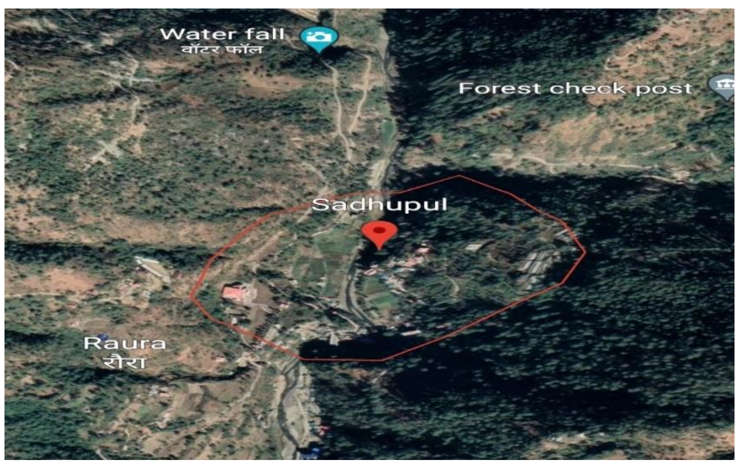


Figure 4.4 Location of Sadhupul village

The Ashwini River originates from the Chail Peak, and its water quality is influenced by this location. Additionally, the elevation of Sadhupul, which is situated at an altitude of 1198 meters above sea level, also affects the water quality of the river. The longitude and latitude of Sadhupul are 34.9702223 and 77.1053593, respectively, which provides an important geographical reference point for understanding the location and characteristics of the Ashwini River.

4.1.2 Data collection and treatment

Water quality refers to the physical, chemical, and biological properties of water and how they impact its suitability for different applications. The quality of water is an essential aspect to consider as it directly affects human health, agriculture, industry, and the environment. This 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 this study based on their significant influence on the features of water quality. These variables are pH, hardness, total dissolved solids, chlorides, turbidity, and dissolved oxygen. These parameters are crucial indicators of water quality, and their levels could affect the suitability of the water for various applications. The study used a limited 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. The results of this study could help to assess the suitability of the water for various purposes, such as drinking, irrigation, and industrial uses. The study could also identify potential sources of contamination and help to develop strategies for protecting and improving the water quality in the area. 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.

4.2 Overview of artificial intelligence model

4.2.1 Artificial neural network (ANN) model

A class of machine learning model called artificial neural networks (ANNs) is modelled after how the human brain functions. ANNs are made up of linked, layered neurons that transmit the results to neurons in the layer below to form the output after each gets data from a variety of other neurons and mathematically processes it. 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 perceptron. A perceptron takes in one or more inputs, multiplies each input by a weight, adds them together, and applies an activation function to produce an output. The activation function is typically a non-linear function, which allows the network to learn complex relationships between inputs and outputs. There are several types of ANNs, including feedforward networks, recurrent networks, and convolutional networks. Feedforward 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. Convolutional 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. Training an ANN involves adjusting the weights between neurons so that the network produces the desired output for a given input. This is typically done using a process called backpropagation, 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. ANNs have been used for a wide range of applications, including image and audio recognition, natural language processing, and financial modeling. They have been particularly successful in tasks like object recognition and speech recognition, where they have achieved human-level performance in some cases. However, ANNs can be computationally expensive to train and can require large amounts of data to achieve good 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. As research in this field continues, it is likely that ANNs will continue to play an important role in many areas of artificial intelligence and machine learning.

It is possible to forecast and monitor water quality using an artificial neural network (ANN). The steps for implementing an ANN model for tracking water quality are as follows:

Data collection: 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. 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. Collecting comprehensive and accurate data on water quality is crucial for ensuring the safety of drinking 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. 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 overfitting. Finally, the testing set is used to evaluate the performance of the trained model on unseen data. 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. 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 hyperparameters to identify the optimal configuration, while cross-validation involves evaluating the performance of the model on different subsets of the data to prevent overfitting.

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. The goal is to optimize the ANN to achieve the highest possible accuracy on unseen data while avoiding overfitting.

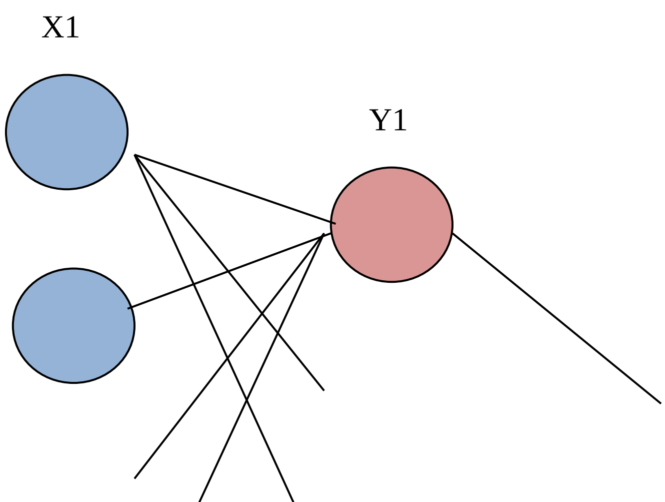
Model testing: To analyze the performance of an artificial neural network (ANN) model, a testing set can be used to evaluate its F1-score, accuracy, precision, and recall. Additionally, other machine learning models like k-nearest neighbor (KNN) and decision tree (DT) can be used to compare their performance 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. 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 model is shown as below

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

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

Where x represents the observed data, y is the predicted data and n is the number of observations.



