**Water quality estimation from the given dataset (IOT approach)**

## **Abstract**

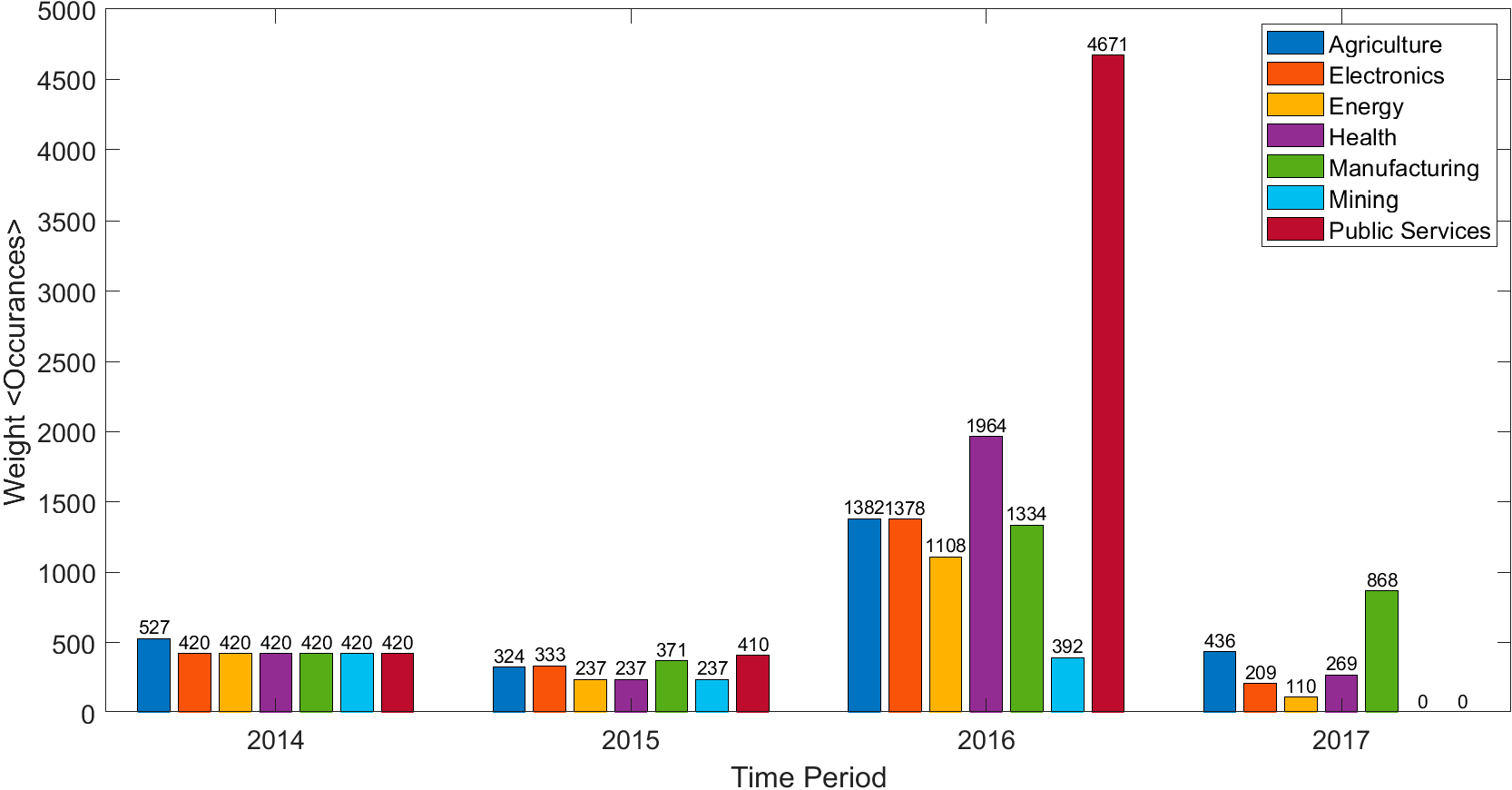
Water is a quintessential element for the survival of mankind. Its variety of uses means that it is always in a constant state of demand. The supply of water most primarily comes from large reservoirs of water such as lakes, streams, and the ocean itself. As such, it is good practice to monitor its quality to ensure it is fit for human consumption. Current water quality monitoring is often carried out in traditional labs but is time consuming and prone to inaccuracies. Therefore, this paper aims to investigate the feasibility of implementing an Arduino-based sensor system for water quality monitoring. A simple prototype consisting of a microcontroller and multiple attached sensors was employed to conduct weekly onsite tests at multiple daily intervals. It was found that the system works reliably but is reliant on human assistance and prone to data inaccuracies. The system however, provides a solid foundation for future expansion works of the same category to elevate the system to being Internet of Things (IoT) friendly.

## 1. Introduction

The Internet of Things, otherwise known as IoT in the simplest sense, refers to the concept of connecting physical devices, machines, software, and objects to the Internet [[**1**](https://www.mdpi.com/2076-3298/8/1/6#B1-environments-08-00006)]. In a broader sense, it is a dynamic and global network infrastructure, in which intelligent objects and entities are used in conjunction with actuators, electronics, sensors, software and connectivity to enhance connection, collection and data exchange [[**2**](https://www.mdpi.com/2076-3298/8/1/6#B2-environments-08-00006)]. This type of network generally has a large number of nodes that interact with the environment and exchange data, whilst reacting to events or triggering actions to exert control or change upon the physical world. By sharing and acting on shared data contributed by individual parts, an IoT system would be greater than the sum of its parts [[**3**](https://www.mdpi.com/2076-3298/8/1/6#B3-environments-08-00006)]. Each network node is considered smart and consumes little resources such as data processing and data storage power as well as energy consumption.

