



CSE 499
One Semester Long Single Project

*An IOT Based Smart Filter to Predict Safe Drinking Water
Supply for School Children*

AUTHOR

Md. Shafiuzzaman Khan
ID: 181014055

SUPERVISOR

Dr. Muhammad Golam Kibria
Associate Professor, Department of CSE

A project report submitted in partial fulfillment of the requirements
for the degree of Bachelor of Science in Computer Science and Engineering

Department of Computer Science and Engineering
University of Liberal Arts Bangladesh
688 Beribadh Road Mohammadpur, Dhaka – 1207, Bangladesh

CERTIFICATE OF APPROVAL

The project report entitled “An IOT Based Smart Filter to Monitor Safe Drinking Water Supply for School Children by Using Machine Learning” is submitted to the **Department of Computer Science and Engineering**, University of Liberal Arts Bangladesh, Dhaka in partial fulfillment of the requirements for the degree of the Bachelor of Science.

Dated: 6th February, 2022

Abul Kalam Al Azad, PhD
Associate Professor and Acting Head
Department of Computer Science and Engineering
University of Liberal Arts Bangladesh

Signature & Date

Muhammad Golam Kibria, PhD
Associate Professor
Department of Computer Science and Engineering
University of Liberal Arts Bangladesh

Signature & Date

DEDICATION

I am dedicating my project to my parents and my supervisor who had faith, trust, and unconditional support throughout the semester. Specially dedicated my project to my mother who gave me relentless support.

Especially thanking my teachers and partners for their patience, support, love, and perseverance during the ups and downs to accomplish this project.

I am dedicating my report to them and ALLAH who is glorious.

Acknowledgement

I would want to convey my heartfelt gratitude to Dr. Muhammad Golam Kibria, my project supervisor for allowing me to collaborate with him on this project. I owe him a great debt of gratitude for his direction and regular supervision, as well as for giving vital project information.

The Department of Computer Science and Engineering is also a significant part since it has given me the chance to enhance my talents and grow as a person. It goes without saying that the essential assistance I've received from the department's whole professors and staff has had a significant positive impact on my undergraduate studies.

I would want to express my gratitude to my family members for their support and encouragement during this whole project, as well as during the last four years of my undergraduate studies. I would not have been able to complete this project without my parents' love and care, and I am grateful to them for that.

Md. Shafiuzzaman Khan

University of Liberal Arts Bangladesh

Date: 6th February, 2022

ABSTRACT

In recent times Internet of Things (IOT) is widely frequently utilized in several research fields for monitoring, collecting, and analyzing data from the perspective of helping people. In Bangladesh, water pollution has become one of the most serious threats for children in recent years and in most areas, the drinking water is mostly contaminated and polluted. Humans, especially children can be affected by various water-borne diseases due to drinking polluted water. If drinking water contamination is recognized early enough, appropriate steps can be implemented and possibly dangerous situations can be minimized. The quality of water should be checked in real-time to ensure a constant supply of pure drinking water. Innovative and smart systems to monitor drinking water contamination are becoming interesting and fascinating these days with the help of advanced sensors, the Internet of Things, clouding, and machine learning technologies. With advancements in communication technologies, smart water quality monitoring solutions are becoming more popular. This project provides a comprehensive summary of recent research on the subject of smart drinking water quality monitoring for school children. This project proposes a low-cost water quality monitoring system leading to new technologies such as Machine Learning, the Internet of Things, and Cloud Computing that can replace traditional quality monitoring methods.

Keyword: PH Sensor, Turbidity Sensor, Temperature sensor, Arduino Uno, Wi-Fi module, Cloud Storage, Machine learning, LSTM.

List of Figures

Figure 1: Block Diagram	10
Figure 2: SDG Goal 6- Clean Water and Sanitation	11
Figure 3: Statistics of Safe Water Supply in Asia	12
Figure 4: Students at Kharikshyan Governmental School	13
Figure 5: Timeline of Project Management	14
Figure 6: Smart Water Quality Monitoring System with Cost Effective Using IOT	15
Figure 7: IOT Based Smart Water Quality Monitoring System	16
Figure 8: Internet of Things Enabled Real-Time Water Quality Monitoring System	17
Figure 9: Smart Water Quality Monitoring System	18
Figure 10: Water Quality Monitoring System	19
Figure 11: System Design Model of Water Quality Monitoring System	20
Figure 12: Water Monitoring in Large Aquatic Area	21
Figure 13: System Architecture	22
Figure 14: Arduino UNO R3	23
Figure 15: PH Sensor	24
Figure 16: Temperature Sensor DS18B20	25
Figure 17: Turbidity Sensor	26
Figure 18: ESP 8266 Wi-Fi Module	27
Figure 19: Home Filter	28
Figure 20: Data Collection by Arduino IDE	29
Figure 21: A Complete Data Flow Diagram on Predictive WQM System	30
Figure 22: Neural Network	31
Figure 23: Recurrent Neural Network (RNN)	32
Figure 24: A Structure of Long-Short Term Memory Neural Network	33
Figure 25: IOT Sensors Setup for Data Collection	34

List of Figures

Figure 26: A Circuit Diagram of PH Sensor	35
Figure 27: A Circuit Diagram of DS18B20 Temperature Sensor	36
Figure 28: A Circuit Diagram of Turbidity Sensor	37
Figure 29: Implemented Code	38
Figure 30: Water Quality Dataset 2005-2021	39
Figure 31: Finding Null Value	40
Figure 32: Comparing Target Water Quality with Total Count	41
Figure 33: Data Analysis by Mean, Max and Min	42
Figure 34: Hourly Temperature Condition in 2005-2021	43
Figure 35: Daily turbidity Condition in 2005-2021	44
Figure 36: Daily PH Condition in 2005-2021	45
Figure 37: Correlations of Turbidity and Temperature	46
Figure 38: Correlations of PH and Temperature	47
Figure 39: Box Plot of Temperature in Month	48
Figure 40: Box Plot of PH in Month	49
Figure 41: Box Plot of Turbidity in Hour	50
Figure 42: Predicting Temperature by LSTM	51
Figure 43: Predicting PH by LSTM	52
Figure 44: Predicting Turbidity by LSTM	53

List of Tables

Table 1: WHO Authorized Water Parameters and Range	40
--	----

TABLE OF CONTENT

TITLE	PAGE
CERTIFICATE OF APPROVAL	ii
DEDICATION	iii
ACKNOWLEDGEMENT	iv
ABSTRACT	v
LIST OF FIGURES	vi-vii
LIST OF TABLES	vii
 Chapter 1: Introduction	
1.1 Introduction	1
1.2 Motivation	2
1.3 Rationale of the Study	2
1.4 Research Questions	4
1.5 Expected Output	7
1.6 Project Management and Finance	7
1.7 Report Layout	8
 Chapter 2: Background	
2.1 Terminologies	10
2.2 Related Works	11
2.3 Comparative Analysis and Summary	21
2.4 Scope of the Problem	21
2.5 Challenges	22
 Chapter 3: Research Methodology	
3.1 An IOT Based Smart Filter to Predict Safe Drinking Water Supply by Using Machine Learning	23
3.2 Instrumentation	24
3.3 Data Collection Procedure	28
3.4 Statistical Analysis	29
3.5 Applied Mechanism	31
3.6 Implementation Requirements	32

Chapter 4: Experimental Results and Discussion

4.1 Experimental Setup	35
4.2 Experimental Results and Analysis	41

Chapter 5: Impact on Society, Environment and Sustainability

5.1 Impact on Society	49
5.2 Impact on Environment	50
5.3 Ethical Aspects	51
5.4 Sustainability Plan	51

Chapter 6: Conclusion and Implication for Future Research

6.1 Summary of the Study and Conclusions	52
6.2 Implication for Further Study	52

References

Appendix

Chapter 1: Introduction

1.1 Introduction

Clean and Safe drinking water is an essential thing for our entire life and wellbeing. To survive and healthy life, access to safe drinking water is one of the basic amenities for humankind and mostly for children. In heavily populated areas safe drinking water is most needed to reduce the habit of drinking unsafe water. In Bangladesh, groundwater is mostly considered as drinking water among the various sources in both rural and urban regions. The resources of fresh drinking water are sufficient in Bangladesh but those resources are increasingly polluted by dangerous contaminants because of rapid industrialization and significant population development. In the aspect of the present situation, school children are in dangerous situations when they mostly drink unsafe water while they attend school. Floods and cyclones are two natural phenomena that have an impact on drinking water quality. Consumers are suffering from various diseases after consuming these harmful contaminants from unsafe drinking water especially children are dying in significant numbers every year. World Health Organization (WHO) estimated that roughly 80% of diseases are caused by contaminated water either directly or indirectly. There are many water parameters that determine the quality of safe drinkable water and those parameters could be both physiochemical and microbiological. Predicting the water quality by using a machine learning algorithm and with the help of IOT devices would be very helpful for school children. This smart system can forecast the water quality by using LSTM (Long Short-Term Memory) time series algorithm.

1.2 Motivation

In Bangladesh every year about 45 thousand children die because of various water diseases such as cholera, fluorosis, dysentery, typhoid, and polio. They are mostly school students. Most of the time those children drink water from the filters while they attending in school. These filters can purify water for a specific period of time but after a period of time, the germ kill kit that mainly kills all germs stops functioning or does not work properly. So at that time children start to drink unsafe water. This situation made me think a lot so I decided to work on that. So that after completing my project I can apply this knowledge practically in schools where they use filters to drink water for their children. Also, I did not notice such works in Bangladesh that have started working in the aspect of children's health. That is why this made me inspired to think about the health of Bangladeshi Children's and I want to start my project on that criteria. My objective is to make an IOT-based smart filter system to monitor safe drinking water supply for school children. This intelligent process will predict the quality of drinking water and can read the condition continuously of water by using some sensors in filters that are used in schools in Bangladesh. By sensing the basic parameters of water, this smart system will notify users via mobile phone or email. By using a machine-learning algorithm model, this system will work intelligently. It will analyze past data and will give predictive decisions if any changes are important to take.

1.3 Rationale of the Study

This smart system will monitor the quality of water and predict the quality for future by using some basic water parameter sensors. It will measure the PH value by PH sensor, as well as turbidity of water will be measured by turbidity sensor and the temperature of water will be monitored by DS18B20 waterproof sensor. All sensors will be connected with a microcontroller Arduino UNO R3. The Sensors data will pass through the Wi-Fi protocol ESP 8266 to the clouding System. All data will be stored in this API and a machine learning model which is made my LSTM algorithm will predict the future data of filter water by observing

present data. If the parameters of water may cross the threshold, this model will catch it and the user will be notified by email.

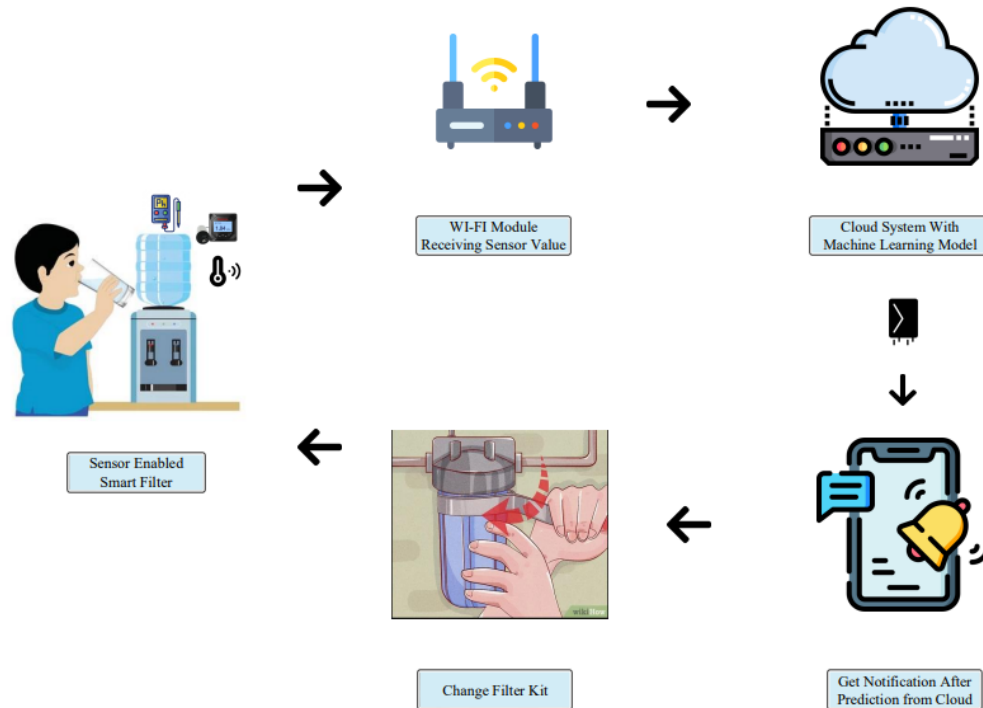


Figure 1: Block Diagram

The block diagram is showing whole scenario about my project. This smart filter is for school students where sensors are attached with the filter and the cloud system will be worked with the help of machine learning model LSTM. School students will be able to drink water from this filter until the water is safe.

1.4 Research Questions

1.4.1 Sustainable Development Goal 6: Clean Water and Sanitation

Improvement of clean and safe drinking water is one of the important goals of SDG among 17 goals. The United Nations General Assembly set 17 Sustainable Development Goals in 2015, with a 2030 deadline, based on the principle of "leaving no one behind." This goal says to ensure the availability and sustainable management of clean and safe drinking water. My project will ensure to drink safe drinking water by using this smart filter system and will help to move forward to achieve this SDG goal in the aspect of school children's health. [18]



Figure 2: SDG Goal 6 – Clean Water and Sanitation

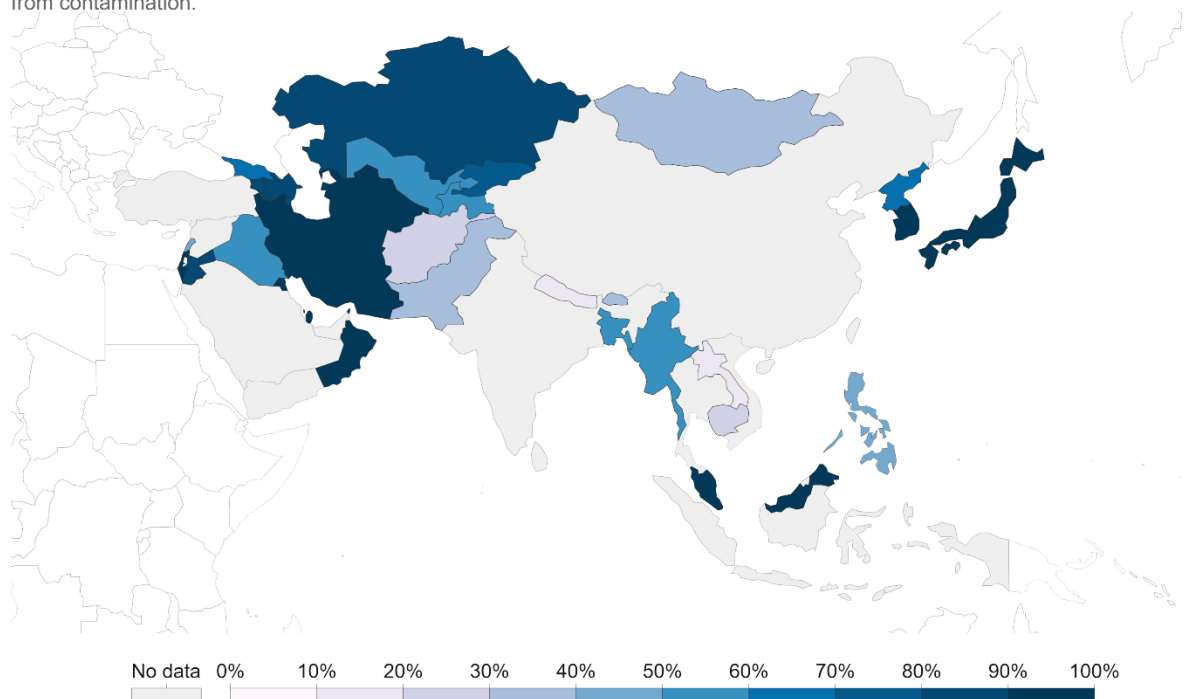
1.4.2 Safe Drinking Water

In Bangladesh, a large number of people are not getting access to pure safe drinking water. Bangladesh was found to be one of the poorest performing countries in South Asia, following only Nepal and Afghanistan in a global study on the health status of 188 countries based on health-related Sustainable Development Goals (SDG) parameters. This analysis simply proves that without addressing the crucial components of safe drinking water, sanitation, and hygiene, no country can make lasting progress toward global health goals. According to WHO and UNICEF Bangladesh confirmed only 58.51% of safe areas where people are staying to drink safe water. This is the percentage of properly managed drinking water service, accessible when needed, and free of contaminants. This is one of the reasons to choose this project to fulfill the SDG goal to hundred percent in between 2030.