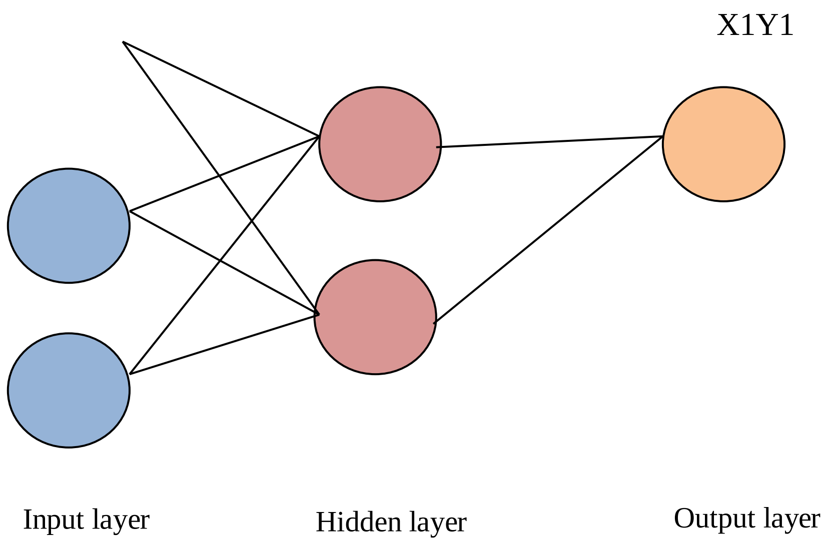


Figure 4.5 ANN MODEL

The model's network structure is set up to allow for a systematic flow of information. Signals are moved from the input layer (independent variables) to the hidden layer for processing before being transmitted to the output layer via a system of weighted connections.

4.2.2 LSTM (Long Short-Term Memory)

The main purpose of creating recurrent neural networks (RNNs) with long short-term memory (LSTM) is to solve the problem of fading gradients in traditional RNNs. When the error gradient in the back propagation strategy shrinks too much, it becomes challenging for the network to learn long-term associations. By including a memory cell that may selectively recall or input information based on the input data, LSTMs address this issue. An LSTM's architecture consists of an input layer, an output layer, and one or more LSTM layers with memory cells. Each memory cell is composed of three gates: an input gate for controlling the flow of new data input, a forget gate for deciding which data to keep or discard, and an output gate for controlling the flow of data output. The gates are managed by sigmoid activation functions that output values between 0 and 1 where 0 represents "forget" or "close" and 1 denotes "input" or "open". LSTMs have proven to perform better than traditional RNNs and other machine learning techniques in a range of applications, making them a popular choice for time series analysis and prediction jobs.

Long short-term memory (LSTM) networks, which Hochreiter and Schmidhuber created in 1997, shattered accuracy records in a number of application fields. About 2007 saw a revolution in speech recognition as LSTM started to outperform traditional models in a number of speech applications. The steps for implementing an ANN model for tracking water quality are as follows:

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. 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. Collecting comprehensive and accurate data on water quality is crucial for ensuring the safety of drinking water, protecting aquatic ecosystems, and monitoring the impacts of human activities on water resources.

Water quality index and classification:

The water quality index (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.

Table 4.1 Standard values of parameters according to WHO

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

$$WQI = \frac{\sum_{i=1}^{N} qi \times xi}{\sum_{i=1}^{N} xi}$$

Where N is the number of parameters, qi is the quality rating scale of each parameter and xi is the unit weight of each parameter.

Following equation can be used to calculate qi and xi:

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

$$xi = \frac{K}{Si}$$

Where Pi is measured values of parameters and P ideal is the ideal value of parameters and Si is the standard value of parameters.

Preprocessing method:

An important stage in preparing data for machine learning is data normalization. Normalization's goal is to scale down input values and output variables to a single scale so they can be compared consistently. One of the most commonly 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 minimum and maximum values of each variable are identified, and the values are rescaled to lie between 0 and 1. This is done by subtracting the minimum value from each value and dividing by the difference between the maximum and minimum values. This results in a new set of values that are all within the range of 0 to 1. For example, suppose we have a dataset containing temperature readings between -20 and 40 degrees Celsius. The minimum temperature is -20 degrees, and the maximum temperature is 40 degrees. To normalize this data using min-max normalization, we subtract -20 from each temperature value and divide by 60 (the difference between 40 and -20). The resulting values will fall within the range of 0 and 1, and we can use them to train an artificial neural network (ANN) model. 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.