The term Internet of Things was initially coined in 1999 by Kevin Ashton, an expert in digital innovation [[**4**](https://www.mdpi.com/2076-3298/8/1/6#B4-environments-08-00006)]. Since then, there has been a significant growth and development in the IoT industry because IoT provides a platform that creates opportunities for people to connect devices and control them with big data technology. [**Figure 1**](https://www.mdpi.com/2076-3298/8/1/6#fig_body_display_environments-08-00006-f001) below shows the seven industries that are mainly affected by the growth of IoT over the period of late 2014 to early 2017, indicated by the weight which represents the occurrences of investment. The industries affected are namely: Manufacturing, Agriculture, Public Service, Health, Electronics, Energy, and Mining [[**5**](https://www.mdpi.com/2076-3298/8/1/6#B5-environments-08-00006)]. IoT integration into manufacturing operations have been repeatedly emphasized by governments using the term Industrial IoT (IIoT) to produce fully intelligent, connected and autonomous manufacturing plants.

Agriculture benefits from IoT’s real-time operation for optimizing productivity at reduced costs. A recent study by Cipolla et al. (2019) shows that an IoT system can be used to monitor soil moisture and electrical conductivity, as well as surface and groundwater electrical conductivity, to optimize the management of both irrigation and drainage system [[6](#_bookmark24)]. Installation and moni- toring of sensor devices in public services allowed for intelligent transportation system and traffic management, optimal water and electricity management. A water flow driven sensor network can be deployed without much expense and maintenance can be used to reduce the time needed to detect leakage or contamination in urban water distribution systems [[7](#_bookmark25)]. Very low applied risks in IoT implementation in cities are the reason why the Public Service industry is heavily influenced by IoT. Electronics are able to share information through internet connections. Health service can provide better healthcare system and medical data through qualitative analysis in diagnosis.



**Figure 1.** Trends of Industry Impacted by Internet of Things (IoT). Adapted from [[5](#_bookmark23)].

IoT is not limited to public uses only but can also be used privately. With a central integrated IoT system, the home atmosphere can be adjusted by the pressing of a button, be it temperature, air control, or ambient music. Furthermore, there is the option for smart home security systems which can incorporate cameras, motion detectors, and locks, to notify home owners immediately if the system suspects burglary or intrusion of property. Household IoT systems are able to understand the user’s life habits and appropriately evolve and adapt into a smart housekeeper through constant self-perception and self- checking [[4](#_bookmark22)]. Having all of these features adds convenience, customization, security, and ease of use to life at home [[8](#_bookmark26)].

Evidently, IoT minimizes human efforts in many life aspects whilst promoting efficient resource utilization. It guarantees high speed, accurate quality data with secure processing and better client or user experience [[9](#_bookmark27)]. These imply, amongst other advantages, the re- liability and validity of data, performance, security and privacy. Table [1](#_bookmark1) shows that IoT units are becoming increasingly popular for not just consumer use, but also business and industries and are projected to rise at a steady rate for the coming years [[10](#_bookmark28)].

Lakes and streams are the planet’s most important freshwater system. According to their immediate environments, they are ecosystems and natural life habitats and form part of the food chain from vegetative material to animals to humankind. Rivers are complex life support systems that operate on a thin line of sustainability [[11](#_bookmark29)]. In recent history, the sharp increase in the human population has resulted in a considerable increase in the need for freshwater worldwide. Coupled with other factors such as global warming and anthropogenic inputs (pollution from municipal and industrial wastewater discharge), it is not unbelievable to say that the quality of water is now a major concern for experts around the world. To comprehend the effort and investment needed to obtain fresh drinking water, it is necessary to first understand the fundamental problems faced by freshwater systems.

**Table 1.** IoT units installed by categories (millions of units).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Category** | **2016** | **2017** | **2018** | **2019** |
| Consumer | 3963.0 | 5244.3 | 7036.3 | 12863.0 |
| Business: Cross-Industry | 1102.1 | 1501.0 | 2132.6 | 4381.4 |
| Business: Vertical-Specific | 1316.6 | 1635.4 | 2027.7 | 3171.0 |
| Grand Total | 6381.8 | 8380.6 | 11196.6 | 20415.4 |

The quality of surface water is largely affected by natural processes as well as man- made impacts, whereas surface water runoff is a seasonal phenomenon largely affected by climate; anthropogenic discharges represent a constant polluting source to rivers and streams. The Environmental Protection Agency (EPA) attributes Nonpoint Pollution Sources (NPS) as the reason that America’s lakes, rivers and estuaries, in general, remain polluted. Some of the NPS can be prevented, but much of it is a result of the combination of rain, melting snow and irrigation systems [[12](#_bookmark30)]. All three of these events mean that water picks up all types of debris and pollutants in its path to waterways. Water runoff from parking lots, industries, farmlands, and suburb carries oil, gasoline, pesticides, sewage and various other contaminants into water supplies, lakes, rivers and eventually the oceans. Trash, plastic bottles and other refuse also are carried away by floods and rainstorms. These pollutants can have a negative and devastating impact on vegetation and aquatic ecosystems. Thus, activities that can generate NPS include, but are not limited to [[13](#_bookmark31)]:

* 1. Sediment washing from agriculture.
  2. Deadly viruses and bacteria from animal grazing.
  3. Construction works.
  4. Aftermath of natural disasters particularly floods, tornadoes, hurricanes, and tsunamis.
  5. Gasoline and oil from recreational boating.
  6. Old and leaky septic systems.
  7. Urban runoffs from homes and landfills.
  8. Chemicals from household mismanagement and so on.