Share of the population using safely managed drinking water, 2020

A safely managed drinking water service is defined as one located on premises, available when needed and free from contamination.

Our World
in Data



Source: WHO/UNICEF Joint Monitoring Programme for Water Supply, Sanitation and Hygiene (JMP) OurWorldInData.org/water-access • CC BY

Figure 3: Statistics of Safe Water Supply in Asia

1.4.3 Arsenic Contamination in Drinking Water



Figure 4: Students at Kharikshyan Government School in Chittagong Hill Tracts

In 2013, more than 97 percent of Bangladesh's population had access to better water sources, indicating that the country has achieved tremendous progress in this area. However, only 34.6 percent of people have access to safe drinking water. Between 2000 and 2012, the percentage of people who drank arsenic-contaminated water fell from 26.6 percent to 12.4%. Despite this, Bangladesh continues to have the world's highest proportion of people exposed to arsenic contamination. Despite progress, 19.4 million people continue to drink water that contains arsenic levels that exceed national health regulations. Difficulties to safe drinking water, as well as sanitation, have a considerable impact on children's health, nutrition, education, safety, and other outcomes. Every year, 45,000 children die from diarrhea. According to World Health Organization research, over 45,000 children under the age of five die every year in Bangladesh as a result of diarrhea caused by contaminated water. That is why I have made my decision to make a system which will help school children at least when they will stay in school. Because the chances to drink contaminated water in school is very high than at home. [25]

1.5 Expected Output

It will ensure the water quality to save children from life-threatening diseases such as diarrhea, cholera, fluorosis, bone deformities, typhoid, dysentery, polio, etc. This smart system will forecast and notify users by analyzing data whenever the filter will not be able to purify the water correctly or if any dangerous element accidentally come through the filter. This smart system is cost-effective and anyone from rural or urban areas can afford this. So if it can develop properly it can save lots of children's lives all around our country. This smart system will ensure accuracy in measurement.

1.6 Project Management and Finance

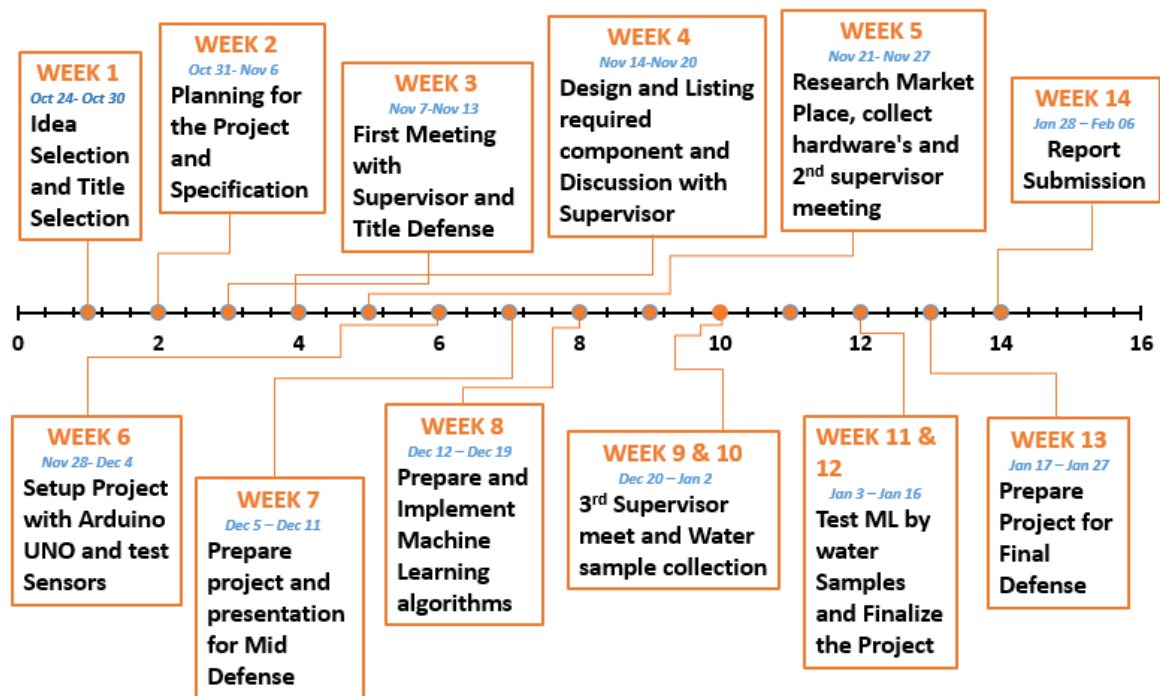


Figure 5: Timeline of Project Management

Whenever I have confirmed about to work on this project I started to learn about IOT devices and their interactions through Arduino IDE software. I have decided to work on three sensors that are financially manageable for me. I have collected PH sensor, turbidity sensor, and temperature sensor for data collection, and Arduino IDE R3 microcontroller. This instrument was manageable and financially affordable at all. Then I collected water data for around 2 months for my project. Then I set a data set for my machine learning. After the mid defense, I started my work on machine learning to predict by analyzing the given dataset. I trained my model by giving 15.5 years' data. This dataset has been made on the basis of the first 2 months' data that I took from my home filter. Then I prepared the remaining data for my project. Otherwise, a little data will not be good, and also we cannot expect a good output from the machine learning model. I used the LSTM algorithm to forecast the water quality. And in the last, I prepared a report on the basis of my whole work.

1.7 Report layout

Here I am giving my project layout which I will cover in the next chapters:

Chapter 2: Here I will provide all of my project backgrounds. Here I will explain why I chose this project and why I want to develop it, as well as some similar work that has previously been done, and what impact our country will have, as well as what will be the best solution for our country.

Chapter 3: In this chapter, I will also introduce which components, techniques, algorithms, and instruments I used in my project. I will discuss the dataset and procedure I have done for my project. There will be system architecture, problem statement, and proposed methodology of my project, and also some basic diagrams will be provided in this section.

Chapter 4: In this chapter, I will discuss the experimental setup of my project. I will show all sensor's setup and their circuit connection in this section. I also discuss the machine learning

algorithm I have used in my project. There I will show all statistics of my given dataset. It will also show the co-relations of PH, turbidity, and temperatures with each other.

Chapter 5: This section will be covered with the social impact of my project. How much this will contribute in social development I will discuss them. It will also cover the ethical issues and health aspects. The sustainability of my project also will be discussed in this chapter.

Chapter 6: In the End, I will conclude my discussion with a whole summary. I will summarize my entire project's work and discuss some of the project's future plans.

Chapter 2: Background

2.1 Terminologies

IOT - Internet of Things. The internet of things is a combined network of physical objects where they all are set in with sensors, internet, software, and other technologies to connect each other and also to exchange various data over the internet.

SDG – Sustainable Development Goal. Development that fulfills present needs without affecting future generations' ability to meet their own is referred to as sustainable development. The major goal is to end poverty, protect the environment, and ensure that everyone lives in peace and prosperity by 2030.

LSTM – Long Short Term Memory. Long short-term memory is a deep learning architecture based on artificial recurrent neural networks (RNNs). LSTM model is well suited to classify, processing and forecasting on the basis of time series of data.

Sensor - Sensors are electronic devices that detect and react to changes in their surroundings. For example, temperature, motion, light, and pressure are all examples of possible inputs.

Arduino IDE – Arduino IDE is a software where it mainly used to interface IOT electronics devices by giving formal instructions. This instruction has been written in C and C++ programming language. This software normally provides lots of outputs by following the common instructions.

PH - PH is a function of the hydrogen ion concentration, which indicates whether a solution is acidic or alkaline. Typically, the pH scale ranges from 0 to 14.

Turbidity - Turbidity indicates the existence of water cloudiness caused by a large number of small particles that are typically invisible to the naked eye.

ADC – Analog to digital convertor. It can receive random input as waves, particles, electrons, heat and convert them into digital output to understand the concentration and amounts of these particles.

2.2 Related Works

Safe drinking water is both a basic need for good health and a basic human right. In many places of the world, freshwater is already in short supply. Increased population, urbanization, and climate change will make it much more restrictive in the next century. Drinking water quality refers to the relationship between the composition of water and the impact of natural and human processes. Chemical compounds are introduced into the water supply system through leaks and cross-connections, resulting in deterioration of drinking water quality.

Sathis and Gandla used four sensors in their IOT-based smart system to monitor water quality. These four sensors are pH, turbidity, water level, and temperature sensor. They have used an ultrasonic sensor to measure the water level and DHT-11 is used to measure the temperature of the water. Those sensors are connected to a microcontroller unit (Arduino Mega) which is the main processing unit in this system. For data transmission, they developed the system with ESP8266 WIFI module which is connected with PC. [1]

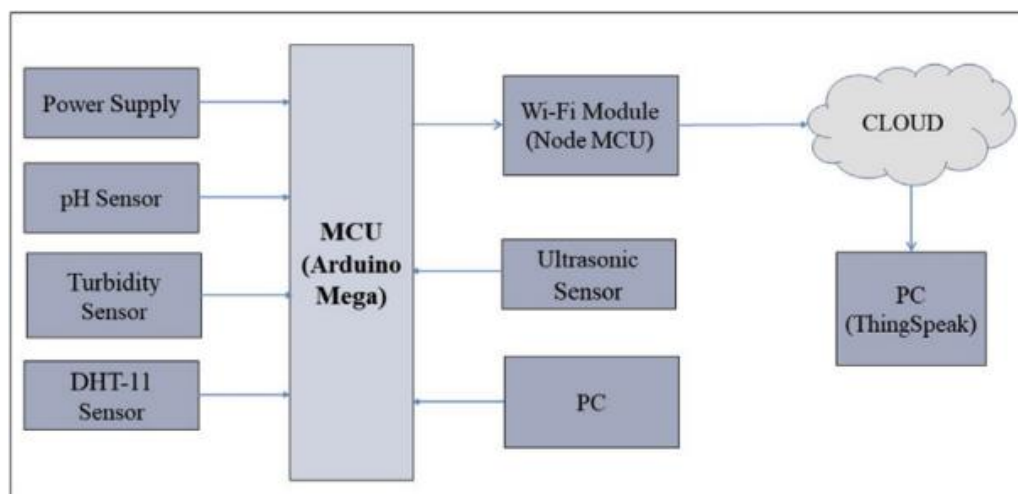


Figure 6: Smart water quality monitoring system with cost-effective using IoT

Varsha, Anjitha, Akshay, Aruna, Jagadeesh, and Audre used a total of six sensors when they established their smart water quality monitoring system. They took pH sensor, temperature,

turbidity, conductivity, humidity, and CO₂ sensor in their system. They used a core controller which is the heart of the system and this system controls all operations. All of these sensors are connected with this core controller and through this, it gives the sensor measurements from the water. This core controller also compares these sensor data with the standard values and gives notification to the end-users by a wireless module they used. [2]

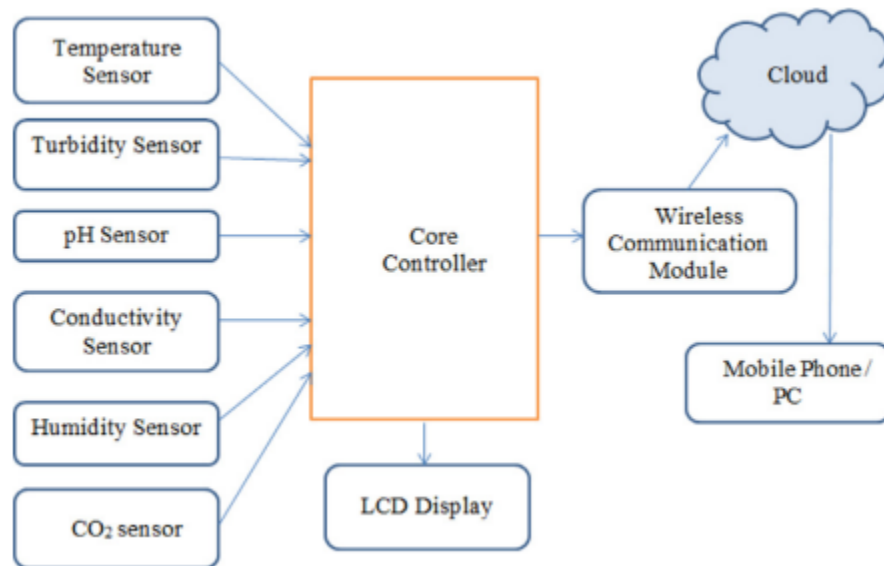


Figure 7: IoT based smart water quality monitoring system

Geetha and Gouthami made their research on water quality monitoring systems, which will monitor the water quality in real-time. They mentioned the main reason for unsafe water is the overexploitation of natural resources by human beings. The advanced technology for industrialization and agriculture is responsible for this situation. Poor water quality spreads disease, kills people, and discourages socioeconomic development. According to the water resource information system in India, in 2017 around 5 million had died because of waterborne diseases. Farmers use fertilizers and pesticides and those are washed through soil by rainwater and spread water diseases when they ended up in water resources. They made the system to

monitor the water quality in real-time where they used key parameters pH, conductivity, turbidity, and water level detector sensor. The controller is performing as a central unit in this system which is attached with these sensors. The sensors values can be shown by LCD display and those sensor data will have sent to the desktop-based cloud system. A message will have sent to the user's smartphone from the cloud if the quality of water exceeds its range. [3]

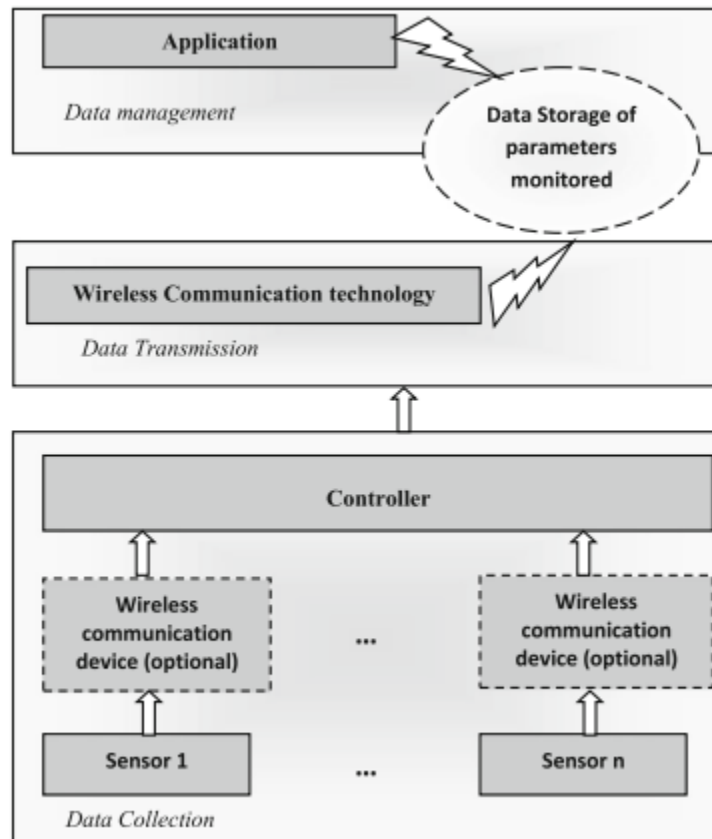


Figure 8: Internet of things enabled real-time water quality monitoring system

Prasad, Mamun, Islam, Haqva designed a smart water quality monitoring system. They used pH, conductivity, and temperature Sensors, an analog to digital converter (ADC), a microcontroller, SD storage, and a GSM module to make up the entire system. The acquired data can be saved onboard using an SD card or transferred to a File Transfer Protocol (FTP) or

cloud server. For this project, a cloud server is used in conjunction with a local workstation for data analysis. The sensors, microcontroller, and GSM communication were all part of a comprehensive packaged set that was used for this project. To attain the desired result, the system is designed with a sleep mode that allows the device to rest for 15 minutes after an hour of continuous readings. [4]