$$x = \frac{x - xmin}{xmax - xmin}$$

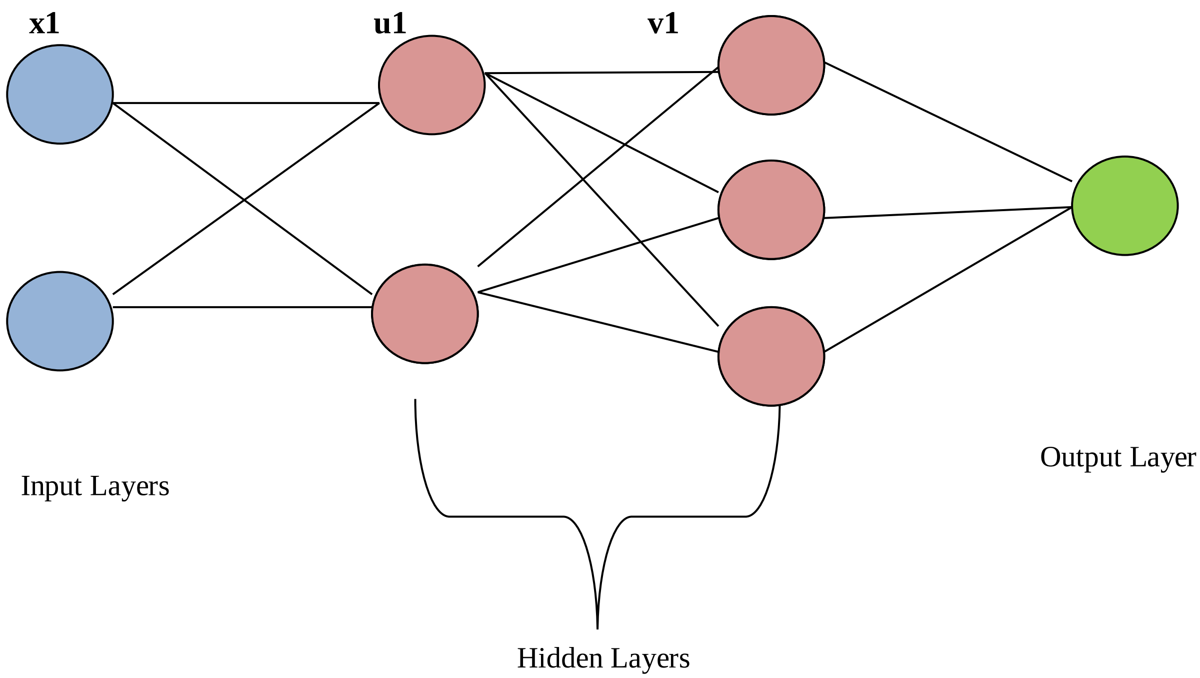


Figure 4.6 LSTM model

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:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(yiobserved - yiestimated)$$

Where yi is the observed and estimated value.

Root-mean square error:

$$RMSE = \sqrt{\sum_{i=1}^{n}\frac{(yi\ observed - yiestimated)^2}{n}}$$

Coefficient of Correlation:

$$R = \frac{n(\sum_{i=1}^{n} yiob \times yiest) - (\sum_{i=1}^{n} yiob)(\sum_{i=1}^{n} yiest))}{\sqrt{n(\sum_{i=1}^{n} yiob)^\wedge 2 - (\sum_{i=1}^{n} yiest)^\wedge 2}}$$

4.3 PARAMETERS AND SENSORS

4.3.1 Parameters used during sampling of data are as follows:

1. **pH:** pH which ranges from 0 to 14, is a gauge of its acidity or alkalinity. While pH values below 7 are acidic and those above 7 are alkaline, pH 7 is neutral. Different elements like dissolved minerals, pollution, and organic matter have an impact on the pH of water. The ideal pH range for drinking water and aquatic life is 6.5 to 8.5. Facilities for wastewater treatment modify the pH to make sure it is within the acceptable range for disposal.

Maintaining the wellbeing of aquatic ecosystems and guaranteeing the safety of drinking water need regular pH monitoring.

2. **Turbidity:** The amount of cloudiness or haze in water is measured by turbidity. Suspended particles including dirt, organic waste, and algae that scatter light and impair sight are the main culprits. A nephelometer, which determines the quantity of light scattered by the suspended particles, may be used to detect turbidity. As turbidity reduces the amount of light available for photosynthesis and can clog fish gills, it can be harmful to aquatic life. High turbidity can also impact the water's visual qualities and make it more challenging to purify it for drinking.

3. **Dissolved Oxygen:** The amount of oxygen gas dissolved in water is measured by the term "dissolved oxygen." The existence of aquatic life depends on DO because many aquatic creatures, including fish, insects, and bacteria, need oxygen to breathe. A number of variables, including temperature, pressure, salinity, and the presence of photosynthetic organisms, can affect the DO levels in water. Human activities such as pollution and fertilizer enrichment can also have an impact on DO levels. Fish deaths and other detrimental effects on aquatic ecosystems can result from low DO levels. Monitoring DO levels in water is crucial for preserving the ecological balance of aquatic systems and safeguarding the health of aquatic life**.**

4. **Total dissolved solids:** The combined amount of inorganic and organic compounds in water that are too tiny to be removed by filtration is known as total dissolved solids. Minerals, salts, metals, and other substances are included in TDS. TDS levels that are too high in drinking water can have negative effects on flavor, appearance, and health. A number of variables, including natural geology, industrial and agricultural operations, and human waste, can affect TDS levels. Therefore, it's crucial to keep an eye on TDS levels to ensure the water is safe and of high quality.

5. **Hardness:** The amount of calcium and magnesium ions in water is measured by how hard the water is. Hard water can impair the effectiveness of soaps and detergents, produce scaling in pipes and appliances, and harm industrial operations. Consequently, it's crucial to monitor water hardness for both household and industrial applications.

6. **Chloride:** Chlorides in water describes the quantity of chloride ions in water. One of the main inorganic anions present in natural water sources are chloride ions, which can be produced by a variety of processes including industrial processes, natural weathering of rocks, and wastewater discharges. High chloride concentrations in drinking water can have a detrimental influence on taste, and high concentrations in aquatic habitats can harm aquatic life.

4.3.2 Sensor used during sampling:

1. **pH sensor:** The pH of a solution, such as water, may be measured with a pH sensor. Hydrogen ion (H+) concentration in a solution, which defines whether it is acidic or alkaline, is what pH sensors measure. The secondary process' pH must be maintained between 6.3 and 7.85. A pH sensor can identify and report pH value variations.

Figure 4.7 pH sensor

2. **Dissolved oxygen sensor:** The amount of oxygen gas dissolved in water is measured by a device known as a dissolved oxygen (DO) sensor. A membrane is used in the operation of DO sensors to distinguish the sensor element from the water's oxygen content. They are susceptible to influences from variables including temperature, pressure, and interference from other gases in the water, and they need calibration and routine maintenance to assure correct observations. For microorganisms to exist, their habitat must have the proper level of oxygen. The optimal concentration of dissolved oxygen is 2 mg/L.