Climate and seasons have an effect on the baseflow of rivers and streams. Changing trends in rainfall have contributed to water shortages and affected terrestrial habitats due to variation in precipitation patterns and intensity [[14](#_bookmark32)]. Extreme rainfall may cause disasters such as rainstorms, floods, and erosion, all of which have the ability to alter the natural ecosystem of rivers, streams, and lakes. There is a clear relationship between precipitation and decreased river water quality. The regression models used in reference [[15](#_bookmark33)] have demonstrated that bacteria concentrations increase exponentially with observed precipitation. With heavy rainfalls, it is expected that erosion loss in a freshwater ecosystem would be significant, potentially causing an alteration to the depth of the river bed and to the river flow. Sub- jectively however, rainfall does not directly influence sediment discharge, but rather the interplay between rainfall and land-use activities affecting sediment production [[16](#_bookmark34)].

What then defines a healthy river? According to [[17](#_bookmark35)], “a biological system can be considered healthy when its inherent potential is realized, its condition is stable, its capacity for self-repair when perturbed is preserved, and minimal external support for management is needed”. Simply put, a river may be defined to be in good condition if its appearance remains stable and is able to rectify any unnatural changes by itself. When conducting water quality monitoring, there are many indicators to choose from. These may be from a biological perspective such as observation of local aquatic life residing in the body of water, or from a physical and chemical perspective such as soil erosion, stream flow, and sediment

discharge [[18](#_bookmark36)]. For a small-scale river, measurement of a few important parameters can suffice in giving a general idea to river health.

Currently only few studies exist which investigated ground or surface water in Brunei Darussalam [[19](#_bookmark37),[20](#_bookmark38)]. Results of studies and ongoing monitoring activities have highlighted certain level of pollution in the Brunei River. Possible complex contamination scenarios resulting from this may require a range of remediation and assessment measures [[21](#_bookmark39)]. However, natural occurring microbial processes are able to break down even recalcitrant contaminants [[22](#_bookmark40)].The main sources of pollution have been traced to: effluent and sludge from sewage treatment works within the catchment, sullage waste, solid waste, and direct disposal of sewage from the water village and nearby settlements (50%), surface water runoff from the capital’s central area (29%), and point and nonpoint pollutant loads from various sub-catchment uses, including agricultural, residential and industrial uses [[23](#_bookmark41)]. Under the Tenth National Development Plan, upgrades to existing monitoring systems and quality management frameworks as well as installation of new drainage and sewage systems have been proposed by the government [[23](#_bookmark41)]. However, this does not take away the fact that Brunei does not have a method of active online water quality monitoring. The current scheme of monitoring targets the source water where river water is piped to a treatment plant and samples are collected via sampling tap in the treatment plant’s laboratory. Water quality monitoring results can also be compiled in a database for improved decision making [[24](#_bookmark42)]. Laboratory results take time to be processed and, even then, the results might be inaccurate as certain parameters vary onsite and in-lab. Thus, this study is undertaken to lay the foundation for making advancements in the field of online water quality monitoring.

The main objectives of the study are to develop Internet of Things (IoT) systems, con- sisting of multiple sensors, communication link, storage and processing capabilities, energy for powering the device, etc., in order to monitor water quality of rivers/streams and also to identify the causes and factors contributing to water quality issues around the vicinity if any. A testing site in Universiti Brunei Darussalam (UBD) has been chosen for the study where the IoT system was placed and monitored directly. Due to financial constraints, the sensors chosen will only focus on the most important parameters. The study took into account the measurement of defined parameters in order to offer real-time online monitoring feedback to users. The data gathered from different IoT sensors would be used in combination with other data to perform data analysis and the results obtained are used to propose preventive measures on how to minimize the impact of pollution. The foundation laid by this study will be kept in order to develop a fully integrated IoT system in the future.

The next section will provide insights on the research component selection, prototype setup, sensor calibrations, and test site selection. Results are given in Section [3](#_bookmark11), which is followed by discussions in Section [4](#_bookmark13). The final section concludes the paper.

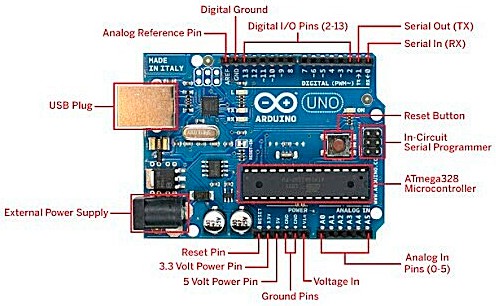
# Design and Development

The development of a simple prototype system fit for water quality monitoring needs to be comprised of the following components:

1. Multiple sensors to collect relevant data from the environment.
2. A central microcontroller loaded with a computer program to read analogue data and convert them to digital output.
3. A portable laptop with relevant software to read the digital data and present the data in an understandable format on a screen, as well as to provide power to the microcontroller.

The main component of the monitoring system is the Arduino UNO R3 Board shown in Figure [2](#_bookmark2). It is a microcontroller board based on the ATmega328 with three important features:

1. A total of 6 analog input pins labelled A0 to A5 to allow up to a maximum of 6 analog sensors to connect directly to the Arduino.
2. A total of 2 power supplies pin labelled 3.3 volts and 5 volts with in-built voltage regulation to provide power to sensors.
3. A USB plug that can be used in conjunction with a USB cable to connect with a microprocessor.



**Figure 2.** Arduino board and its connections.

The system used is an Arduino microcontroller with four accommodating sensors: pH, Temperature, Turbidity, and Total Dissolved Solids (TDS). Sensors were chosen based on ease of use, measurability (of parameters), portability, as well as being economical and cost-effective as a strict budget must be adhered to. The sensors are collectively shown in Figure [3](#_bookmark3).