Figure 9: Smart Water Quality Monitoring System

Prashant, Vamshi, Ramakrishna proposed a model named IOT Based Smart Water Quality Monitoring and Prediction System. They used several sensors like pH, temperature sensor, and water flow detector observer's sensor. This research said, pH is the main parameter when we think about the quality of water. By pH, we can measure the alkalinity level of acidity in water. It is manufactured by the DF robot and we can get accurate data from pH sensors by using a digital meters and it will give reading continuously. This research also shows that we would take pH values between zero to fourteen in zero to eighty-degree Celsius temperature. The

temperature sensor is used to sense the temperature of the water. According to capacity temperature sensors can be eleven types. They choose waterproof ds18b20 for their project which has one wire interface itself. [7]

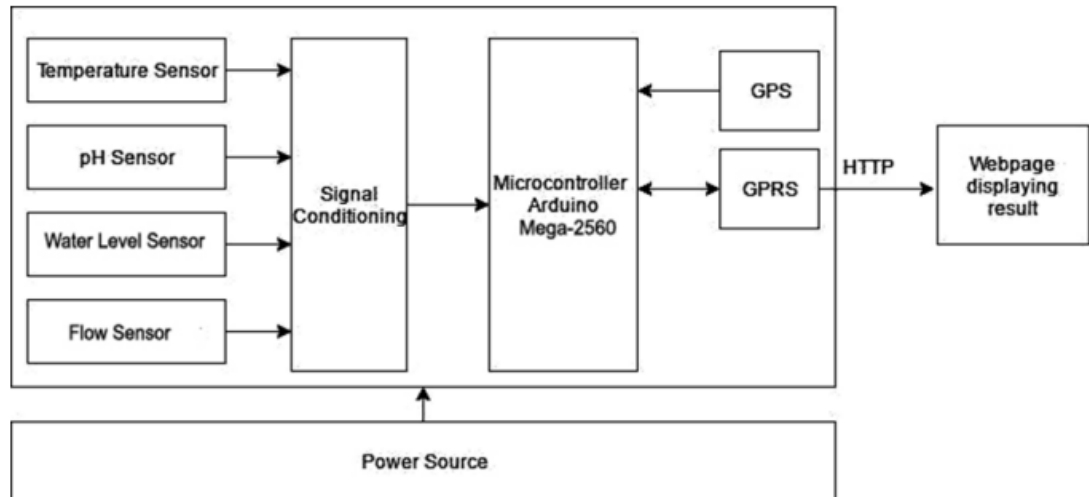


Figure 10: Water Quality Monitoring System

We can determine the acidic power of water by the measure of pH level. The pH of pure drinking water is between 6.5-8.5. Water with lower pH becomes acidic and has a higher corrosive power. Electrical conductivity and total alkalinity have a positive relationship with pH. A lower pH can lead human beings to cancer whereas a higher pH can induce incrustation and chlorination difficulties. In most of the analyses from given samples, the researchers observed that the turbidity level of water is between 0.47 NTU and 6.90 NTU. Because of high turbidity, people are suffering from waterborne diseases and gastric disorders. When the amount of dissolved solids in water is high it indicates the water is polluted. Large-scale agriculture activities, as well as specific geographical factors, can cause high TDS levels in drinking water. TDS is made up of both volatile and non-volatile solids. The TDS level in good quality drinking water should be less than 600 and a study found that the TDS levels in the samples ranged from 171.2 mg/l to 641 mg/l. The existence of chloride in water indicates that the groundwater has been affected by sewage percolation and wastewater. A high quantity of chloride in water suggests the presence of organic waste as a result of wastewater

contamination. When the quantity of chloride increases in water it can cause cardiovascular problems and give the water a harsh flavor. [5]

The developed smart system of Anantha and Dr. Geetha said that the acidity or alkalinity of a solution can be measured by a pH sensor. It is said the logarithm scale for pH is 0-14 but the neutral point is 7. When the value is below 7 it indicates the water would be acidic. The turbidity sensor is mainly used to measure the cloudiness of water. It is mainly used to observe how much the water is clean. It shows some degree when the water loses its transparency. The temperature sensor shows the water temperature. The DS18B20 temperature sensor is used in this system and can measure the temperature from -50 to +125 degrees Celsius. [8]

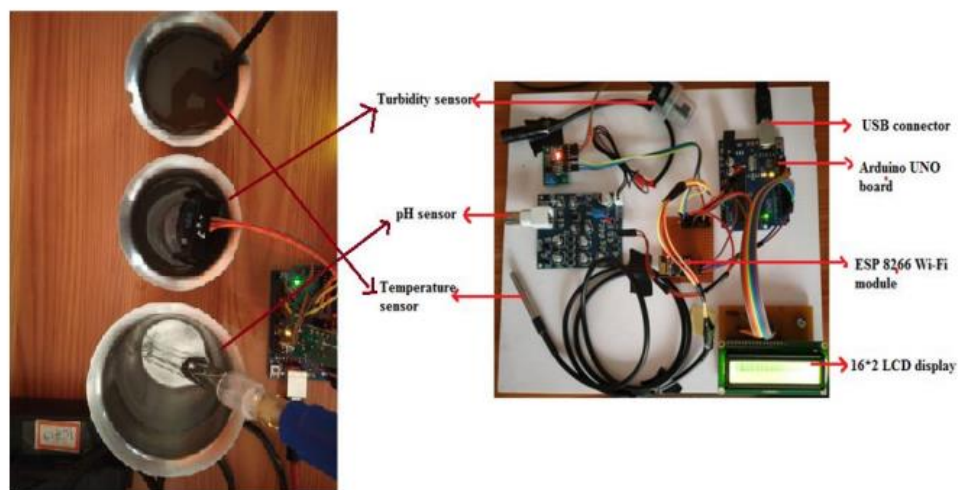


Figure 10: System Design Model of water quality monitoring system

S. Angel, kayalvizhi, Bhavadharini, and Kalpana developed a system that is mainly a machine learning-based water quality measurement monitoring smart system. This intelligent system will detect the water quality automatically and they used three IoT sensors which will help to get desired data. They took pH, turbidity, and conductivity sensors. They implemented a machine learning model in their project. This model will predict and provide decisions with the help of previous data which is called training data. These data set will be analyzed and take decision by the K Means algorithm of machine learning. They gathered all sensor data from tap water, drinking water, lemon water, salt water, mud water to make future decisions accurately. They considered pH range 6.5 to 8 as pure water and less than 7 is acidic. Their

research said the conductivity range of drinking water is 5-50 mS/m. They also included that, the turbidity level of drinking water should be 2NTU and it will not be drinkable if the range becomes greater than 5NTU. [9]

Abhinandhan and Kuldeepak worked on a water quality monitoring system that is IOT and machine learning-based. Their research says, around 200 children are dying due to the drinking water crisis in the world, and mostly 3 million deaths in a year which is huge. To monitor drinking water, they developed a system using machine learning and cloud-based application in their system which will ensure the safe drinking water and protect children from water diseases such as bone deformities and cholera. They collected various data through sensors from water and send these data to Node MCU and Node MCU sending these data to a cloud-based database. They used cloud-based visualization tools to visualize and analyze data. Then the system sends emails to concerned users. [10]

Vaishnavi, Daigavane, and Gaikwad entitled “Water Quality Monitoring System Based on IOT” where they focused to ensure safe drinking water supply in real-time. They mentioned water is polluting all over the world and they want to work on that. They used basic water parameter sensors to observe the real-time condition of the water. They took a pH sensor, flow sensor, turbidity sensor, and temperature sensor, and all these sensors are processed by an Arduino microcontroller. These sensor data will be viewed by a WIFI system. Their research shows the pH range of drinking water is 6 to 8.5 and below 7 indicates the water is acidic. The pH sensor is operated by a 5V power supply connection from the microcontroller. They took a DS18B20 temperature sensor that can measure from -55 to +125 degrees Celsius. They integrated ATmega128P as a microcontroller in their system. They used ESP 2866 Wi-Fi module which is extremely cost-effective and provides a better connection. [11]

Chetna, Shreya, Harshal, and Nitin developed their project on an advanced IOT based water quality monitoring system. According to their research, a huge amount of water is wasted due to poor management and uncontrolled use of water sources. Because of that, there is now a small amount of drinkable water available on the surface of our earth. About 2439 children died in 2018 due to various water diseases such as cholera, typhoid, diarrhea, etc. Their report

says 1.3 core people were affected by these dangerous diseases. The children whose age is below; their death rate is 60% when they are affected by these water-related diseases. A total of 584 people were died in 2018 due to hepatitis. Poor management is one of the reasons for the shortage of drinking water. When the water overflows in water tanks, schools or colleges that time water wastage is occurring. To solve this problem, they created a system that will automatically detect the water level and quality with the help of sensors. Their proposed model consists pH sensor, water flow sensor, and ultrasonic sensor, and these sensors are connected to ESP 2866 micro core controller. They can observe the sensor values through WI-FI which is also built in the WI-FI module ESP 8266. [12]

System Design

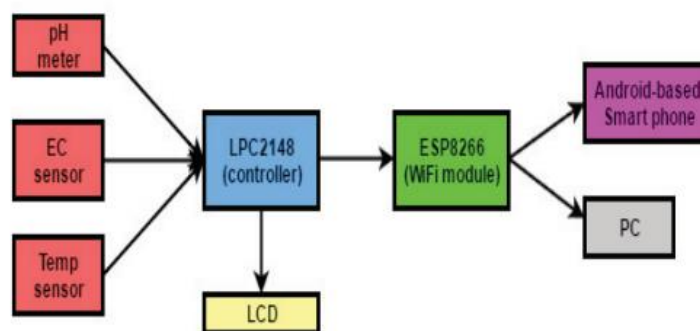


Figure 11: Water Quality Monitoring System

Brinda and Jain created a system that can measure real-time water quality by sensing some water basic parameters. They tended to use pH sensors, temperature, and conductivity sensors in their developed project. They used the ZigBee module to collect all sensor values and send them to the microcontroller. The Microcontroller sends the data with the help of a GSM module to a smartphone or PC without a wire connection. This system also sends a notification to users when the sensor can sense their polluted water is existing. This also can operate the real-time data and the central system can access those data by using the 3G or GSM module.

This project is time-consuming and not economical. They used mainly two data transferring IOT devices ZigBee and GSM Module and connected to the microcontroller. ZigBee mainly sends sensor data to the microcontroller and through the GSM module notifying users by sending messages on smartphones and PC. The sensor data are coming from the analog signal of the microcontroller. There is an ADC, which converts all these signals into digital formats and this digitally formatted data pass by ZigBee to the microcontroller. Then microcontroller process those data and after analyzing all data it sends a message by the GSM module to the users. ESP8266 has been used as a Wi-Fi module in their system which is a low-cost Wi-Fi module and it has Wi-Fi chips on it full of TCP/IP. [13]

M Cho, G Lenin, L Huo, and Prassana established a project on real-time quality measures with the help of the internet of things. They developed their system based on a wireless sensor network in an IOT based environment. They designed their model to observe the pH value, temperature value, water level, turbidity, and carbon dioxide level in the water. Object-oriented program C++ has been used in their software system. They wanted to develop a system that will be cost-effective and consume less power. [14]



Figure 12: Water Monitoring in large Aquatic Area

Alexander, Demetillo, Michelle, and Japitana worked on a system of water quality monitoring with wireless sensor network technology. The system is planned and implemented with the

primary purpose of monitoring water quality, and it is appropriate for use in a vast aquatic region. It can also measure and store data in a database in real-time. This system has a place to provide timely notifications to the system's management or authorities, as well as users. The system's most important function is to ensure that the data collected from the sensor accurately reflects the actual aquatic scenario and that the data is transmitted and delivered in a timely manner as a web display of information or as a short message service (SMS) sent to pre-identified key users. Data collection and storage, data transmission and display (charts, graphs, and datasheets), power management for longer measurement durations, and a maintenance interface for administrators were all integrated into this system. [15]

2.3 Comparative Analysis and Summary

As far as I have checked there are some similar projects that have been done before. Like water quality monitoring and real-time IOT based water quality monitoring system which are a little similar in the sense of gathering water data by IOT devices. But still, there is no work has been done yet as I proposed for school children. There is also some work which was implemented various machine learning models to justify water quality but there is no work found which can forecast water quality for school children. That is why my IOT based smart filter for school children which can predict water quality will be a unique project. Also, my project will fulfill the SDG's 6th goal "Clean Water and Sanitation" from the aspect of Bangladesh by 2030.

2.4 Scope of the Problem

There are many scopes regarding this project. From the aspect of children's health, it is a very necessary project that has to be completed by someone or an organization. Every year a tremendous number of school children suffer water-borne diseases and a large amount of them also die. From research, it is found that when children are staying at home and drinking water they do not affect much but when they go outside, mainly when they go to their schools and drink water from non-hygienic areas or filters they become sick after drinking from these sources. That is why my proposed project is well suited and will be a great achievement if it can apply in every corner school of Bangladesh. After all the 6th goal of sustainable development is to ensure clean drinking water. If our government wants to work on school children's health issues by providing safe clean drinking water my smart filter can be a very effective and well-suited option. This filter will be affordable and cost-effective for everyone. Anyone from rural and urban areas will be able to buy this filter at a reasonable price.

2.5 Challenges

There are several challenges that I have to overcome to complete my project properly. The IOT devices that I'm using all are technical materials. Sometimes they do not work properly if there is any lacking exist in my working procedure. The Arduino codes do not run if all required library is not provided in the Arduino IDE. Sometimes sensors do not show appropriate values in that case the whole circuit connection has to be set again. For my project major challenge is the machine learning model. There is no data set for water quality in online resources by which I can compare with my collected data. That is why the machine learning algorithm that I will use, will not provide accuracy level properly. Because of that I have ready a dataset by my own and train them to predict the future data of filter.

Chapter 3: Research Methodology

3.1 An IOT Based Smart Filter to Predict Safe Drinking Water by Using Machine Learning

In this research, machine learning will be used to predict the future data of filter water. For this prediction, I have to apply a neural network algorithm named LSTM. LSTM is a Long-Short Term Memory that can predict more accurately than its previous version of recurrent neural network (RNN). The cloud system will be an API that will be built with this machine learning model and the data set will update continuously when it will get data through the Wi-Fi module. The sensors also read data continuously when they are all connected to the central microcontroller board, Arduino Uno. The Arduino Uno is the central processing unit of this system. It will send all data to the cloud server through Wi-Fi module ESP 8266. The user will notify by the system when it can predict the water quality reach or cross its threshold.