Figure 4.8 Dissolved oxygen sensor

3. **Turbidity meter:** A turbidity meter is a tool used to gauge how cloudy or turbid a liquid, such as water, is. Turbidity meters measure the quantity of light scattered by liquid-containing particle density. To evaluate the purity of the water, they are frequently employed in wastewater treatment and water quality monitoring. Nephelometric turbidimeters, which employ a light-scattering methodology, and formazin turbidimeters, which use a chemical method, are two examples of the many varieties of turbidity meters that are available. To provide reliable results, turbidity meters need to be calibrated and maintained regularly.

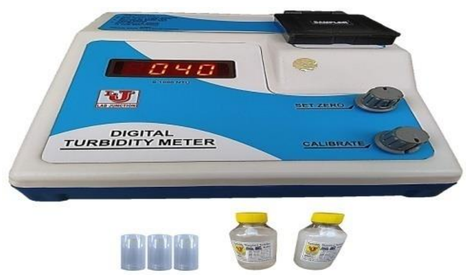


Figure 4.9 Turbidity meter

4. **Total dissolved solids sensor:** It calculates the quantity of dissolved solids in a sample of water. The concentration of dissolved inorganic and organic compounds in water is measured using a TDS (Total Dissolved Solids) sensor. The total number of ions present in the water is measured using TDS sensors employing electrical conductivity. To assure the safety and quality of water, they are frequently employed in industrial applications and water quality monitoring.

Figure 4.10 TDS sensor

4.4 CHEMICAL USED

Table 4.2 Chemicals used during sampling of data

S.NO.	Parameters	Chemical used
1.	Chlorides	Silver Nitrate in presence of dilute nitric acid to form silver chloride
2.	Alkalinity	Methyl Orange as indicator, Sulphuric acid, Phenolphthalein
3.	Hardness	Ethylene Diamine Tetra Acetic acid (EDTA)
4.	Acidity	Methyl Orange as indicator, NaOH

CHAPTER 5

WORK DONE

5.1 Experiment

In this work, the SES preprocessing approach and updated LSTM and ANN models were used to forecast water quality in order to anticipate the features of the water quality in surface water. In this study, we have compared two different models i.e., artificial neural network model and long short-term model. ANN model presents the data in the form of histograms that shows us the correlation between different parameters. But in case of LSTM model, it tells us about water quality index, MSE, RMSE. We can also test the model accuracy by using two different classifiers i.e., KNN (k –nearest neighbor) and DT (decision tree) model.

5.1.1 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. 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.

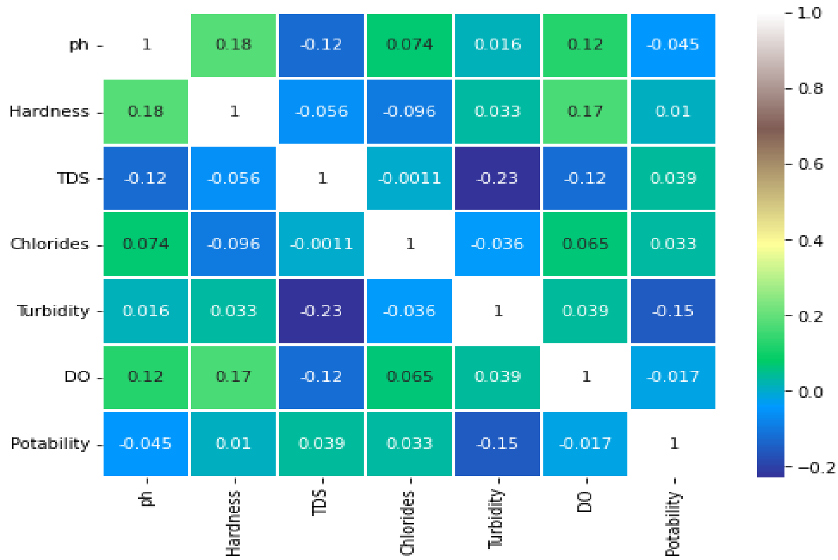


Figure 5.1 Heat Map plot by ANN model

Fig 5.1 talk about the correlation of two different parameters. We 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. 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, the majority of the data points on the graph follow a distinct pattern but a small number of points fall outside of this pattern.

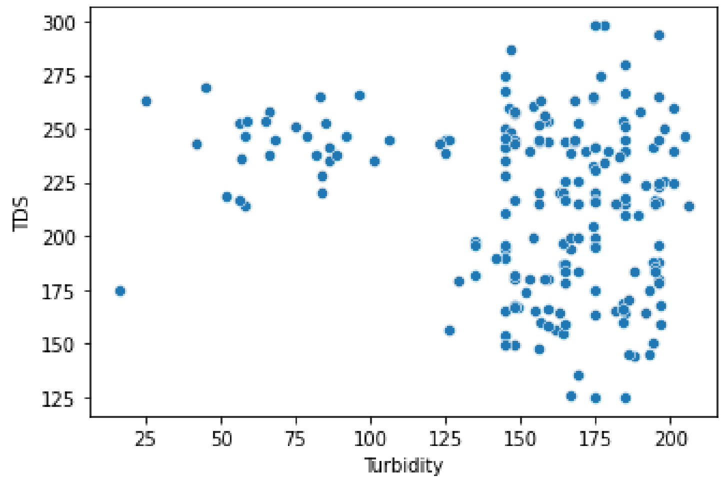


Figure 5.2 Correlation between TDS and Turbidity

5.1.2 Distplot graph: The distplot displays the data distribution of a single variable in comparison to the density distribution.