(**a**) (**b**)



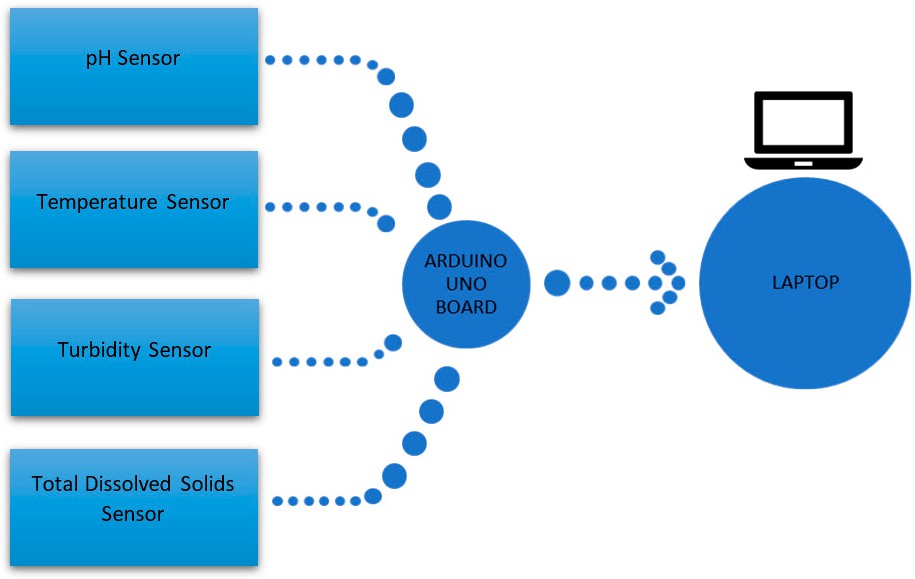
(**c**) (**d**)

**Figure 3.** (**a**) pH sensor, (**b**) temperature sensor, (**c**) turbidity sensor, (**d**) total dissolved solids sensor.

Figure [3](#_bookmark3)a shows the RELAND SUN pH sensor with a temperature range of *−*10 *◦*C to 50 *◦*C. It has a response time of under 5 s but a settling time of nearly a minute for readings to stabilize. Figure [3](#_bookmark3)b shows the DS18B20 temperature sensor that has an operating temperature range from *−*55 *◦*C to +125 *◦*C with accuracy +/*−*0.5 *◦*C (between the range

*−*10 *◦*C to 85 *◦*C). The ReYeBu turbidity sensor shown in Figure [3](#_bookmark3)c, was used to measure turbidity and has an operating temperature of 5 *◦*C~90 *◦*C with a response time of less than 500 ms. The OOTRTY total dissolved solids (TDS) sensor referred to in Figure [3](#_bookmark3)d was used to measure TDS. It has a measurement range of 0~1000 ppm and measurement accuracy of *±*10% F.S. It should be noted that the sensor cannot be used above 55 *◦*C and it is advisable not to place the sensor too close to the edge of a water surface.

Figure [4](#_bookmark4) below shows the connection of the sensors to the Arduino microcontroller and operating laptop.



**Figure 4.** Block diagrams of the connections.

Due to the limited number of power outputs of the Arduino, the power pin of the Arduino was connected to a breadboard to allow powering of multiple devices at the same time, as shown in Figure [5](#_bookmark5) below.

.Both turbidity and pH sensors require calibration to convert the obtained voltage read- ings to the corresponding turbidity and pH readings. To calibrate, different concentrations

of soil and water mixtures and different pH solutions were used to calibrate the turbidity and pH sensors, respectively.

The turbidity sensor was calibrated by measuring several soil and water mixtures made from known masses of soil mixed with 0.6 L of water. The data are tabulated below in Table [2](#_bookmark6).

**Table 2.** Data for turbidity calibration.

**NTU Voltage**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Soil Mass** | **Water** | **Nephelometric Turbidity Average Average** | | |
| **(g)** | **Volume (L)** | **Units (NTU) Readings** | **Reading** | **Reading (V)** |
| 1.0169 | 0.6 | 120.1, 128.3, 140.4, 121.3, 120.6 | 126.14 | 3.97 |
| 2.0190 | 0.6 | 392.0, 398.0, 392.0, 396.0, 400.0 | 395.60 | 3.68 |
| 3.0096 | 0.6 | 407.0, 422.0, 412.0, 428.0, 422.0 | 418.20 | 3.58 |
| 4.0201 | 0.6 | 677.0, 657.0, 690.0, 702.0, 664.0 | 678.00 | 3.25 |

Figure [6](#_bookmark7) below shows the plotted relationship between the average voltage readings and average NTU readings.

Average voltage readings (V)

Average NTU reading

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 1000 |  | | | | | |
| 900 |  |  |  |  |  |  |
| 800 |  |  |  |  |  |  |
| 700 |  |  |  |  |  |  |
| 600 |  |  |  |  |  |  |
| 500 |  |  |  |  |  |  |
| 400 |  |  |  |  |  |  |
| 300 |  |  |  |  |  |  |
| 200 |  |  |  |  |  |  |
| 100 |  |  |  |  |  |  |
| 0 | 3 | 3.2 3.4 | 3.6 | 3.8 | 4 | 4.2 |

**Figure 6.** Relationship between turbidity and voltage.

The pH sensor was calibrated by testing it against three solutions of Atlas Scientific pH samples of pH 10, 7, and 4. Due to a room temperature of 20 *◦*C, there is a slight pH change and actual pH is adjusted according to specification labels. Data from calibration is found below in Table [3](#_bookmark8).

**Table 3.** Data for pH calibration.

|  |  |  |
| --- | --- | --- |
| **pH** | **Voltage [V]** | **Average Voltage [V]** |
| pH 10.06 | 2.20, 2.20, 2.20, 2.19, 2.20, 2.19, 2.21, 2.20, 2.20, 2.20, 2.20 | 2.20 |
| pH 7.02 | 2.69, 2.65, 2.65, 2.64, 2.65, 2.64, 2.65, 2.64, 2.63, 2.64, 2.65 | 2.65 |
| pH 4.00 | 3.01, 3.09, 3.02, 3.02, 3.03, 3.02, 3.03, 3.03, 3.03, 3.04, 3.03 | 3.03 |

Figure [7](#_bookmark9) below shows the plotted relationship between the average voltage readings and average pH readings.