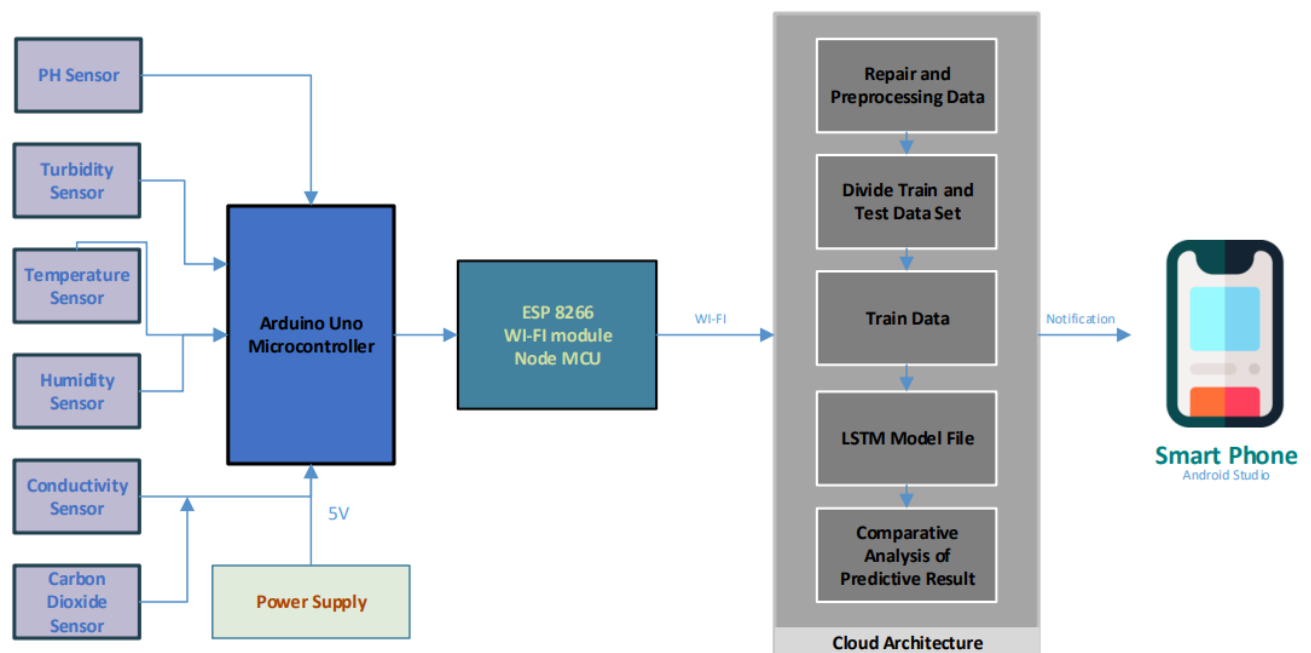


Figure 13: System Architecture

3.2 Instrumentation

For this project, I have been using a microcontroller unit Arduino UNO, three basic parameter sensors and they are pH sensor, Turbidity Sensor and Temperature Sensor. A Wi-Fi module ESP 8266 will be used. These are the physical instrumentations for my project. I will try to show interfacing of pH sensor, temperature sensor and turbidity sensor with Arduino UNO. Arduino Uno is a microcontroller and central unit of my project. I used a USB cable from my PC to power up the Arduino Uno.

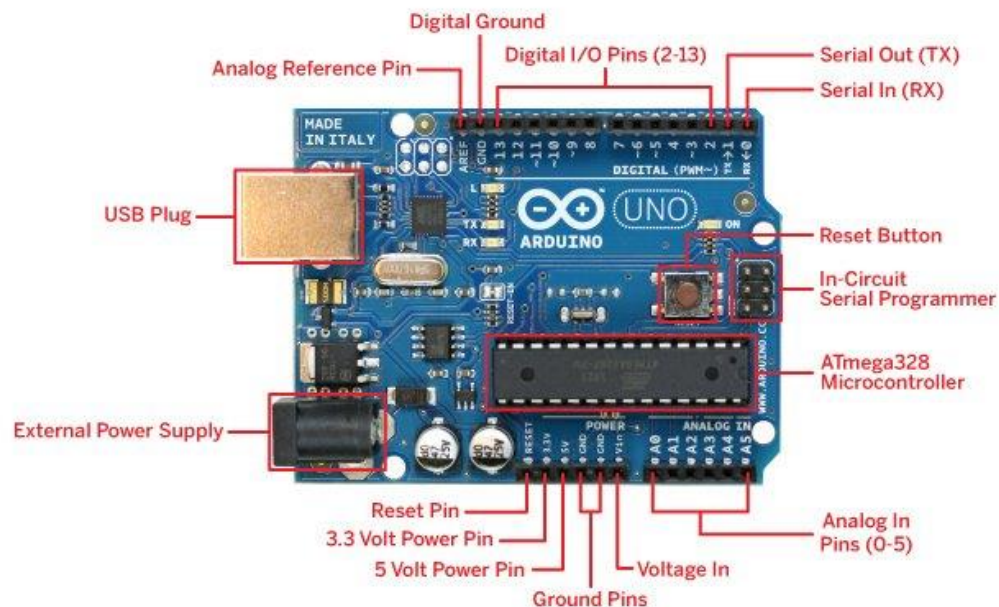


Figure 14: Arduino UNO R3

Arduino Uno: Arduino Uno is a microcontroller board that is built by an ATmega328 microcontroller. There are 14 digital input/output pins, 6 analog inputs, a USB connection, a power jack, and a reset button on this board. To get started, we can use a USB cable. An external power supply also can be used to power the Arduino Uno. An AC-to-DC adapter or a battery can provide external power. The voltage regulator may overheat and destroy the board if we use more than 12V. For Arduino UNO's the recommended range is 5 to 12 volts. The memory of an ATmega328 microcontroller includes only 32 Kilobytes.

PH Sensor: A pH sensor is made up of a probe and includes two electrodes. They are internal reference electrodes and measuring electrodes. Both of these electrodes contain potassium chloride solution with a silver chloride wire. The bulb of this glass sensing electrode is comprised of a unique glass that has been coated with silica and metal salts. The pH is measured by the concentration of hydrogen ions in this glass detecting electrode.



Figure 15: PH Sensor

Temperature Sensor: The temperature sensor used is named DS18B20 and it is a waterproof temperature sensor. It can accurately monitor temperatures in the range of -55 to +125 degrees Celsius. This sensor uses a single-wire communication protocol, so all you need to do to measure the ambient temperature is attach one wire to Arduino. This sensor's power consumption ranges from 3.3 to 5 volts. The sensor out wire, which is yellow, is linked to the Arduino Uno. The register is connected to both the yellow and the red lights.



Figure 16: Temperature Sensor DS18B20

Turbidity Sensor: Turbidity, like smoke in the air, is the cloudiness or haziness of a fluid generated by enormous numbers of small particles that are normally imperceptible to the naked eye. To measure water quality turbidity is a vital indicator. The turbidity level of water was found to be between 0.47 NTU to 6.90 NTU, according to the researchers.



Figure 17: Turbidity Sensor

Node MCU: It's an open-source IOT platform. Espressif System's ESP8266 Wi-Fi Chip on-chip and ESP-12 modulus-based hardware make up the device. The Node MCU Development Board is equipped with Wi-Fi, analog pins, digital pins, and serial communication protocols. The board is used for wireless communication since the technology has progressed to the point where vast networks of low-cost devices can track infrastructure in real-time. [1]

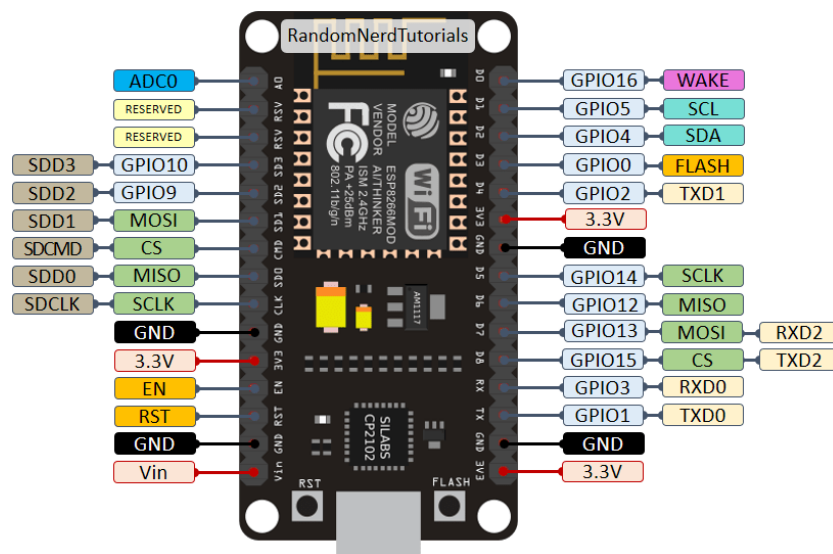


Figure 18: ESP 8266 Wi-Fi Module

There are no other physical instrumentations for this project. The other requirement for the project is a Cloud system which will develop by implementing a database management system and web design and development. For notifications, an android application system has to be made to receive the message or notification from the system. This android system will be made with the help of android studio.

3.3 Data Collection Procedure

For the project, I have to predict the future water quality data, that is why a large number of data has to be collected. At first I started to collect real data from my home filter. I have collected two months' data from the filter. Each day I have collected twenty times data. In two months I gathered total one thousand two hundred data. But for machine learning LSTM algorithm this number of data was not enough. That is why I collected some dataset from kaggle and GitHub. I have collected total 16.5 years' worthy data to train the data set. The large number of data can help to get accurate prediction.



Figure 19: Home Filter

When I collected data, I had considered only pH, Turbidity and temperature only. In the proposed model there also exist extra three parameters but for primary research I took only this three parameter to work on. PH sensor has been used to identify the presence of hydrogen ion in water.

we can determine the acidic power of water by the measure of pH level. The pH of pure drinking water is between 6.5-8. Water with lower pH becomes acidic and has a higher corrosive power. Electrical conductivity and total alkalinity have a positive relationship with pH. A lower pH can lead human beings to cancer whereas a higher pH can induce incrustation and chlorination difficulties. [5]

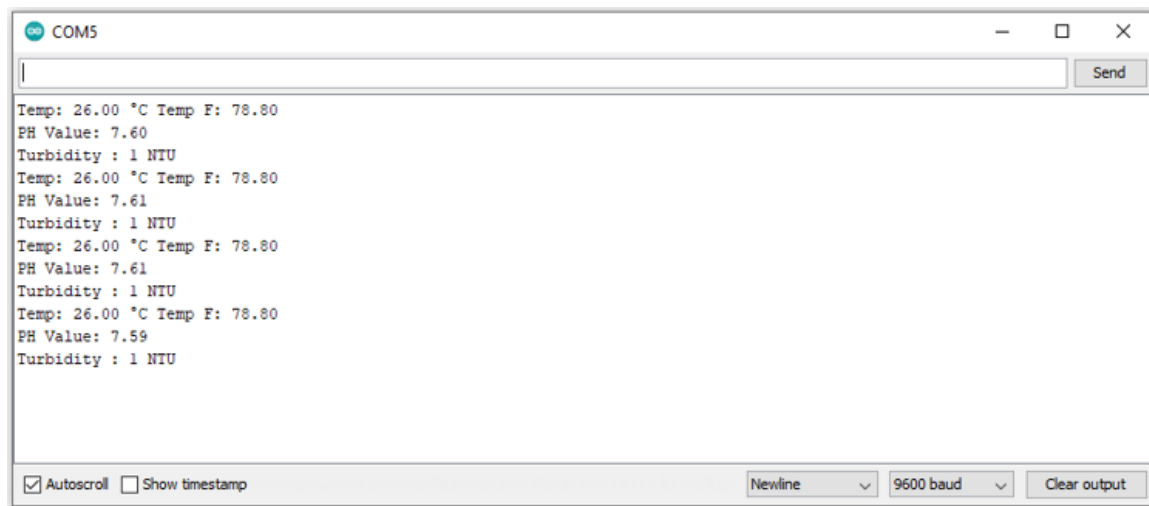


Figure 20: Data Collection by Arduino IDE

In most of the analyses from given samples, the researchers observed that the turbidity level of water is between 0 NTU and 5 NTU. Because of high turbidity, people are suffering from waterborne diseases and gastric disorders. [5]

3.4 Statistical Analysis

According's to World Health Organization, there are many water parameters that determine the quality of safe drinkable water and those parameters could be both physiochemical and microbiological. World Health Organization (WHO) estimated that roughly 80% of diseases are caused by contaminated water either directly or indirectly. Consumers are suffering from various diseases after consuming these harmful contaminants from unsafe drinking water especially children are dying in significant numbers every year. According to WHO, 1984 and APHA, 2000

they set some parameters as a global benchmark, and with the help of these parameter ranges we can recognize contaminants levels in drinking water. **[WHO PAPER]**

Parameters	Units	WHO Standards
Aluminum	mg/liter	0.1-0.2
Ammonia	mg/liter	35
Chloride	mg/liter	200-300
Chlorine	mg/liter	0.6-1.0
Choro phenols	µg/liter	0.1-2.0
Color	TCU	Below 15
Copper	mg/liter	1
Dichlorobenzene	µg/liter	0.3- 30
Ethylbenzene	µg/liter	2-130
Hardness	mg/liter	100–300
Hydrogen Sulfide	mg/liter	0.05-0.1
Manganese	mg/liter	0.3
Monochloramine	mg/liter	Below 5
Monochlorobenzene	µg/liter	10–20
pH	mg/liter	6.5-8
Sodium	mg/liter	200
Styrene	µg/liter	4-2600
Sulfate	mg/liter	250-1000
Toluene	mg/liter	40-120
Total Dissolved Solids	mg/liter	600
Turbidity	NTU	Below 5
Xylenes	µg/liter	300
Zinc	mg/liter	3–5

Table 1: WHO Authorized Water Parameters and Range

3.5 Applied Mechanism

In the rationale of the study, I have described the proposal. According to my proposal, the system will collect all required data from the given sensors. The sensor readings will pass through the Wi-Fi module in the cloud system. Now in the cloud system several steps will happen. This cloud API

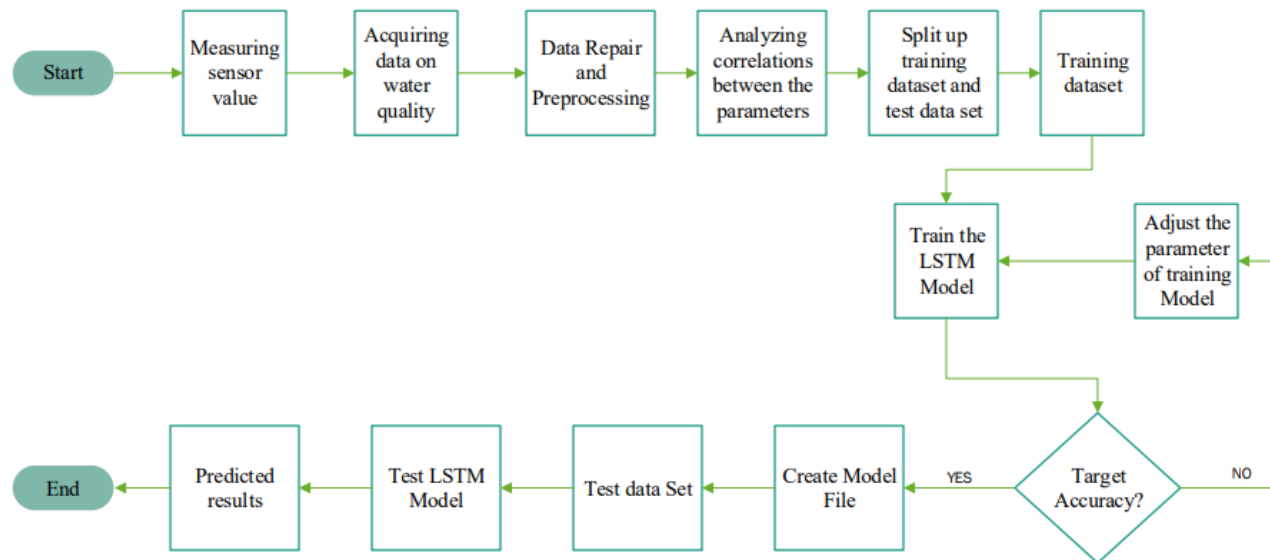


Figure 21: A Complete Data Flow Diagram On Predictive Water Quality Monitoring System

System first measures the sensor value and will obtain all water quality data in the system dataset. Then it will repair and preprocess the data set because there could be a null value in the dataset. It will analyze the relations between the parameters and changes. The dataset is divided into two pieces' train set and test set. The train set will be trained with the LSTM algorithm and will create a model file for use to predict any given data. By using this model file, it will be able to provide predictions for upcoming values in the next few days. This predicted result will be compared with the test data set. By doing a comparative analysis between the test set and train set it will show the accuracy level of its prediction. When the RMSE will low enough, we can say that the prediction capacity of the system is good and it will be effective for this project and satisfy the requirements.