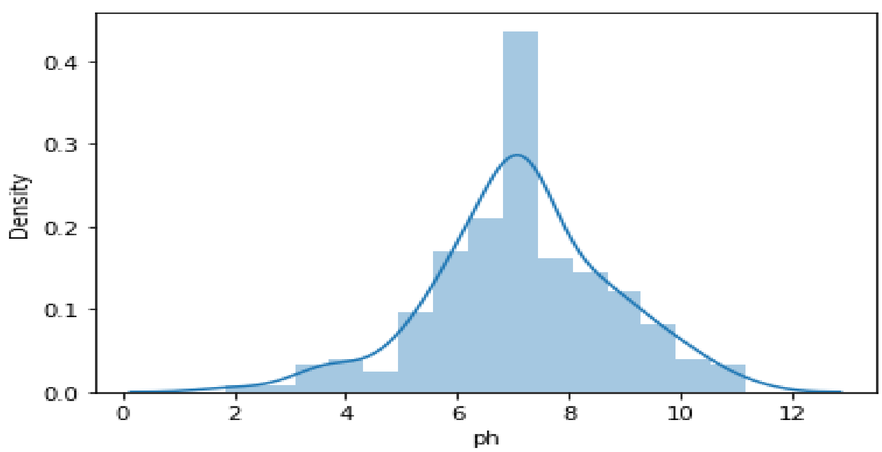


Figure 5.3 Distplot graph

5.1.3 Boxplot graph: Boxplots are used to gauge how evenly dispersed a data set's data are. It creates three quartiles out of the data set. The lowest, maximum, median, first quartile, and third quartile of the data set are shown in this graph.

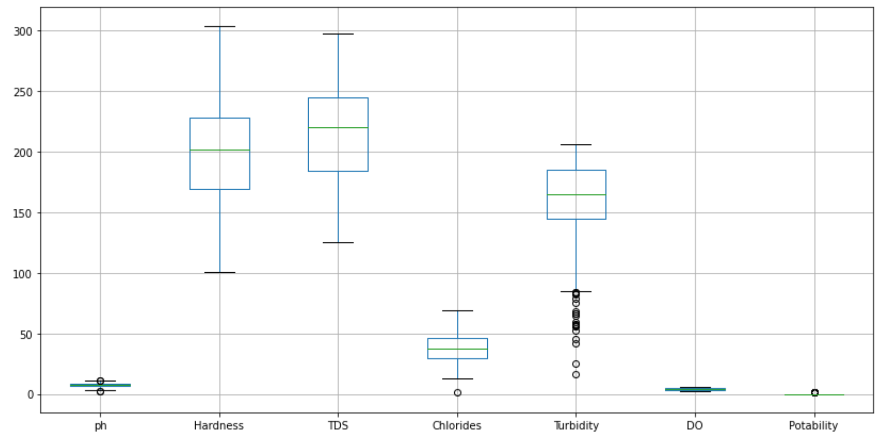


Figure 5.4 Boxplot graph

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

5.1.4 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 judgements 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. After that, it descends the tree to the following node and bases its judgement 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. Following figure shows accuracy of ANN model using 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 5.5 Accuracy of DT model