Programming was done with an Integrated Development Environment (IDE) based in Java with the programming language C/C++. With the IDE, program sketches can be made, compiled, debugged, and uploaded to the Arduino microcontroller board to be executed. References to build the program sketches may be found online and OneWire

pH

and Dallas Temperature libraries needed to be downloaded for the temperature sensor to function.

**Figure 7.** Relationship between pH and voltage

10.6

9.6

8.6

7.6

6.6

5.6

4.6

3.6

2.1

2.2

2.3

2.4

2.5

2.6

2.7

2.8

2.9

3

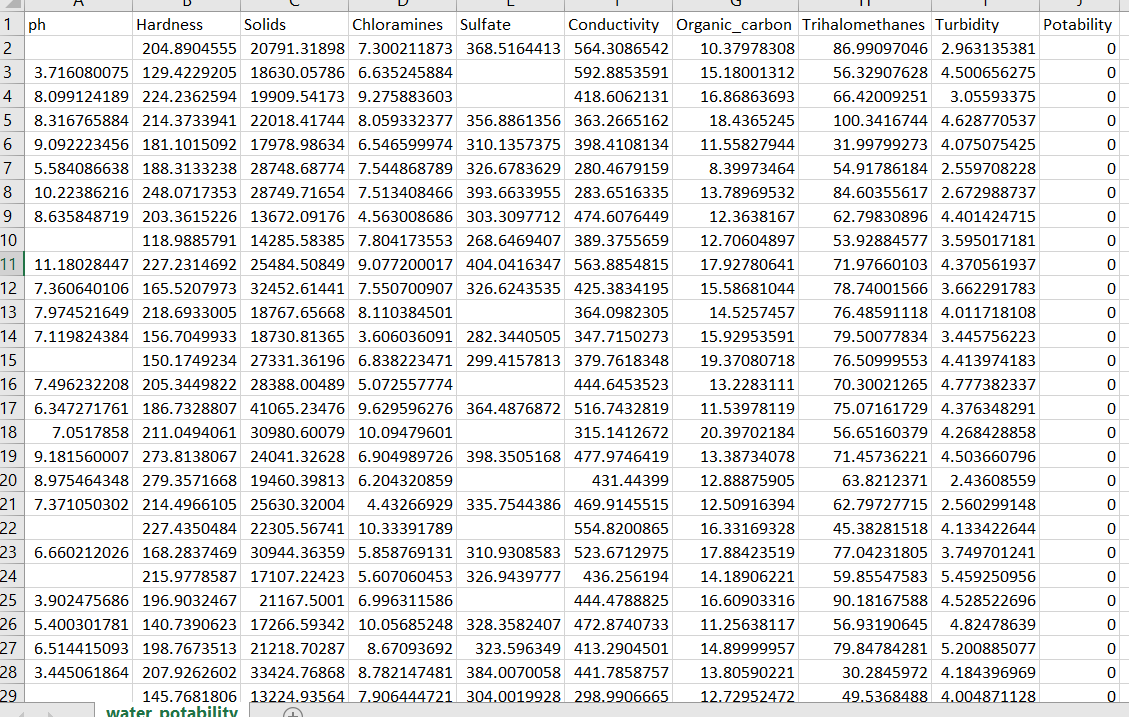
3.1

Average voltage (V)

# Results

Table [4](#_bookmark12) below shows the data gathered during the duration of the test. Data for pH were initially unavailable while the pH sensor was being calibrated. Readings for pH only became available from week 3 onwards after the sensor was readied.

**Table 4.** Data for the duration of the test.



The data from the four weeks were compiled and plotted on graphs to analyze any correlation between parameters and to evaluate the health of the river.

# Discussion

From Figure [9](#_bookmark14), it can be seen that temperature in the morning varies between 24 *◦*C to 27 *◦*C whereas afternoon temperature varies from 25 *◦*C to over 28 *◦*C. Additionally from Table [5](#_bookmark15), the maximum recorded temperature is 28 *◦*C while the lowest is 24 *◦*C, and

the deviation is slight per day. Heating of the earth by the sun is cumulative throughout the day and maximum temperature of the day is achieved during noon to mid-afternoon, between 12 p.m. and 3 p.m., when accumulated solar energy is at its maximum [[25](#_bookmark43)]. In a tropical country such as Brunei, the season is akin to summer all year round and solar radiation is high so it is not unusual for the river temperature to approach 30 *◦*C. Unsteady heating by the sun causes variations in heat accumulation and accounts for fluctuations in daily temperatures. Data for day 13 and 20 are noted for being lower than expected due to the presence of rain earlier in the day.

**Table 5.** Summary of parameters.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Temperature** | **Turbidity** | **TDS** | **pH** |
| Average | 26.48 | 1986.99 | 210.67 | 6.59 |
| Minimum | 24.28 | 820.54 | 19.40 | 0.00 |
| Maximum | 28.23 | 2643.2 | 956.30 | 14.00 |
| Standard 0.98 | | 444.56 | 188.75 | 5.17 |

Deviation

Figure [10](#_bookmark16) shows that turbidity values are higher in the afternoon. The tested stream is artificial, its contents come from underground pipes that flow in the morning but not in the afternoon. According to [[26](#_bookmark44)], when the stream is flowing the kinetic energy of the high-speed stream can resuspend river bottom solids and carry suspended solids which would account for high turbidity values. In a still body of water, suspended solids would settle on the river bottom and give lower turbidity readings. Thus, according to theory, the values should be higher in the morning but Figure [10](#_bookmark16) does not justify this. A possible explanation would be the oversaturation of suspended solids in the afternoon still river. With no flow, the solids remain within the stream and contribute to turbidity readings and there is no sediment transfer downstream. It is also known that turbidity and temperature share a positive correlation. Suspended solids have a higher heat capacity than water and are able to absorb more heat, leading to a higher recorded temperature [[27](#_bookmark45),[28](#_bookmark46)]. Indeed, both temperature and turbidity readings are higher in the afternoon. It should however be noted that the exact turbidity value is difficult to pinpoint due to natural variation in season, local geology, river flow, weather, and climate and readings vary slightly every few seconds when measuring, though a value not exceeding 3000 NTU is deemed favorable.