3.6 Implementation Requirements

1. **C Programming:** The sensors I will use for reading water quality parameters, there will be needed c programming to instruct the compiler how to convert and read data from the sensors and how to achieve the intended data from ADC converter.
2. **Arduino IDE:** The only way to connect Arduino Uno microcontroller and sensors is Arduino IDE software. This software will help to compile the code and the UNO and Wi-Fi module devices will be connected to Arduino IDE.
3. **Neural Network:** To predict my future data the clear concept I have to have is neural network. How this neural network works and its functions in machine learning? Neural network is a computational system that work like human brain where all nodes are interconnected. It works on some algorithms which can recognize some patters and their relations of given data. After recognizing pattern and relations it continues its self-learning and pattern recognitions. This process continues repeatedly. That is why it can easily can give informative results of any given problem.

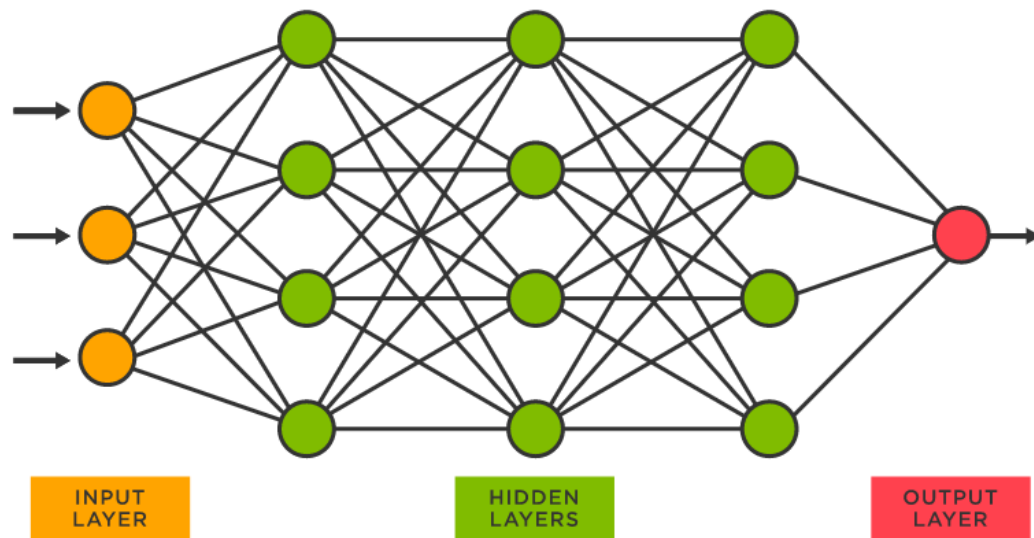


Figure 22: Neural Network

When data get into the neural network system the hidden layers work on to patter recognize, find out correlations and other relations that might be useful to solve any problem statement.

4. **Recurrent Neural Network (RNN):** In the 1980s, the recurrent neural network was created for the first time. Recurrent neural network has been made up its structure with an input layer, hidden layer, and output layer. RNNs have repeating modules in a chain-like topology, with the goal of utilizing these segments as a memory to store valuable information from earlier processing steps. A feedback loop is included in RNNs, which enable the neural network to receive a series of inputs. That means, the output from $t-1$ node is used to fed back to t node as input. As a result, Recurrent neural networks have proven to be successful at learning sequences. [6]

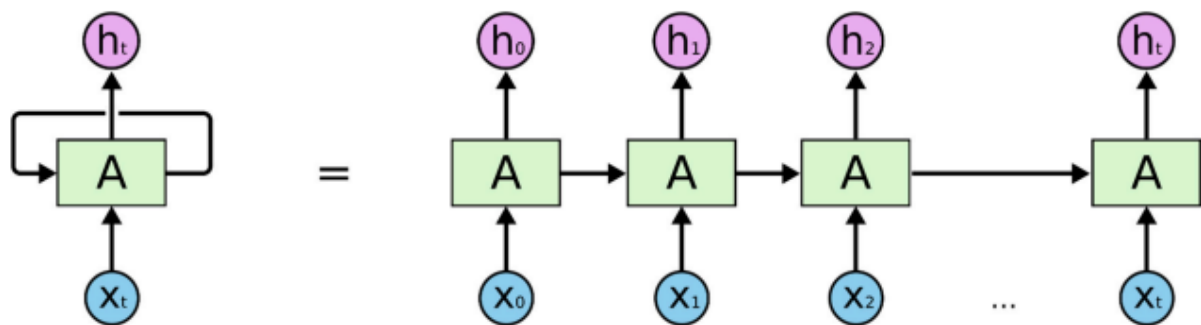


Figure 23: Recurrent Neural Network (RNN)

5. **Keras Dense Layer:** The dense layer may be a neural network layer that's connected deeply, which suggests each neuron within the dense layer receives input from all neurons of its previous layer. The dense layer is found to be the foremost commonly used layer within the models. The dense layer performs matrix-vector multiplication in the background. The values in the matrix seem to be real parameters that the classifier can be used to train and update. The dense layer produces an 'm' dimensional vector as its output. As a result, the dense layer is mostly used to modify the vector's dimensions. Dense layers also apply operations to the vector, such as scaling, rotation, and translation. There are different parameters in the dense layer. They are units, activations, use bias, initializers, regularizes, and constraints. In a dense layer, each and every single neuron can receive inputs from its previous layer neurons.

6. Long-short Term Memory: Hochreiter and Schmidhuber proposed Long Short-Term Memory, a development of RNN, to overcome the existing RNN limitations by adding more interconnections per unit. LSTMs are indeed a type of RNN by default and it can learn long-term correlations and remember information for lengthy periods of time. The LSTM model is implemented in the form of a chain. It is also the recurring module on the other hand, has a different structure. It features four interacting layers with a unique form of communication, rather than a single neural network like a normal RNN. [6]

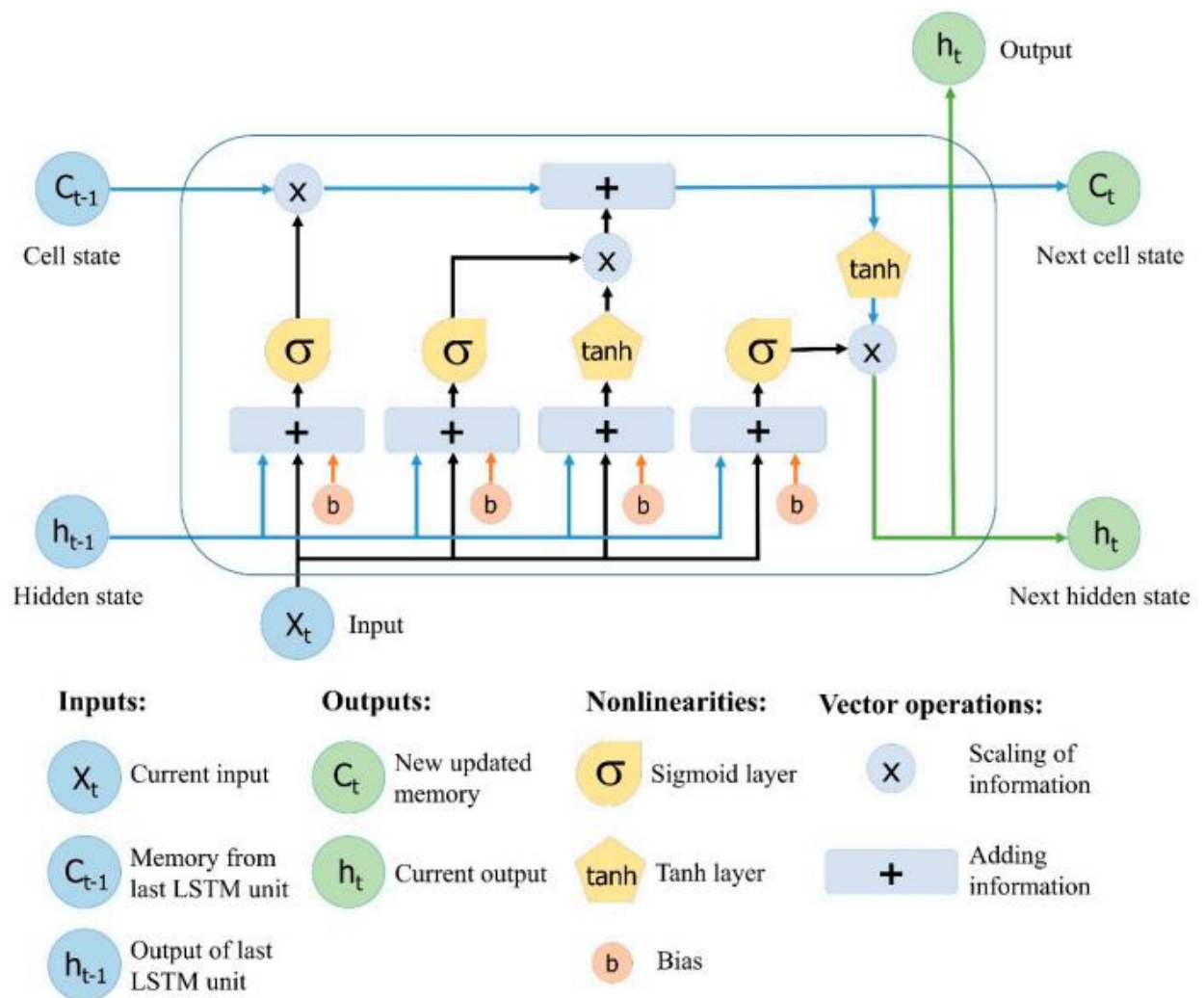


Figure 24: A Structure of Long-Short Term Memory(LSTM) Neural Network

Chapter 4: Experimental Results and Discussion

4.1 Experimental Setup

In my project primarily I have used three sensors to represent three parameters. They are pH sensor to collect hydrogen ions concentration in water, turbidity sensor for measuring the transparency cloudiness of water, and the last I have used to measure the temperature sensor to measure the water temperature in my system. All these sensors were connected to the central microcontroller Arduino Uno R3. This central unit has been used to connect all sensors for collecting water quality data. I have connected this Arduino Uno with the Arduino IDE software with a cable in my computer to run the system and this compiler helped to measure this different concentration by following some coding instructions which has been written in C programming language.



Figure 25: IOT Sensors Setup for Data Collection

The pH of pure drinking water is between 6.5-8.5. The neutral value of drinking water is 7 because it is neither acidic nor alkaline. Water with lower pH becomes acidic and has a higher corrosive power. Electrical conductivity and total alkalinity have a positive relationship with pH. A lower pH can lead human beings to cancer. Here is the circuit diagram of pH sensor and Arduino Uno

where Vcc power supply is connected to Arduino's 5V, ground pin to ground and the output pin of pH sensor is connected to Ao analog pin of Arduino.

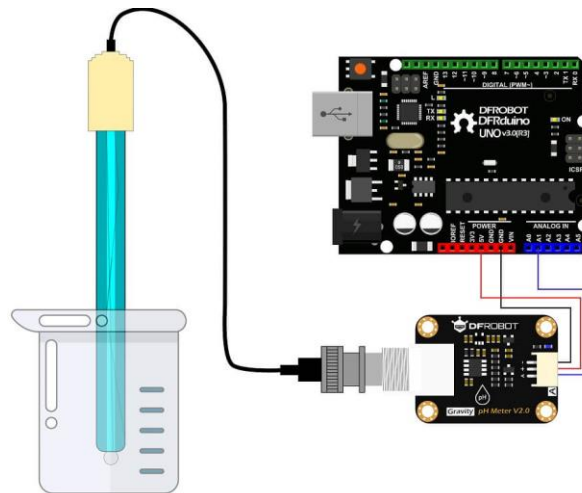


Figure 26: A Circuit Diagram of pH sensor

The temperature sensor circuit schematic is shown below. Red, Yellow, and black are the three wires existing in a DS18B20 temperature sensor. The Arduino Uno's VCC power supply is connected to the red one. This sensor's power consumption ranges from 3.3 to 5 volts. The sensor out wire, which is yellow, is linked to Arduino Uno pin number 2. Both the yellow and red wire are connected to the 4.7k register.

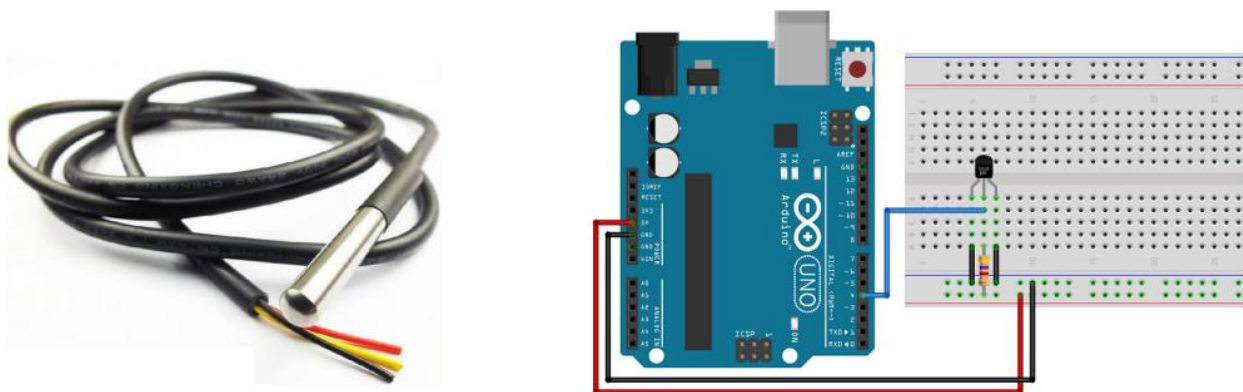
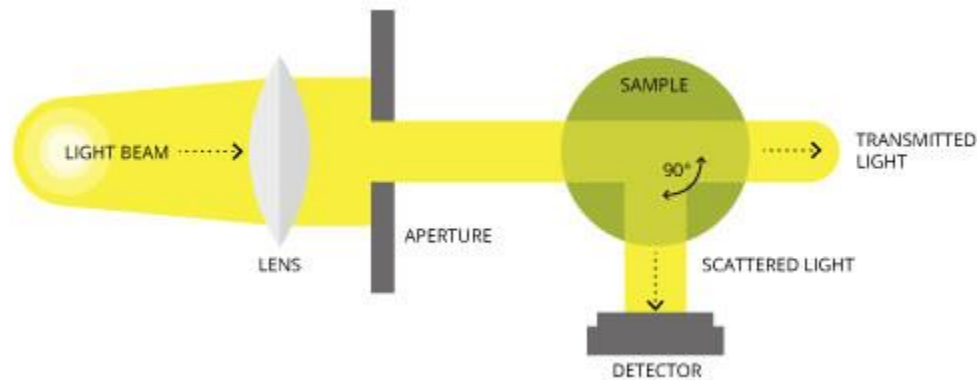


Figure 27: A Circuit Diagram of DS18B20 Temperature Sensor

The turbidity sensor works on the idea that when light passes through a sample of water, the amount of light transmitted is proportional to the amount of soil or cloudiness in the water. The amount of light transmitted reduces as the dirt or cloudiness level rises. The turbidity sensor determines the turbidity of the clean water by measuring the amount of light transmitted.



Turbidity sensor's circuit diagram has been shown below. The power of turbidity is the red one and connected to the Arduino Uno's 5V power. The ground wire is black and it is also connected to the Arduino's ground portion. The output wire is yellow and it is connected to the Analog pin A1 of Arduino Uno.