5.1.5 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 taken into account. The algorithm determines the distances between each new data point and every other data point in the training set in order 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. Following figure shows accuracy of ANN model using KNN classifier.

```
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 5.6 Accuracy of KNN model

But in case of LSTM model, we have found the accuracy of model using DT classifier. Accuracy of model using DT comes out to be 95% which is more than ANN model.

Following figure shows the output generated by LSTM model`

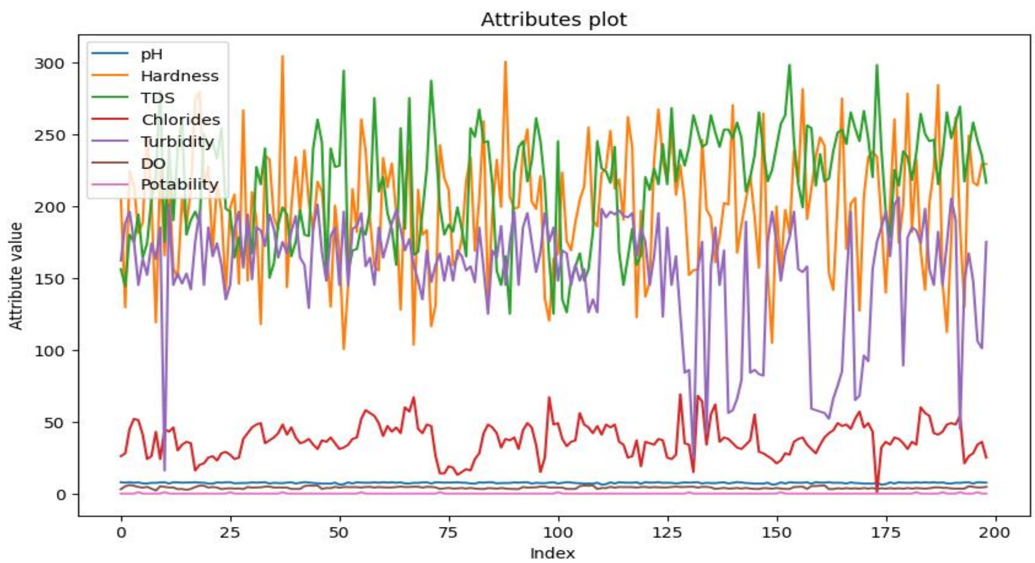


Figure 5.7 Output generated by LSTM Model

Mean square error and Root mean square error:

Table 5.1 MSE and RMSE values for ANN and LSTM model

MODELS	MSE	RMSE
ANN model	0.52	0.60
LSTM model	0.04	0.21

CHAPTER 6

RESULT AND DISCUSSION

It was investigated if artificial intelligence algorithms could replace more traditional techniques for estimating and forecasting water quality. Because of the modelling and forecasting of water quality, the time and resources needed for laboratory analysis have been greatly and crucially decreased. The Ashwini River, which flows through the tiny Himachal Pradesh village of Sadhupul, was the subject of the case study.

With the use of a potent artificial intelligence model, our goal is to create a real-time system and test a fresh strategy for accurately anticipating and classifying water quality. This work proposes merging the discussed artificial intelligence methods to precisely duplicate water levels and quality. This dataset had a total of six parameters. The study came to the conclusion that categorization and forecasting of water quality may be done using LSTM and ANN models. The purpose of this study was to show how the LSTM and ANN models may be used to forecast the quality of surface water.

Based on previous measurements, we provide predictions for future water quality levels in this analysis. 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, or relatively close to actual values. A lower MSE indicates that the model is better at predicting the output values. The average squared difference between the anticipated values and the actual values is what this metric measures.

However, the MSE value in the ANN model is also less than 1, though 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.

CHAPTER 7

CONCLUSION

In recent years, the use of artificial intelligence (AI) models in monitoring and evaluating water quality has become increasingly popular. This is due to the fact that 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. One of the major findings of the study was that Long short-term memory (LSTM) and Artificial Neural Networks (ANN) are the most commonly used AI models for water quality monitoring and evaluation in the past decade. 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. ANN, on the other hand, 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. The study also found that Iran and Southeast Asia account for the majority of 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 R2 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. The study also examined the input parameters and output measures that were used in surface water quality monitoring and evaluation. The most commonly studied parameters included pH, turbidity, Total Dissolved Solids (TDS), hardness, chlorides, and dissolved oxygen. These parameters are all important indicators of water quality and can be used to identify potential sources of contamination or other issues. 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. While the findings of the study shed light on the use of neural networks in surface water quality prediction, 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. While the study found that LSTM and ANN are the most commonly used neural network models, there are many other topologies that have the potential to improve accuracy and performance. Future studies could explore the use of Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Deep Belief Networks (DBN), 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. Future studies could focus on these regions to identify the unique challenges and opportunities for using neural networks in water quality prediction. In addition to these issues, there are also several other areas where future research could be focused. There is also a need for more research on the use of neural networks in real-time water quality monitoring systems, which could help improve the speed and accuracy of water quality data collection. Overall, the study provides valuable insights into the use of neural networks in surface water quality prediction. However, there are still many issues that need to be addressed to improve the accuracy and applicability of these models. By addressing these issues and conducting further research in this area, it is likely that the accuracy and applicability of neural network models in water quality prediction will continue to improve in the future. 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.

REFERENCES

1. P. Zeilhofer, L. V. A. C. Zeilhofer, E. L. Hardoim, Z. M. Lima, and C. S. Oliveira, “GIS applications for mapping and spatial modeling of urban-use water quality: a case study in District of Cuiabá, Mato Grosso, Brazil,” *Cadernos de Saúde Pública*, vol. 23, no. 4, pp. 875–884, 2007.View at: Publisher Site | Google Scholar

2. M. A. Kahlown, M. A. Tahir, and H. Rasheed, *National Water Quality Monitoring Programme, Fifth Monitoring Report (2005–2006)*, Pakistan Council of Research in Water Resources Islamabad, Islamabad, Pakistan, 2007, http://www.pcrwr.gov.pk/Publications/Water%20Quality%20Reports/Water%20Quality%20Monitoring%20Report%202005-06.pdf.

3. UN water, “Clean water for a healthy world,” Tech. Rep., Development, 2010, https://www.undp.org/content/undp/en/home/presscenter/articles/2010/03/22/clean-water-for-a-healthy-world.html.View at: Google Scholar