Figure [11](#_bookmark17) shows an increasing trend in dissolved solids in the river over the course of the sampling period. To begin understanding this trend, the relationship between total dissolved solids, conductivity, and temperature must first be clarified. Total dissolved solids and conductivity are directly proportional to one another. Pure water cannot hold any electrical charge but water that contains minerals and salts can. Any body of water

that contains dissolved salts and mineral will give a conductivity reading. Conductivity of water (and by extension dissolved solids) is expected to increase with temperature. For each 1 *◦*C increment, conductivity rises by 2–4%. Temperature influences conductivity by increasing ion mobility and dissolvability of many salts and minerals [[29](#_bookmark47)]. Figure [12](#_bookmark18) shows that average temperature increased from Day 4 to 12, likewise there was a general increase in TDS values in the same period. In this period, the correlation coefficient was 261 ppm/*◦*C in the morning and 177 ppm/*◦*C in the afternoon, indicating a big increase in total dissolved solids per temperature increment. Furthermore, the data from Day 15 to 19 where there was a decrease in average temperature also resulted in a general decrease in TDS values in the same period. During this period, the correlation coefficient was 106 ppm/*◦*C in the morning and 117 ppm/*◦*C in the afternoon, indicating a marginal decrease in total dissolved solids per temperature decrease. Total dissolved solids (TDS) is usually low for freshwater sources, at less than 500 ppm [[30](#_bookmark48)] which is agreed by most of the data here.

As can be seen from Figure [12](#_bookmark18) and Table [5](#_bookmark15), the pH values recorded were inaccurate. The afternoon data for day 17 and morning data for day 19 lie within the extremes of the pH range of 0–14. A reason for low water pH values can be related to stimulation treatments or CO2 sequestration operations performed in the oil and gas industry [[31](#_bookmark49),[32](#_bookmark50)] however which have not been conducted in this area. The remaining pH values vary significantly each day and onsite readings can change drastically every second. This is a stark contrast to its stability when being calibrated under lab conditions. A logical explanation would be that this is due to the artificial nature of the stream where its contents are influenced by local underground discharge pipes. The pipes may have contained high amounts of hydrogen or hydroxy ions. Another possible explanation would be the decomposition of leaves in the river bottom which would release carbon dioxide, allowing the formation of carbonic acid. Carbonic acid can lose one or both its hydrogen ions as shown in the equations below:

Equation (1): Formation of carbonic acid

*CO*2 (*g*) + *H*2*O* (*l*) *H*2*CO*3 (*aq*) (1)

Equation (2): Loss of first hydrogen ion

*H*2*CO*3 (*aq*) *HCO*3*−* (*aq*) + *H*+ (*aq*) (2)

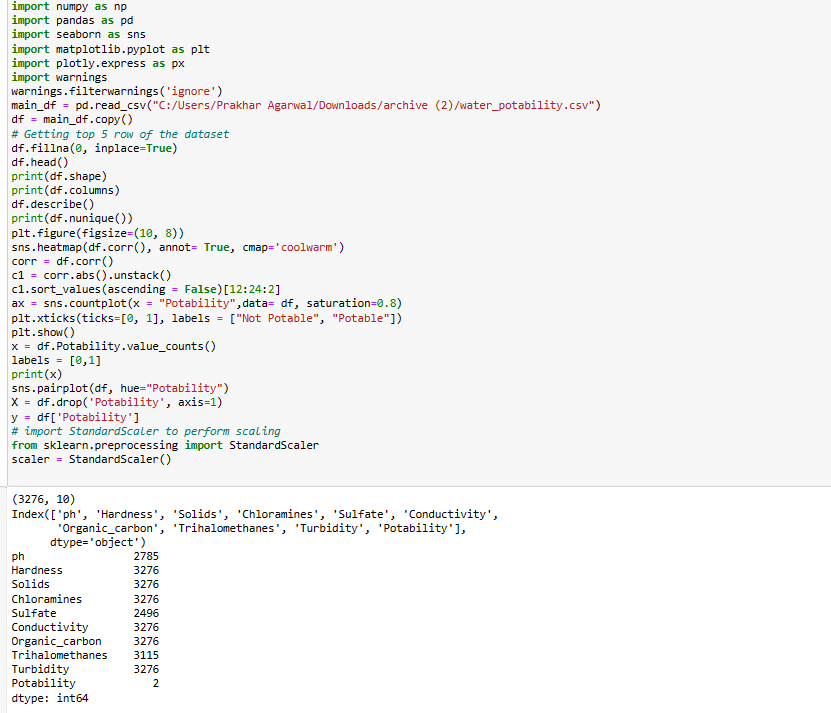
Equation (3): Loss of second hydrogen ion