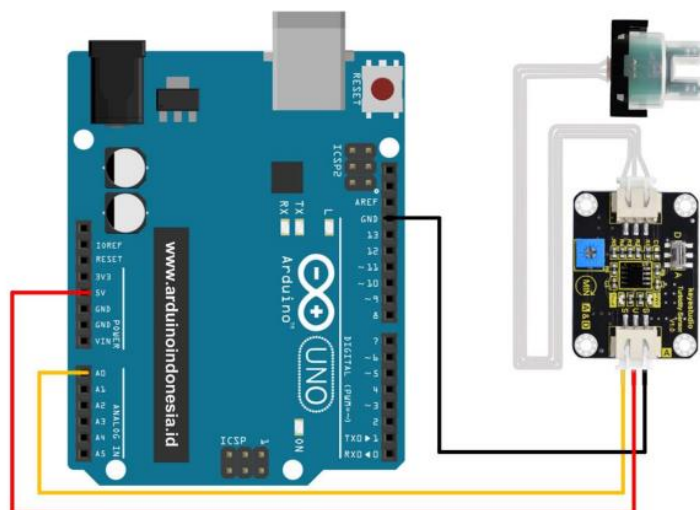


Figure 28: A Circuit Diagram of Turbidity sensor

Arduino Code: Part 1

I have collected the Arduino codes from different sources and modified them because not all codes were correct to run. Sometimes they didn't work because of missing some variables and functions. So that I have made functions and modified names with proper variable. All these codes have been taken from GitHub and YouTube descriptions. But these codes were available individually I merge them all together for three sensors of my project. That is how my project run successfully and I started to collect data through this code.



```
UNO_CODE

// for temperature sensor
#include <OneWire.h>
#include <DallasTemperature.h>
#define ONE_WIRE_BUS 2
OneWire oneWire(ONE_WIRE_BUS);
DallasTemperature sensors(&oneWire);
DeviceAddress insideThermometer;

// for ph sensor
float calibration_value = 21.34 - 0;
int phval = 0;
unsigned long int avgval;
int buffer_arr[10],temp;
float ph_act;

// for turbidity Sensor
int sensorPin = A0;
float volt;
float ntu;

void setup(void)
{
  Serial.begin(9600);
  sensors.begin();
  sensors.isParasitePowerMode();
  sensors.getAddress(insideThermometer, 0);
  printAddress(insideThermometer);
  sensors.setResolution(insideThermometer, 9);
}

void printTemperature(DeviceAddress deviceAddress)
{
  float tempC = sensors.getTempC(deviceAddress);
  if(tempC == DEVICE_DISCONNECTED_C)
  {
    return;
  }
  Serial.print("Temp: ");
  Serial.print(tempC);
  Serial.print(" °C");
  delay(1000);
  Serial.print(" Temp F: ");
  Serial.println(DallasTemperature::toFahrenheit(tempC));
  //Serial.print("F");
  delay(1000);
}

void loop(void)
{
  sensors.requestTemperatures();
  printTemperature(insideThermometer);
  for(int i=0;i<10;i++)
  {
    buffer_arr[i]=analogRead(A0);
    delay(30);
  }
}
```

Figure 29: Implemented Code

Arduino Code: Part 2

```
}
for(int i=0;i<9;i++)
{
for(int j=i+1;j<10;j++)
{
if(buffer_arr[i]>buffer_arr[j])
{
temp=buffer_arr[i];
buffer_arr[i]=buffer_arr[j];
buffer_arr[j]=temp;
}
}
}
avgval=0;
for(int i=2;i<8;i++)
avgval+=buffer_arr[i];
float volt=(float)avgval*5.0/1024/6;
ph_act = -5.70 * volt + calibration_value;

Serial.print("PH Value: ");
Serial.println(ph_act);
delay(2000);

volt = 0;
for(int i=0; i<800; i++)
{
```

```
volt = 0;
for(int i=0; i<800; i++)
{
    volt += ((float)analogRead(sensorPin)/1023)*5;
}
volt = volt/800;
volt = round_to_dp(volt,2);
if(volt < 2.5){
    ntu = 3000;
}else{
    ntu = -1120.4*square(volt)+5742.3*volt-4353.8;
}
Serial.print("Turbidity : ");
Serial.println(ntu);
delay(1000);
}

void printAddress(DeviceAddress deviceAddress)
{
    for (uint8_t i = 0; i < 8; i++)
    {
        if (deviceAddress[i] < 16) Serial.print("");
    }
}
```

Dataset: I have made a data set for my project which has been created to make a machine learning model predict the future data of water quality. To make this dataset I have taken help from Kaggle and GitHub where so many datasets are available. I mainly took turbidity parameter data from these sources. But the pH data I have taken physically from my home filter and also the remaining data I have made manually to run the model first. Because to predict future data I have to make such a model where it needs a huge number of data. It is not possible to collect a huge number of data from a single source. That is why I made this dataset manually to test and research purposes. I will apply Long-Short Term memory deep learning algorithms to find out the real pattern and the correlations between the data. I have taken the temperature data set from different sources of google. There was no exact dataset of every single day of the previous 20 years. That is why I had to search in which year and in every month in that year how much average temperature was recorded. From these results, I have made an understandable temperature dataset to readily complete the dataset. The target column has been made on the basis of water quality conditions. If any of the parameters cross its threshold value, then the target value is set as 0. When every parameter is in safe range then the target column showed as 1; that is safe to drink.

	A	B	C	D	E
1	Date	Temperature	pH	Turbidity	Target
2	1/1/2005 0:00	22.5	6.4	2.963135	0
3	1/1/2005 1:00	22.5	6.4	4.500656	0
4	1/1/2005 2:00	22.5	6.4	3.055934	0
5	1/1/2005 3:00	22.5	6.4	4.628771	0
6	1/1/2005 4:00	22.5	6.4	4.075075	0
7	1/1/2005 5:00	22.5	6.4	2.559708	0
8	1/1/2005 6:00	21	6.4	2.672989	0
9	1/1/2005 7:00	21	6.4	4.401425	0
10	1/1/2005 8:00	21	6.4	3.595017	0
11	1/1/2005 9:00	21	6.4	4.370562	0
12	1/1/2005 10:00	21	6.4	3.662292	0
13	1/1/2005 11:00	21	6.4	4.011718	0
14	1/1/2005 12:00	21	6.4	3.445756	0
15	1/1/2005 13:00	21	6.4	4.413974	0
16	1/1/2005 14:00	21	6.4	4.777382	0
17	1/1/2005 15:00	21	6.4	4.376348	0
18	1/1/2005 16:00	24	6.4	4.268429	0
19	1/1/2005 17:00	19.5	6.4	4.503661	0
20	1/1/2005 18:00	19.5	6.4	2.436086	0
21	1/1/2005 19:00	19.5	6.4	2.560299	0
22	1/1/2005 20:00	19.5	6.4	4.133423	0
23	1/1/2005 21:00	19.5	6.4	3.749701	0

	A	B	C	D	E
145523	8/8/2021 9:00	21.5	8.12	3.543696	0
145524	8/8/2021 10:00	21.5	8.12	4.673548	0
145525	8/8/2021 11:00	21.5	8.12	3.622379	0
145526	8/8/2021 12:00	21.5	8.12	3.821456	0
145527	8/8/2021 13:00	21.5	8.12	3.226349	0
145528	8/8/2021 14:00	21.5	8.12	3.860084	0
145529	8/8/2021 15:00	21.5	8.12	3.045612	0
145530	8/8/2021 16:00	21.5	8.12	4.364426	0
145531	8/8/2021 17:00	21.5	8.12	3.824734	0
145532	8/8/2021 18:00	21.5	8.12	4.921196	0
145533	8/8/2021 19:00	21.5	8.12	3.95053	0
145534	8/8/2021 20:00	21.5	8.12	3.765694	0
145535	8/8/2021 21:00	21.5	8.12	3.691947	0
145536	8/8/2021 22:00	21.5	8.12	3.399726	0
145537	8/8/2021 23:00	21.5	8.12	4.683335	0
145538	8/9/2021 0:00	21.5	8.12	3.383049	0
145539	8/9/2021 1:00	21.5	8.12	4.903553	0
145540	8/9/2021 2:00	21.5	8.12	3.757799	0
145541	8/9/2021 3:00	21.5	8.12	3.455045	0
145542	8/9/2021 4:00	21.5	8.12	3.958543	0
145543	8/9/2021 5:00	21.5	8.12	4.362792	0
145544	8/9/2021 6:00	21.5	8.12	2.593834	0

Figure 30: Water Quality Dataset 2005 – 2021

4.2 Experimental Results & Analysis

I have used Jupyter Notebook for analysis my dataset and also it will be used to train this data set and will be testify with test dataset. There I tried to analysis my data set from various aspects.

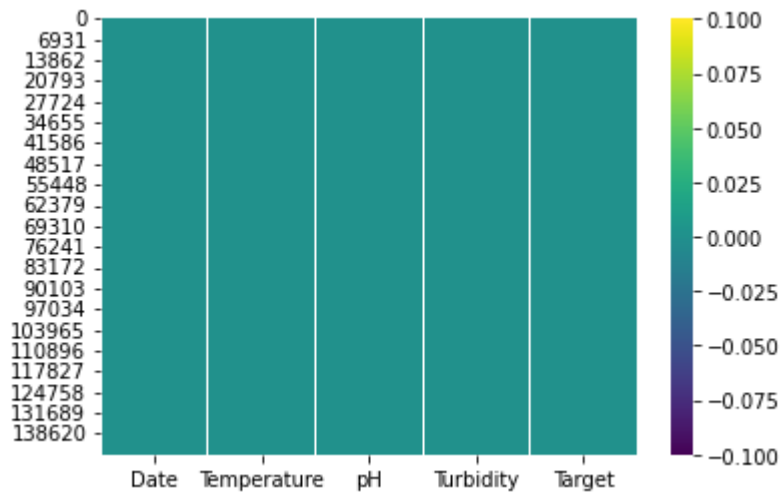


Figure 31: Finding Null Value

In my given dataset, total one lakh forty-five thousands five hundred forty-three (1,45,543) has been used and it is representing total 16.5 years' data. From the hit map I didn't found any null value in the dataset.

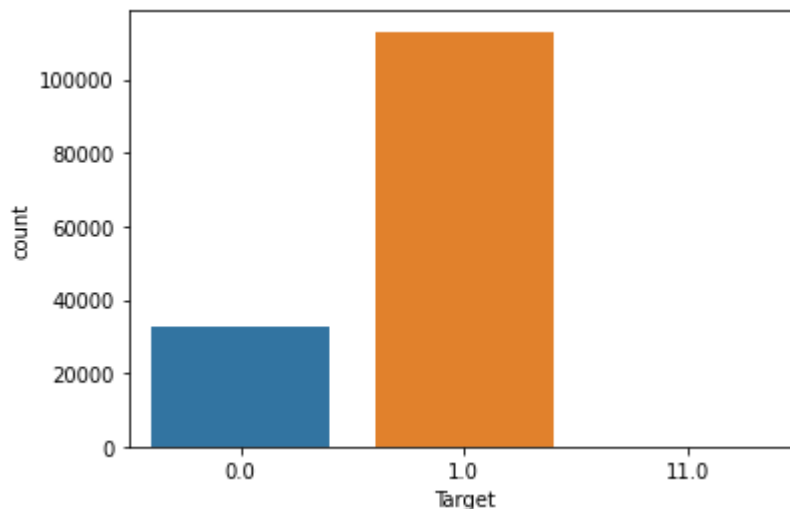


Figure 32: Comparing Target Water Quality with Total Count

By plotting all data, it found more than one lakh times that the water quality is fine to drink and more than thirty thousand time it is showing that, the water quality is not in a drinkable condition. The mean value of temperature, pH and turbidity is 23.57, 7.14, and 3.96. The standard deviation of each parameter are 2.689873, 0.561964, and 0.779953. The minimum value of temperature is

	Temperature	pH	Turbidity
count	145543.000000	145543.000000	145543.000000
mean	23.574027	7.139604	3.967637
std	2.689873	0.561964	0.779953
min	17.000000	6.100000	1.450000
25%	22.000000	6.630000	3.440300
50%	24.000000	6.980000	3.955091
75%	25.500000	7.650000	4.501025
max	29.500000	8.340000	6.739000

Figure 33: Data analysis by mean, max and min

17.0 and maximum value is 29.5. The minimum and maximum value of pH level was 6.1 and 8.34. The maximum and minimum value of turbidity was 6.74 and 1.45. The total counted value all of them was 145543.

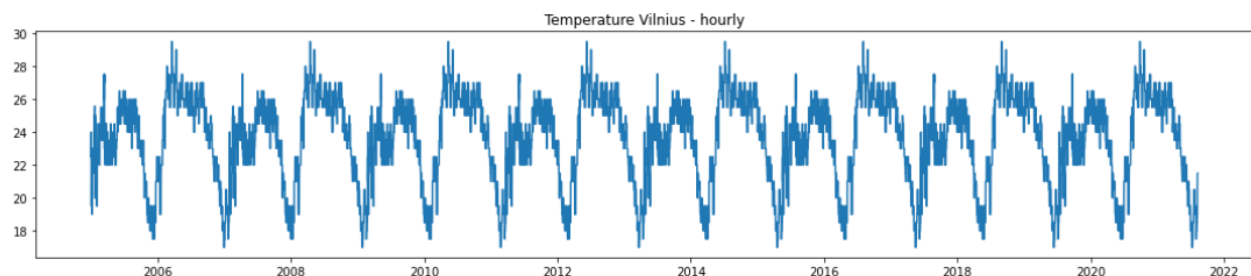


Figure 34: Hourly Temperature Condition in 2005-2021

From the research the condition of hour temperature has been found and it easily understandable form the graph.

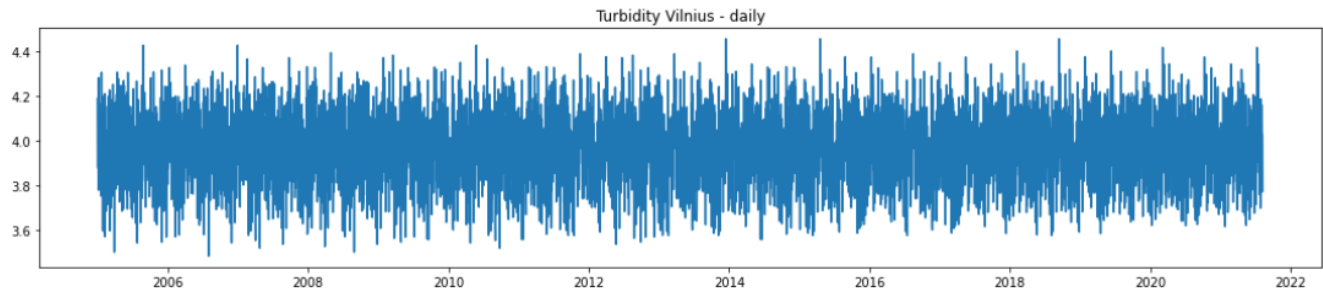


Figure 35: Daily Turbidity Condition in 2005-2021

The daily turbidity condition has been found from the dataset analysis. This graph showing that the limit of turbidity was in between 3 to 4.5 all these years. The standard value turbidity is 0-5. A Nephelometric Turbidity Unit is the measurement unit (NTU). Low turbidity readings imply clear water, while high turbidity values suggest poor water quality.