4. K. Farrell-Poe, W. Payne, and R. Emanuel, *Water Quality & Monitoring*, University of Arizona Repository, 2000, http://hdl.handle.net/10150/146901.

5. T. Taskaya-Temizel and M. C. Casey, “A comparative study of autoregressive neural network hybrids,” *Neural Networks*, vol. 18, no. 5–6, pp. 781–789, 2005.View at: Publisher Site | Google Scholar

6. C. N. Babu and B. E. Reddy, “A moving-average filter based hybrid ARIMA-ANN model for forecasting time series data,” *Applied Soft Computing*, vol. 23, pp. 27–38, 2014.View at: Publisher Site | Google Scholar

7. X. Zhang, N. Hu, Z. Cheng, and H. Zhong, “Vibration data recovery based on compressed sensing,” *Acta Physica Sinica*, vol. 63, no. 20, pp. 119–128, 2014.View at: Publisher Site | Google Scholar

8. M. M. S. Cabral Pinto, C. M. Ordens, M. T. Condesso de Melo et al., “An inter-disciplinary approach to evaluate human health risks due to long-term exposure to contaminated groundwater

near a chemical complex," *Exposure and Health*, vol. 12, no. 2, pp. 199–214, 2020.View at: Publisher Site | Google Scholar

9. M. M. S. Cabral Pinto, A. P. Marinho-Reis, A. Almeida et al., "Human predisposition to cognitive impairment and its relation with environmental exposure to potentially toxic elements," *Environmental Geochemistry and Health*, vol. 40, no. 5, pp. 1767–1784, 2018.View at: Publisher Site | Google Scholar
10. Y. C. Lai, C. P. Yang, C. Y. Hsieh, C. Y. Wu, and C. M. Kao, "Evaluation of non-point source pollution and river water quality using a multimedia two-model system," *Journal of Hydrology*, vol. 409, no. 3-4, pp. 583–595, 2011.View at: Publisher Site | Google Scholar
11. J. Huang, N. Liu, M. Wang, and K. Yan, "Application WASP model on validation of reservoir-drinking water source protection areas delineation," in *2010 3rd International Conference on Biomedical Engineering and Informatics*, pp. 3031–3035, Yantai, China, October 2010.View at: Publisher Site | Google Scholar
12. I. R. Warren and H. K. Bach, "MIKE 21: a modelling system for estuaries, coastal waters and seas," *Environmental Software*, vol. 7, no. 4, pp. 229–240, 1992.View at: Publisher Site | Google Scholar
13. D. F. Hayes, J. W. Labadie, T. G. Sanders, and J. K. Brown, "Enhancing water quality in hydropower system operations," *Water Resources Research*, vol. 34, no. 3, pp. 471–483, 1998.View at: Publisher Site | Google Scholar
14. G. Tang, J. Li, Z. Zhu, Z. Li, and F. Nerry, "Two-dimensional water environment numerical simulation research based on EFDC in Mudan River, Northeast China," in *2015 IEEE European Modelling Symposium (EMS)*, pp. 238–243, Madrid, Spain, October 2015.View at: Publisher Site | Google Scholar
15. L. Hu, C. Zhang, C. Hu, and G. Jiang, "Use of grey system for assessment of drinking water quality: a case S study of Jiaozuo city, China," in *2009 IEEE International Conference on Grey Systems and Intelligent Services (GSIS 2009)*, pp. 803–808, Nanjing, China, November 2009.View at: Publisher Site | Google Scholar

16. E. Batur and D. Maktav, “Assessment of surface water quality by using satellite images fusion based on PCA method in the Lake Gala, Turkey,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 5, pp. 2983–2989, 2019.View at: Publisher Site | Google Scholar

17. S. Jaloree, A. Rajput, and G. Sanjeev, “Decision tree approach to build a model for water quality,” *Binary Journal of Data Mining & Networking*, vol. 4, pp. 25–28, 2014.View at: Google Scholar

18. J. Liu, C. Yu, Z. Hu et al., “Accurate prediction scheme of water quality in smart mariculture with deep Bi-S-SRU learning network,” *IEEE Access*, vol. 8, pp. 24784–24798, 2020.View at: Publisher Site | Google Scholar

19. H. Liao and W. Sun, “Forecasting and evaluating water quality of Chao Lake based on an improved decision tree method,” *Procedia Environmental Sciences*, vol. 2, pp. 970–979, 2010.View at: Publisher Site | Google Scholar

20. L. Yan and M. Qian, “AP-LSSVM modeling for water quality prediction,” in *Proceedings of the 31st Chinese Control Conference*, pp. 6928–6932, Hefei, China, July 2012.View at: Google Scholar

21. A. Solanki, H. Agrawal, and K. Khare, “Predictive analysis of water quality parameters using deep learning,” *International Journal of Computers and Applications*, vol. 125, no. 9, pp. 29–34, 2015.View at: Publisher Site | Google Scholar

22. X. Li and J. Song, “A new ANN-Markov chain methodology for water quality prediction,” in *2015 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–6, Killarney, Ireland, July 2015.View at: Publisher Site | Google Scholar

23. A. A. M. Ahmed and S. M. A. Shah, “Application of adaptive neuro-fuzzy inference system (ANFIS) to estimate the biochemical oxygen demand (BOD) of Surma River,” *Journal of King Saud University - Engineering Sciences*, vol. 29, no. 3, pp. 237–243, 2017.View at: Publisher Site | Google Scholar

24. Y. Khan and C. S. See, “Predicting and analyzing water quality using Machine Learning: a comprehensive model,” in *2016 IEEE Long Island Systems, Applications and Technology Conference (LISAT)*, pp. 1–6, Farmingdale, NY, USA, April 2016.View at: Publisher Site | Google Scholar

25. J. Yan, Z. Xu, Y. Yu, H. Xu, and K. Gao, “Application of a hybrid optimized BP network model to estimate water quality parameters of Beihai Lake in Beijing,” *Applied Sciences*, vol. 9, no. 9, p. 1863, 2019.View at: Publisher Site | Google Scholar

26. H. R. Maier, A. Jain, G. C. Dandy, and K. P. Sudheer, “Methods used for the development of neural networks for the prediction of water resource variables in river systems: current status and future directions,” *Environmental Modelling & Software*, vol. 25, no. 8, pp. 891–909, 2010.View at: Publisher Site | Google Scholar

27. S. Lee and D. Lee, “Improved prediction of harmful algal blooms in four major South Korea’s rivers using deep learning models,” *International Journal of Environmental Research and Public Health*, vol. 15, no. 7, p. 1322, 2018.View at: Publisher Site | Google Scholar