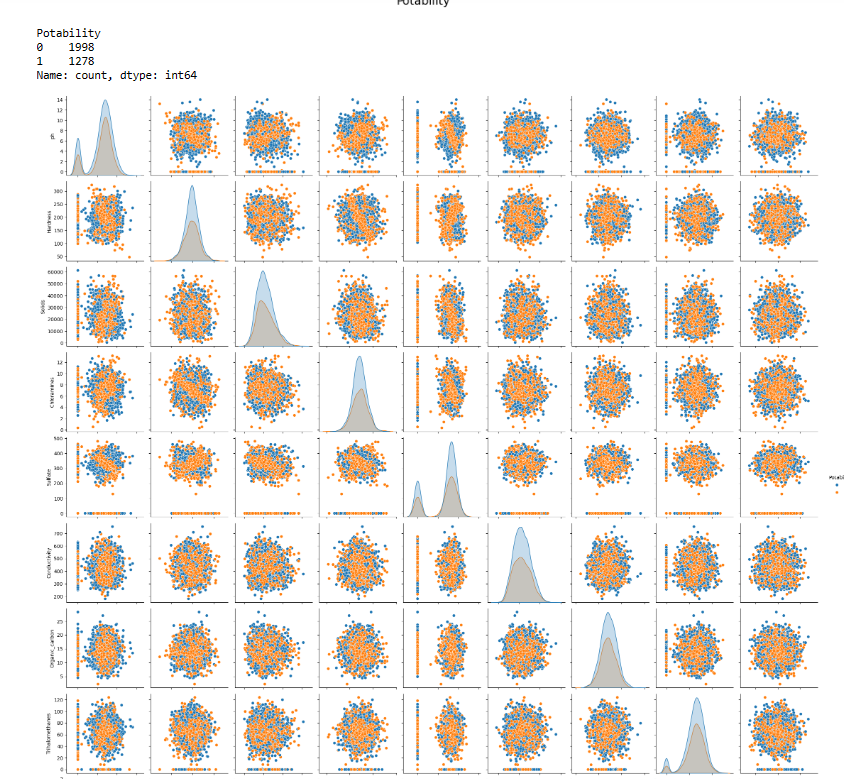
*HCO*3*−* (*aq*) *CO*3 2*−*(*aq*) + *H*+ (*aq*) (3)

The release of hydrogen ions will decrease the pH of the river. However, at higher pH the equilibrium will shift towards the left side, promoting formation of carbonic acid and the resulting decrease in hydrogen ions will increase the pH of the river [[33](#_bookmark51)]. Regardless, the shift in pH due to equilibrium should not be too great and, admittedly, the data gathered for pH is inadequate and more time might have made for better data comparison and evaluation. There were no data for pH for the first two weeks while the pH sensor was still being calibrated.

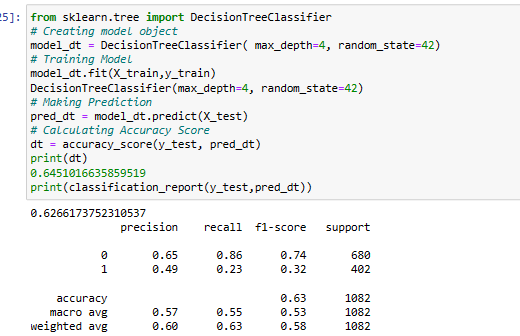
According to a report by Tziortzioti et al. (2019), the system they used is nearly identical to the same setup that we had used, apart from different sensors and an enclosed box to protect the Arduino microcontroller and sensor modules [[34](#_bookmark52)]. From their experiment it can be seen that the readings they obtained are relatively stable, whereas, in the present setup, the readings of pH differ greatly every second. Taking into consideration that the pH sensor was stable during calibration, it may be surmised that the artificial nature of the sample stream may be the cause of the fluctuations, though an experiment in normal water may be needed to justify this claim. The results that they obtained align quite well with our present study. It is also noted that they experience the same problem with the turbidity sensor as in this study, the sensor cable being too short, the issue of the sensor floating on the water surface instead of being fully submerged, and sensor malfunctioning due to water infiltration into sensor core. Additionally, their turbidity sensor suffered from rust due to seawater exposure and needed to be replaced.