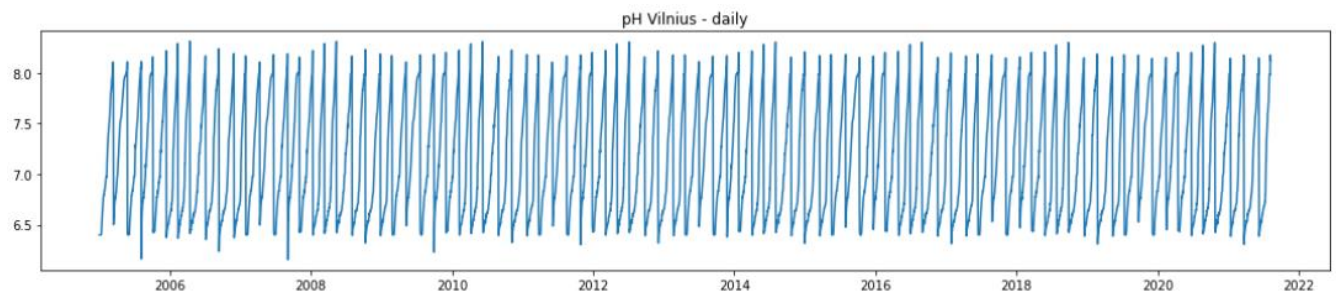


Figure 36: Daily pH Condition in 2005-2021

Pure drinking water has a pH of 6.5 to 8. Because it is neither acidic nor alkaline, drinking water has a neutral value of 7. Water with a lower pH level becomes acidic and more corrosive. The daily pH condition has been found from the dataset analysis. This graph showing that the limit of pH was in between 6 to above 8 all these years. Water that isn't in the appropriate pH range of 6.5 to 8, especially if it's alkaline, isn't always dangerous. Very alkaline water, on the other hand, might have a disagreeable taste or smell, as well as damage pipes and water-carrying appliances. Freshwater with a pH of less than 6.5 is more likely to be toxic, making it unsafe to drink. Metal pipes can also be damaged by it.

Dependencies and Correlation of pH and Turbidity with Temperature:

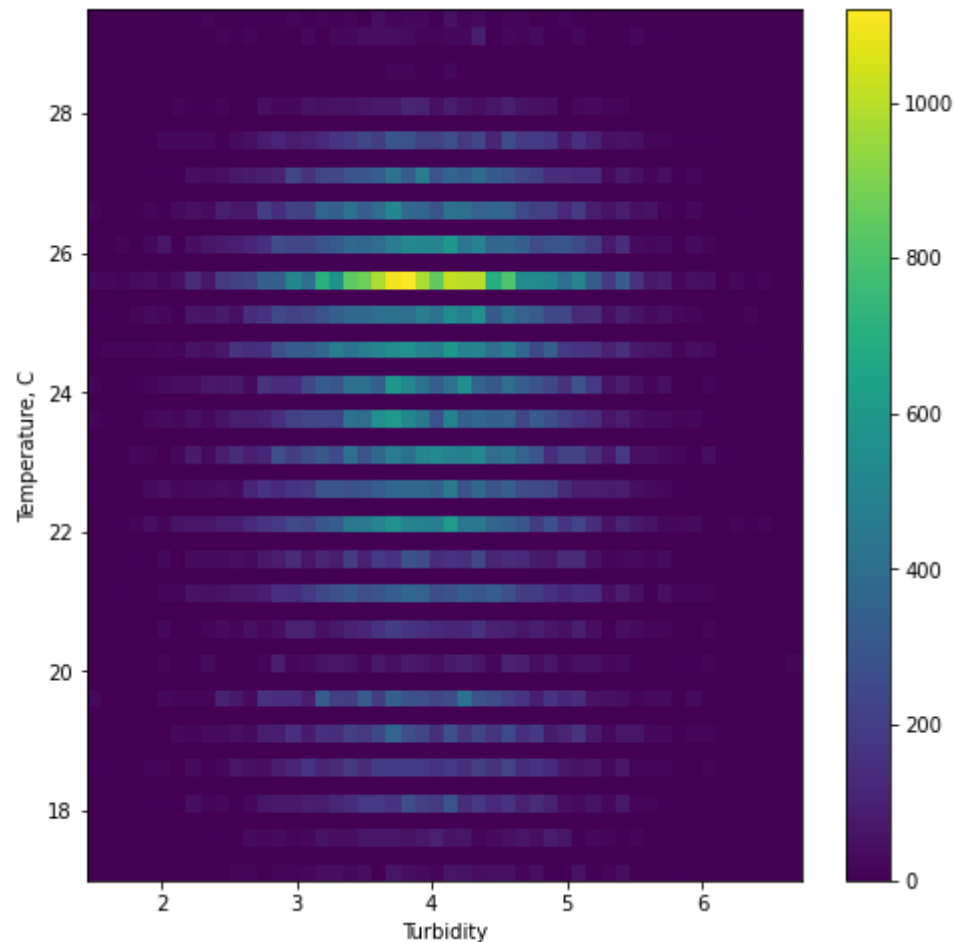


Figure 37: Correlations of Turbidity and Temperature

This plotting shows that the presence of turbidity entities is different in different temperature conditions. The entities have been showing in between 0- 1000. It is showing that the cloudiness of water is low in low temperature when the rate is 18 degrees Celsius to 22 degrees Celsius. But the cloudiness of water is increasing with temperature simultaneously. This plotting shows that the transparency becomes low or cloudiness becomes high when the temperature ranges between 24-28 degrees Celsius.

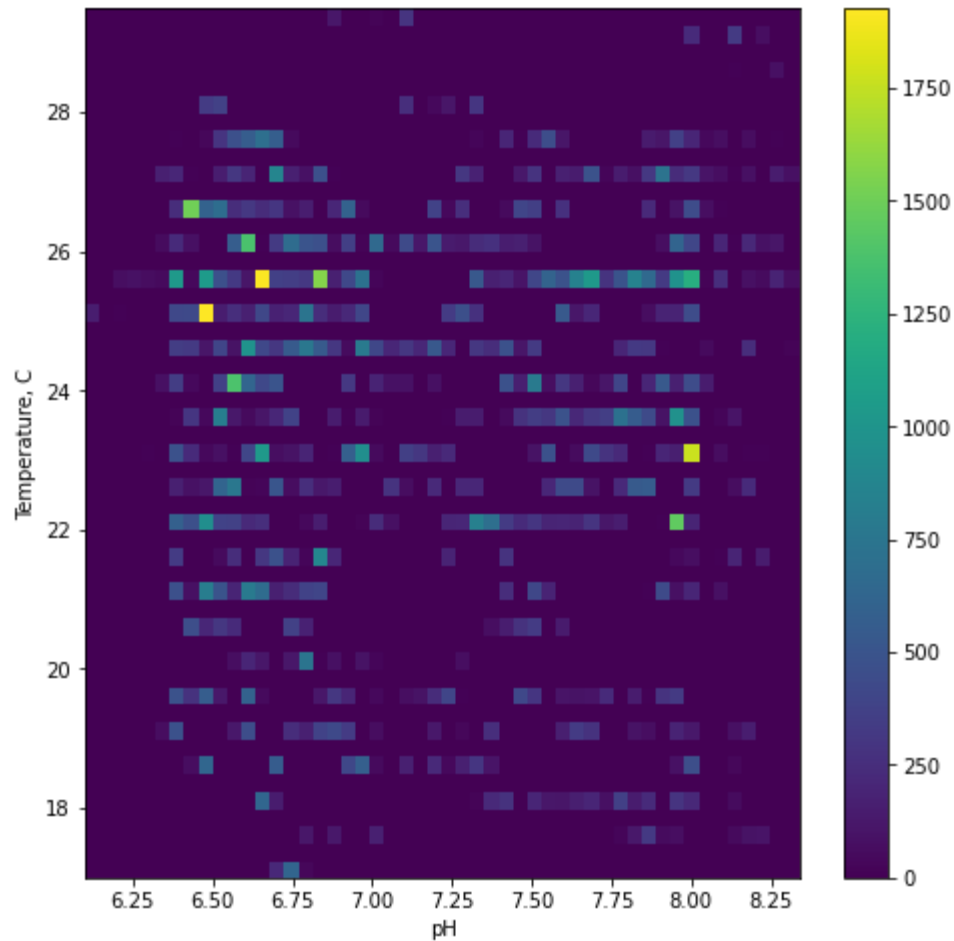


Figure 38: Correlation of pH and Temperature

pH means the presence of hydrogen ions and hydroxide ions in water. This plotting it is showing that in the range of 18 degrees to 24 degrees' Celsius temperature the concentration of hydrogen ion is the literally same condition and the range is 7-7.25. The neutral range is 7 for drinking water. From this plotting, it is clearly showing that after increasing temperature from 24 to 26 the concentration of hydrogen ion literally is not in stable condition. It is incredibly showing its variation in high temperature. In some research, it is proved that the concentration of hydrogen ions in pure water will be stable and the same. So from the analysis on my dataset, it is also showing a stable concentration of hydrogen ions in the range of 7 to 7.25. Also, the World health organization has been said the standard range is 7 for drinking. My research also proves that is information is authentic.

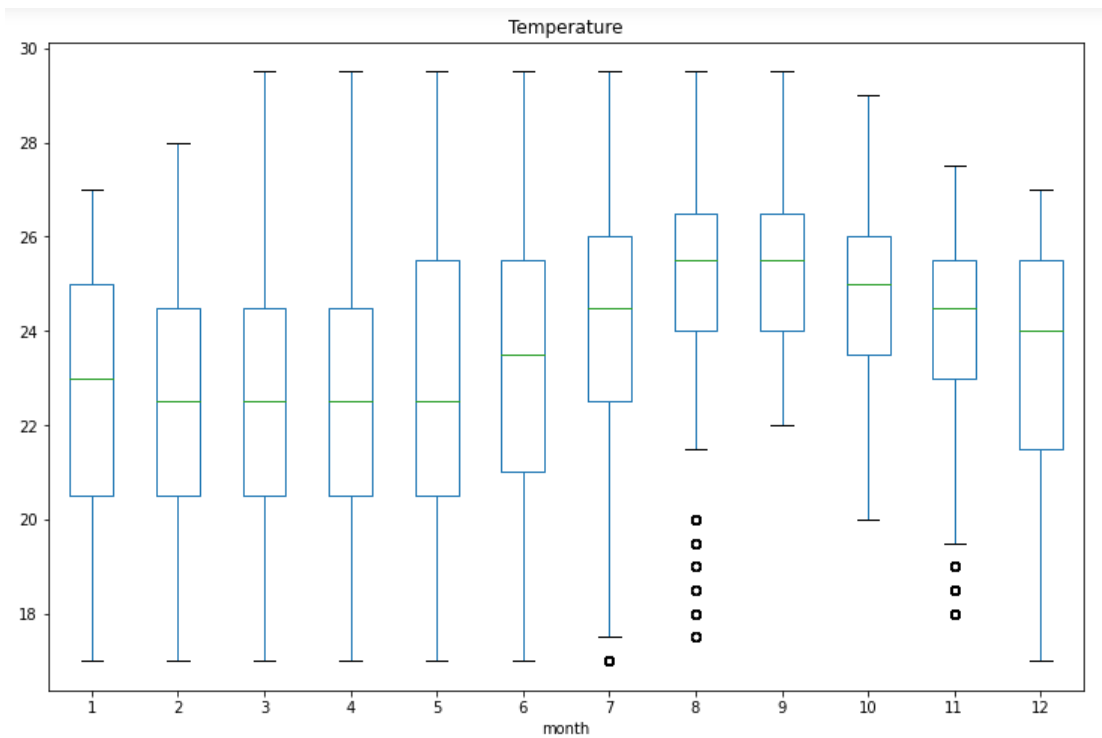


Figure 39: Box Plot of Temperature in Month

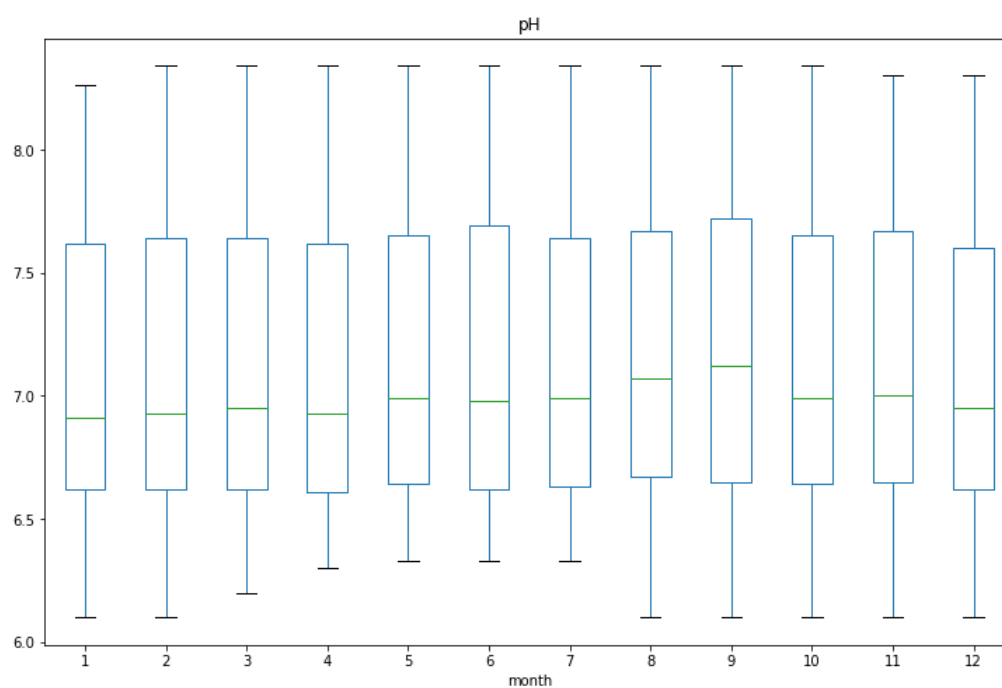


Figure 40: Box Plot of pH in Month

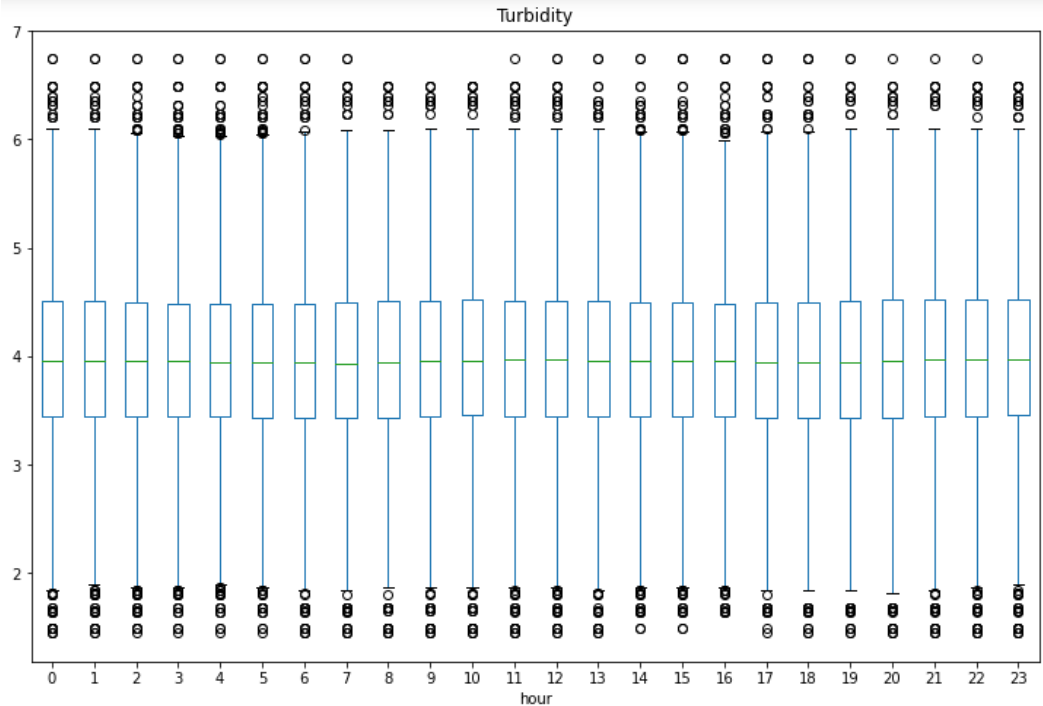


Figure 41: Box Plot of turbidity in Hour

The box plots in Figures 39, 40, and 41 are showing the monthly and hourly conditions of temperature, pH, and turbidity in all these years. In figure 39 the box plot temperature shows their average range in 12 months of every year. The range was in a safe condition. This is also showing that in months 8 and 11 there is no continuous temperature in the range of 18-22 degrees Celsius. Figure 40 also shows the monthly pH condition every year. The pH range was mostly in the safe range above 6.5 and 7. The safe range is 7 for drinking. The box plotting for turbidity in figure 41 has been shown on a daily hour basis for each year and month. This plotting shows that the turbidity of the given dataset was on 4 a regular basis. There is no continuous value in the range of 0 to 2 and also above 6. Though the safe range of turbidity is always 0-5 NTU. This analysis has been showing that the range of turbidity was in good condition all these years in maximum.