28. U. Shafi, R. Mumtaz, H. Anwar, A. M. Qamar, and H. Khurshid, “Surface water pollution detection using internet of things,” in *2018 15th International Conference on Smart Cities: Improving Quality of Life Using ICT & IoT (HONET-ICT)*, pp. 92–96, Islamabad, Pakistan, October 2018.View at: Publisher Site | Google Scholar

29. Z. Ahmad, N. A. Rahim, A. Bahadori, and J. Zhang, “Improving water quality index prediction in Perak River basin Malaysia through a combination of multiple neural networks,” *International Journal of River Basin Management*, vol. 15, no. 1, pp. 79–87, 2016.View at: Publisher Site | Google Scholar

30. V. Ranković, J. Radulović, I. Radojević, A. Ostojić, and L. Čomić, “Neural network modeling of dissolved oxygen in the Gruža reservoir, Serbia,” *Ecological Modelling*, vol. 221, no. 8, pp. 1239–1244, 2010.View at: Publisher Site | Google Scholar

31. N. M. Gazzaz, M. K. Yusoff, A. Z. Aris, H. Juahir, and M. F. Ramli, “Artificial neural network modeling of the water quality index for Kinta River (Malaysia) using water quality variables as predictors,” *Marine Pollution Bulletin*, vol. 64, no. 11, pp. 2409–2420, 2012.View at: Publisher Site | Google Scholar

32. H. Z. Abyaneh, “Evaluation of multivariate linear regression and artificial neural networks in prediction of water quality parameters,” *Journal of Environmental Health Science and Engineering*, vol. 12, no. 1, p. 40, 2014.View at: Publisher Site | Google Scholar

33. M. K. Yusoff, A. Z. Aris, H. Juahir, and M. F. Ramli, "Artificial neural network modeling of the water quality index for Kinta River (Malaysia) using water quality variables as predictors," *Marine Pollution Bulletin*, vol. 64, no. 11, pp. 2409–2420, 2012.View at: Publisher Site | Google Scholar

34. H. Z. Abyaneh, "Evaluation of multivariate linear regression and artificial neural networks in prediction of water quality parameters," *Journal of Environmental Health Science and Engineering*, vol. 12, no. 1, p. 40, 2014. View at: Publisher Site | Google Scholar

35. M. Sakizadeh, "Artificial intelligence for the prediction of water quality index in groundwater systems," *Modeling Earth Systems and Environment*, vol. 2, no. 1, p. 8, 2016.View at: Publisher Site | Google Scholar

36. M. I. Yesilnacar, E. Sahinkaya, M. Naz, and B. Ozkaya, "Neural network prediction of nitrate in groundwater of Harran Plain, Turkey," *Environmental Earth Sciences*, vol. 56, no. 1, pp. 19–25, 2008.View at: Publisher Site | Google Scholar

37. M. Bouamar and M. Ladjal, "A comparative study of RBF neural network and SVM classification techniques performed on real data for drinking water quality," in *2008 5th International Multi-Conference on Systems, Signals and Devices*, pp. 1–5, Amman, Jordan, July 2008. View at: Publisher Site | Google Scholar

38. N. Marir, H. Wang, G. Feng, B. Li, and M. Jia, "Distributed abnormal behavior detection approach based on deep belief network and ensemble SVM using spark," *IEEE Access*, vol. 6, pp. 59657–59671, 2018. View at: Publisher Site | Google Scholar

39. Z. M. Fadlullah, F. Tang, B. Mao, J. Liu, and N. Kato, "On intelligent traffic control for large-scale heterogeneous networks: a value matrix-based deep learning approach," *IEEE Communications Letters*, vol. 22, no. 12, pp. 2479–2482, 2018. View at: Publisher Site | Google Scholar

40. S. Maiti and R. K. Tiwari, "A comparative study of artificial neural networks, Bayesian neural networks and adaptive neuro-fuzzy inference system in groundwater level prediction," *Environmental Earth Sciences*, vol. 71, no. 7, pp. 3147–3160, 2014. View at: Publisher Site | Google Scholar

41. C. Min, “An improved recurrent support vector regression algorithm for water quality prediction,” *Journal of Computational Information*, vol. 12, pp. 4455–4462, 2011. View at: Google Scholar

42. R. Das Kangabam, S. D. Bhoominathan, S. Kanagaraj, and M. Govindaraju, “Development of a water quality index (WQI) for the Loktak Lake in India,” *Applied Water Science*, vol. 7, no. 6, pp. 2907–2918, 2017. View at: Publisher Site | Google Scholar

43. G. Srivastava and P. Kumar, “Water quality index with missing parameters,” *International Journal of Research in Engineering and Technology*, vol. 2, no. 4, pp. 609–614, 2013. View at: Publisher Site | Google Scholar

44. S. Tyagi, B. Sharma, P. Singh, and R. Dobhal, “Water quality assessment in terms of water quality index,” *American Journal of Water Resources*, vol. 1, no. 3, pp. 34–38, 2013. View at: Publisher Site | Google Scholar

45. A. A. Al-Othman, “Evaluation of the suitability of surface water from Riyadh Mainstream Saudi Arabia for a variety of uses,” *Arabian Journal of Chemistry*, vol. 12, no. 8, pp. 2104–2110, 2019. View at: Publisher Site | Google Scholar

46. T. H. H. Aldhyani, M. Alrasheedi, A. A. Alqarni, M. Y. Alzahrani, and A. M. Bamhdi, “Intelligent hybrid model to enhance time series models for predicting network traffic,” *IEEE Access*, vol. 8, pp. 130431–130451, 2020. View at: Publisher Site | Google Scholar

47. S. Jaloree, A. Rajput, and G. Sanjeev, “Decision tree approach to build a model for water quality,” *Binary Journal of Data Mining & Networking*, vol. 4, pp. 25–28, 2014.View at: Google Scholar

48. J. Liu, C. Yu, Z. Hu et al., “Accurate prediction scheme of water quality in smart mariculture with deep Bi-S-SRU learning network,” *IEEE Access*, vol. 8, pp. 24784–24798, 2020.View at: Publisher Site | Google Scholar

49. H. Liao and W. Sun, “Forecasting and evaluating water quality of Chao Lake based on an improved decision tree method,” *Procedia Environmental Sciences*, vol. 2, pp. 970–979, 2010.View at: Publisher Site | Google Scholar

50. L. Yan and M. Qian, “AP-LSSVM modeling for water quality prediction,” in *Proceedings of the 31st Chinese Control Conference*, pp. 6928–6932, Hefei, China, July 2012.View at: Google Scholar