**MACHINE LEARNING BASED MODEL FOR PREDICTING THE QUALITY:**

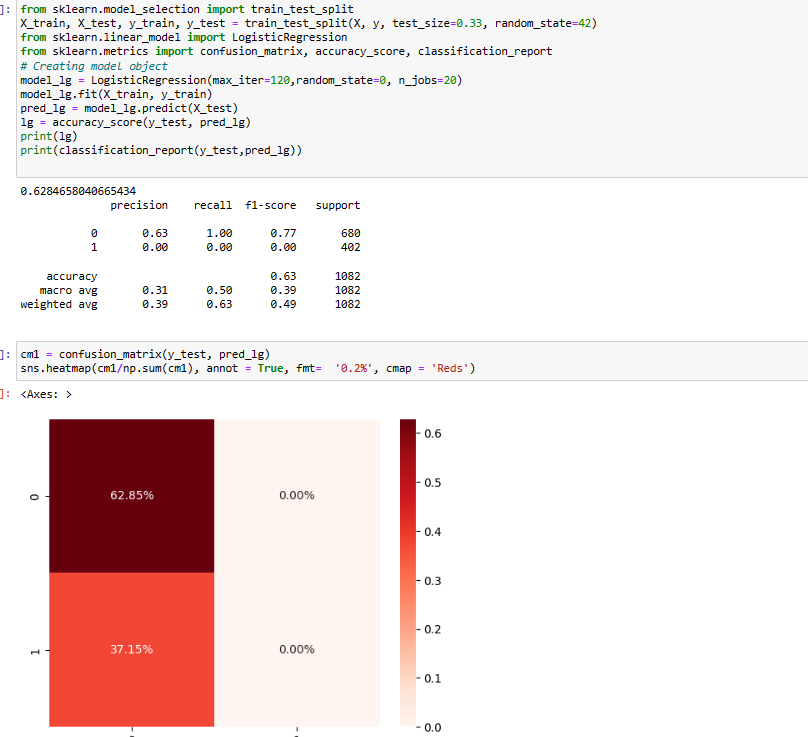




**DECISION TREE CLASSIFIER:**



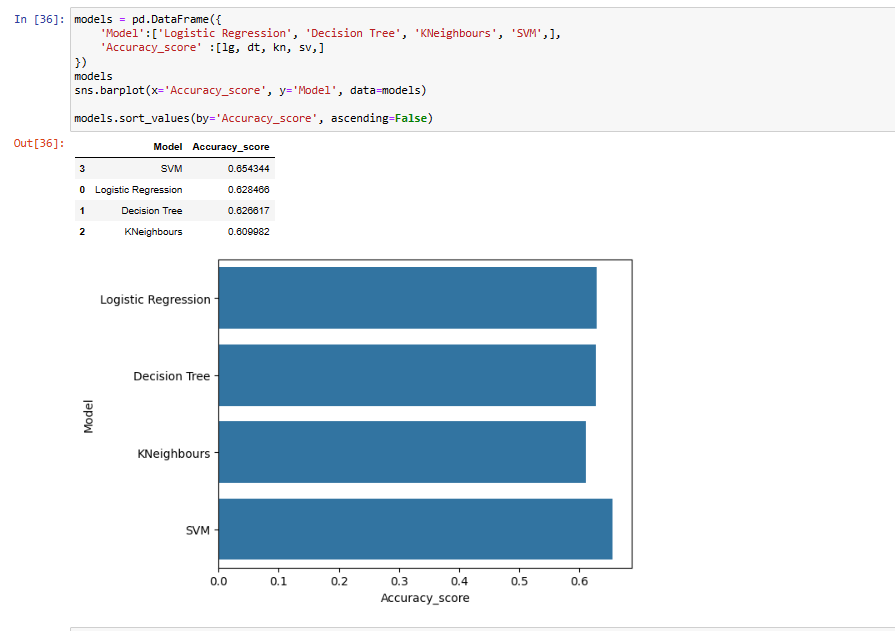
**LINEAR REGRESSION MODEL:-**



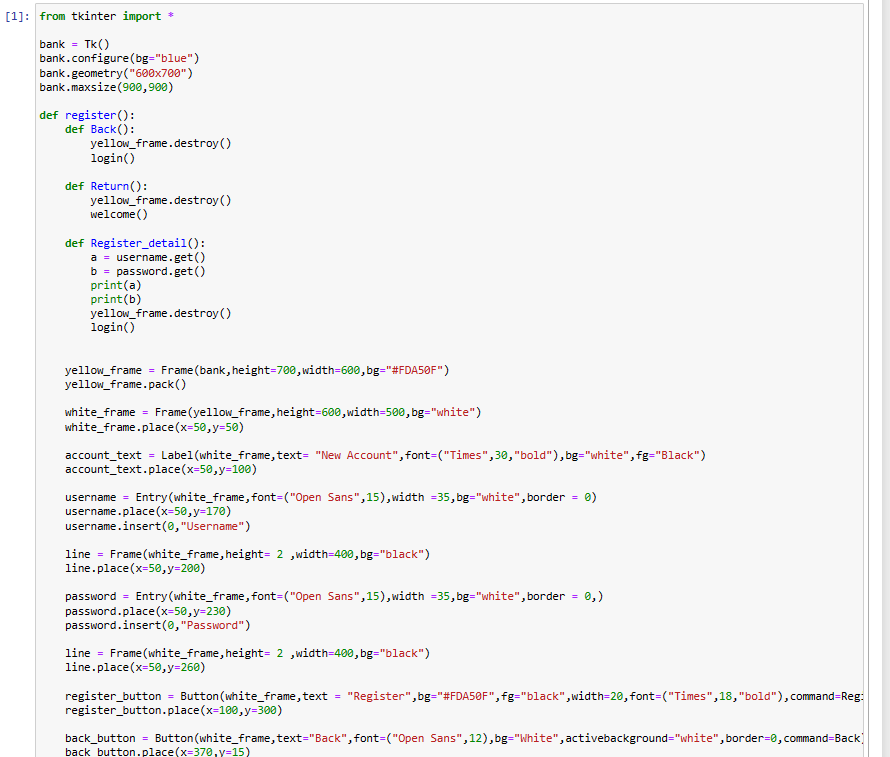
**K NEAREST NEIGHBOUR:**

****

**FINAL CONCLUSION: -**



**GUI BASED INTERFACE TO SEE THE RESULTS: -**



# 

# OUTPUT:-

# 

# Conclusions:

The results seem to indicate that the stream is healthy, however more tests are required to further determine data validity and operability of the current system before the system can be deployed elsewhere. The prototype yet lacks many capabilities and thus require upgrades and expansions; namely, to give it the ability to transmit data through a wireless network to a remote laptop or mobile at any given time and location, and a stronger memory drive or the ability to store data in databases. To that end, a Raspberry Pi 4 Model B could be used in conjunction with the current system. The Pi offers ground-breaking increases in processor speed, multimedia performance, memory, and connectivity. It is outfitted with dual-band wireless LAN and Bluetooth which can be routed to a dedicated monitor which is vital in wireless networking. Additionally, a SIM card would allow it to store quantities of data from the Arduino. The combined central system must then be encased to prevent potential detrimental drawbacks from exposure to environmental weather and conditions. Among the sensors, the turbidity sensor must be improved or replaced altogether. An opening at the top of the turbidity sensor allows water to enter if the sensor is lowered too deep into the stream and the influx causes a direct disturbance in readings. Its short cable meant that it is also severely limited to use at the water edges. A suitable replacement could be the digital type RELIHONES turbidity sensor that is waterproof and offers a longer cable. Implementation of additional sensors in the project extension will allow for monitoring of extra parameters, particularly of concentration of various ions. Such data may be helpful for better determining river health through calculation of Water Quality Index (WQI) or other related indexes. Finally, changes could be made to the testing environment and schedule. Conducting the test at additional locations and taking more frequent data readings should provide for much needed additional data and accuracies in determining river health and validating the monitoring system.

# 