Water Quality Prediction Model by using Long-Short Term Memory:

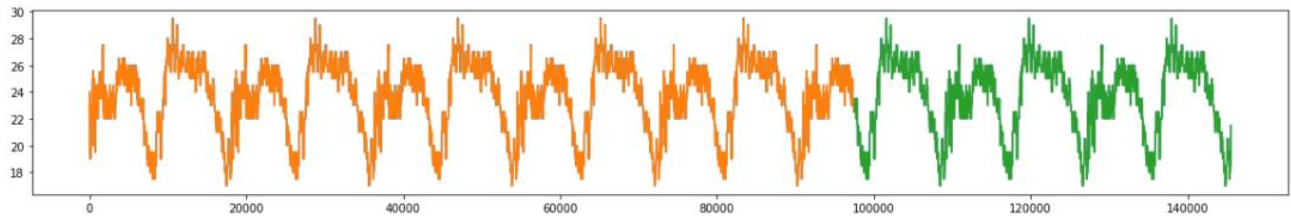


Figure 42: Predicting Temperature by LSTM

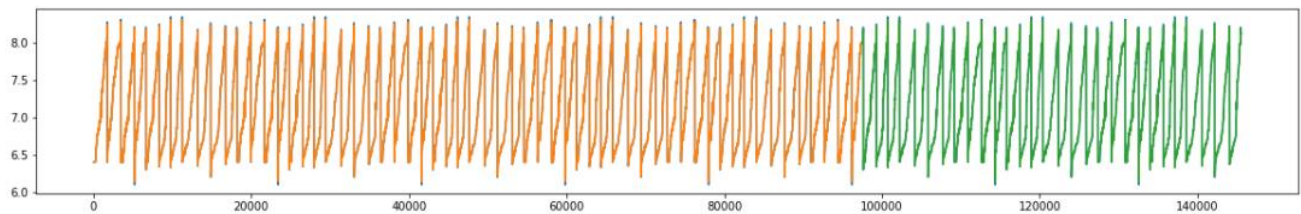


Figure 43: Predicting pH by LSTM

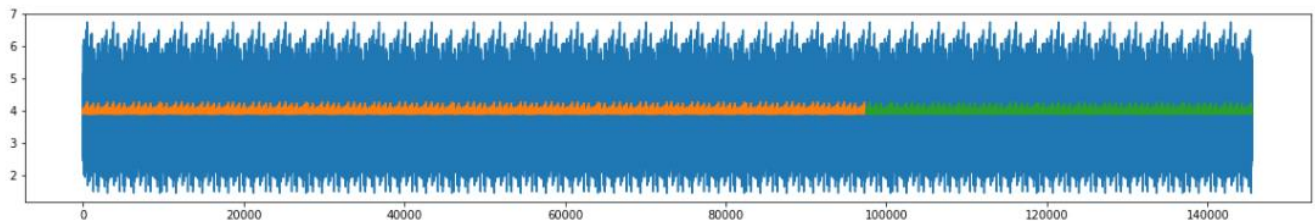


Figure 44: Predicting Turbidity by LSTM

A neural network algorithm LSTM has been used to predict the future water quality data accurately. LSTM can predict better than its previous version recurrent neural network (RNN). In this system, I have trained the dataset with LSTM. The train score for temperature, pH and turbidity is 0.18 RMSE, 0.05 RMSE, 0.78 RMSE. The Test score for temperature, pH and turbidity is 0.17 RMSE, 0.05 RMSE, 0.78 RMSE. This model indicates that their prediction level is so accurate and these models can be used effectively.

Chapter 5: Impact on Society, Environment and Sustainability

5.1 Impact on Society

The impact of safe drinking water on our society is huge from the aspect of children's health. So that my smart filter can contribute a large portion to save the children's health from dangerous water-borne diseases. In Bangladesh, the access to safe drinking water is only 34.6% according to UNICEF. So still it is too low in many areas. My smart filter will be able to provide help to children from these areas where safe drinking water is not guaranteed.



Every year, thousands of children die because of a direct result of water-borne sickness. When children without access to safe water drink from contaminated sources, diseases like dysentery, cholera, polio, hepatitis A, diarrhea, and typhoid can easily be transmitted. However, one of the most important effects of this smart system will be, it will provide and ensure the supply of clean water that aids in the prevention of water-borne diseases. When school children start to drink from that smart filter, it will become trustworthy to everyone. The children will motivate their parents to have this filter in their home also. That is how this can contribute to a large portion of Bangladesh. This will also be cost-effective for every middle-class family to afford.

5.2 Impact on Environment

This Smart Filter can contribute to the environment very effectively. From research, it has been said that a plastic bottle takes approximately 450 years to decompose. In this long period of time, it harms our environment in many ways. Most of the time's schoolchildren carry water bottles for drinking purposes with them when they come out from the home to attend school. They throw away these plastic bottles here and thereafter use this. This becomes a large number of bottles after a few days when ultimately it is impossible to count the number of these plastics. These plastics affect our environment very easily in our country because the garbage collection system in our society has not been grown well and it's easily affecting our ecosystem.



This smart filter will reduce the percentage of carrying water bottles of students while they arrive at school. This smart filter will become a trustworthy system when they start to drink fresh and safe water continuously. It will demotivate to carry the tendency of plastic water bottles because they will be able to drink fresh and clean water from this smart filter without investing any money. When the trend of bringing water bottles reduce, the environment will not affect because of these plastics anymore. The eco system of our country will not affect anymore and the water bottles will not move around the world. So that the environment can breathe and won't change significantly.

5.3 Ethical Aspects

Since this smart system was created to raise awareness among our country's citizens about the dangers of unclean drinking water and to protect our country's children. As a result, conducting any type of business using this system is completely forbidden. Sharing any incorrect information about this smart system can cause significant harm to others, so the school administration should take precautions when disclosing any information regarding water quality to the public. The authority has to be concerned about the filter kit change while the system will notify them to change the filter otherwise it can cause serious health issues for this school students.

5.4 Sustainability Plan

In the long run, sustainability is constancy. It's not uncommon for a system to become obsolete; no one uses that smart filter any longer. These are referred to as "unsustainability." Many electric smart systems are still giving the same service as they were when they were first installed. This is referred to as a sustainable system. As this smart filter was designed for the benefit of a society's children, it will have to maintain the same level of service for an extended period of time. Because this smart technology will be utilized to safeguard the safety of drinking water and the well-being of youngsters. This technology can provide safe drinking water as long as there is a need. If this system cannot detect the quality of water it will not be useful any longer without any doubt.

Chapter 6: Conclusion and Implication for Future Research

6.1 Summary of the Study and Conclusions

- For research the main component is data. For data collection, I collected all IOT materials including Arduino Uno R3, Wi-Fi module ESP 8266, PH Sensor, Turbidity Sensor, and Temperature Sensor Ds18b20.
- To run all these devices together I have connected all these IOT devices with Arduino Uno and Arduino Uno taking instructions from the Arduino IDE software. I have installed this Arduino IDE software on my PC and connected it with Arduino UNO by a cable connection.
- After Connection, I have made a dataset in a CSV file. Then I started to collect data every hour. Sometimes I collected 2 or 3 times in an hour, and that is how I collected 1400 data manually in two months. But for proper research, I needed a large amount of data. That is why I collected different datasets from Kaggle and GitHub resources. Then I prepared a proper dataset of 16 years to testify my research properly.
- After preparing my dataset, I started my study and research on machine learning. I prepared my dataset and preprocessed all data before training it. I choose a neural network algorithm named Long-Short Term memory to research my topic. I need to predict the future data of water quality, that is why this neural network model has been used to predict the upcoming data from my smart filter.
- I have trained three different models with LSTM for pH, temperature, and turbidity individually. Because LSTM cannot train all parameters together, that is why I trained my parameters individually.
- The Models I have created by using LSTM algorithm, they are giving an accurate level of prediction. I have divided my data set into train dataset and test dataset. After the prediction, I tested my predicted data with test dataset. The test score is for temperature, pH and turbidity is 0.17, 0.05 and 0.78. The difference between train data and test data is less than one for each model, which is comparatively good.
- Finally, I can say that, the model I have created it can predict the future data of my smart filter to save thousands of school children's life from water borne diseases.

6.2 Implication for Further Study

- Since I started my work on this project I have faced very challenging situations because all this work I have done alone and I had no previous experience in this field. Most of the time the devices didn't work properly and also I didn't find out any proper instruction on the IOT devices that is why it took time to calibrate all the devices.
- I have made this project where I didn't add the cloud system yet which was included in my project proposal. But I have planned to continue my work on this. I will complete my cloud system for storing datasets and implementing machine learning in this portion.
- When the cloud system will add, my smart system will become a fully automated system where the data will be sent to the cloud system, and from the cloud, the users will get notification when the water quality gets down.
- I applied an LSTM algorithm for predicting. It took so long times for individual three sets of training. I didn't create a model where the system can detect the threshold value from the predictive data. It will take time to create another model for this function, for the time limitations I could not complete this. But I am continuing this work, I will complete another model to catch the threshold value from the prediction.
- After completing these steps, I have the intention to write a paper and publish my work. This fully automated system can contribute undoubtedly a big role to save children's life from dangerous water-borne diseases. This smart system will also fulfill the 6th goal of Sustainable Development Goals (SDG). So I am hopeful to receive a governmental donation to make it successfully and I want to deliver this smart filter all over the area of our country.

References

1. Sathish Pasika, Saj Teja Gandla, “Smart water quality monitoring system with cost-effective using IOT”, Heliyon 6, 2020.
2. Varsha, Anjitha, Akshay, Aruna, Jagadeesh, Audre, “IoT based smart water quality monitoring system”, Global Transitions Proceedings, 2021.
3. Geetha, Gouthami, “Internet of things enabled real time water quality monitoring system”, Springer Open, 2017.
4. Prasad, Mamun, Islam, Haqva, “Smart Water Quality Monitoring System”, Research Gate, 2015.
5. Roohi, A.R. Siddiqui, “Assessment of Physiochemical Characteristics of Drinking Water Quality in Allahabad Metropolitan City, India”, The Oriental Anthropologist, 2019.
6. Xuan, Sungho, Giha, “Application of Long-short Term Memory(LSTM) Neural Network for Food Forecasting”, research gate, July 2019.
7. Prashant, Vamshi, Ramakrishna, “IOT Based Smart Water Quality Monitoring and Prediction System”, IJEAT, 2019.
8. Anantha, Dr. Geetha, “IOT based Real-Time Water Quality Monitoring System using smart Sensors”, IRJET, 2020.
9. Monitoring Using Machine Learning Algorithm”, European Journal of Molecular & Clinical Medicine, 2020.
10. Abhinandhan, Kuldeepak, “Water Quality Monitoring using IOT and Machine Learning”, Abhinandhan E Journal of Engineering Research and Application, 2019.
11. Vaishnavi, Daigavane, Gaikwad, “Water Quality Monitoring System Based on IOT”, Research India Publications, 2017.
12. Chetna, Shreya, Harshal, Nitin, “Advanced Water Quality Monitoring with IoT”, IRJET, 2020.
13. Brinda, Jain, “Real-Time Water Quality Monitoring System using Internet of Things”, International Conference on Computer, Communications and Electronics, 2017.
14. M Cho, G Lenin, L Huo, Prassana, “Real-time water quality system in internet of things”, The Engineering Society, 2020.

15. Alexander, Demetillo, Michelle, Japitana, “A system for monitoring water quality in a large aquatic area using wireless sensor network technology”, Sustainable Environment Research, 2019.
16. Mohammad, Talha, Subhasish, Abhijit, Manjur, Nurul, Karl, Shahadat, “IoT Based Real-time River Water Quality Monitoring System”, Elsevier, 2019.
17. Senthilkumar, Priyadharshini, Sangeetha, Sathiya, Sanjana, "IoT based Filter Testing and Monitoring", ICEECT, 2020.
18. Soundarya, Prathyusha, Niharika, Karthick, Suresh, "Intelligent IoT Based Water Quality Monitoring System", International Journal of Applied Engineering Research, 2017.
19. Pooja, Siddharth, Sanjay, "ASSESSMENT OF WATER QUALITY", International Journal of Scientific & Engineering Research, 2014.
20. Tahera, Fatema, Fahmida, Tridib, Sabuj, Digbijoy, Milan, Md Akramul, Mahfuzar, "Water Quality Index for measuring drinking water quality in rural Bangladesh: a cross-sectional study" Journal of Health Population and Nutrition, 2016.
21. J Lu, F Yuan, "The effect of drinking water quality on the health and longevity of people-A case study in Mayang, Hunan Province, China”, International Conference on Water Resource and Environment, 2017.
22. Taruna, Alankrita, "Assessment of water quality and its effects on the health of residents of Jhunjhunu district, Rajasthan: A cross-sectional study", Journal of Public Health and Epidemiology, 2013.
23. Yirdaw, Bamlaku, "Drinking water quality assessment and its effects on residents’ health in Wondo genet campus, Ethiopia", Environmental system research, 2016.
24. Peiyue, Jianhua, "Drinking Water Quality and Public Health", Exposure and Health, 2019.
25. Azza, Al-Khatib, Maher, "Quality of Drinking Water from Springs in Palestine: West Bank as a Case Study", Journal of Environmental and Public Health, 2019.
26. Daud, Nafees, Shafaqat, Rizwan, Raees, Shakoor, Arshad, Shahzad, Deebe, Murad, Malook, Shui, Drinking Water Quality Status and Contamination in Pakistan, BioMed Research International, 2017.
27. Patrick, Cristina, Villanueva, "Drinking Water Quality and Human Health: An Editorial", International Journal of Environmental and Public Health, 2019.

Appendix

Md.Shafiuzzaman Khan_181014055

ORIGINALITY REPORT

10%

SIMILARITY INDEX

5%

INTERNET SOURCES

5%

PUBLICATIONS

6%

STUDENT PAPERS

PRIMARY SOURCES

1

[dspace.daffodilvarsity.edu.bd:8080](https://dspace.daffodilvarsity.edu.bd/8080)

Internet Source

1%

2

Roohi Rawat, A. R. Siddiqui. "Assessment of Physiochemical Characteristics of Drinking Water Quality in Allahabad Metropolitan City, India", The Oriental Anthropologist: A Bi-annual International Journal of the Science of Man, 2019

Publication

1%

3

Submitted to University of Ulster

Student Paper

1%

4

Submitted to Southern New Hampshire University - Continuing Education

Student Paper

1%

46	ebin.pub Internet Source	<1 %
----	---	------

47	emmadonnan.org Internet Source	<1 %
----	---	------

48	tukl.seecs.nust.edu.pk Internet Source	<1 %
----	---	------

49	www.bgsu.edu Internet Source	<1 %
----	---	------

50	"Neural Information Processing", Springer Science and Business Media LLC, 2017 Publication	<1 %
----	---	------

51	S. Geetha, S. Gouthami. "Internet of things enabled real time water quality monitoring system", Smart Water, 2017 Publication	<1 %
----	--	------

52	Alexander T. Demetillo, Michelle V. Japitana, Evelyn B. Taboada. "A system for monitoring water quality in a large aquatic area using wireless sensor network technology", Sustainable Environment Research, 2019 Publication	<1 %
----	--	------
