3.1 SYSTEM ARCHITECTURE

The following figure shows a system architecture of this study. In the physical layer there are some sensors which is connected to the network layer that has Arduino UNO and its WiFi module. All the sensor reading will be shows on ThingSpeak that users can access from their personal computer or mobile phone.

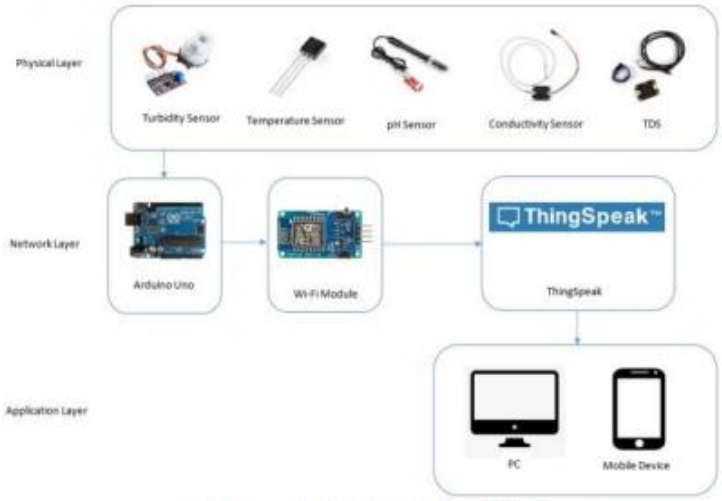

Figure 1 System Architecture

3.2 HARDWARE

3.2.1 ARDUINO UNO

Arduino is an open source development platform that is relatively simple and easy to use with high-level language which is similar to Java and C language.

Figure 2 Arduino UNO

3.2.2 ESP 8266

The ESP 8266 is a Wi-Fi module which permits microcontrollers access to a Wi-Fi arrangement and offers an independent and total Wi-Fi organizing arrangement.

Figure 3 ESP 8266

3.2.3 TEMPERATURE SENSOR

LM35 is used to measure the temperature of water for this study. This sensor can measure from -55ºC to 150ºC and

the level of accuracy is very high if it operated at optimal temperature and humidity levels.

Figure 4 LM35

3.2.4 pH Sensor

The pH sensor that is used in this study is SKU: SEN0161 to determine hydrogen ion concentration in water and demonstrate its alkalinity and acidity. The pH value for each liquid is between 0 to 14. The pH sensor has 3 pins that are power, ground and output.

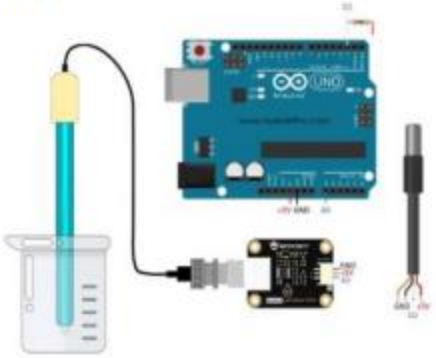

Figure 5 SKU: SEN0161 Connection

3.2.5 TDS SENSOR

Total Dissolved Solid refers to any minerals, salts and metal that dissolved in water. This sensor has 2 electrodes apart that will be inserted into the water to measure the electrical charge (EC) and the result will be converted into ppm.

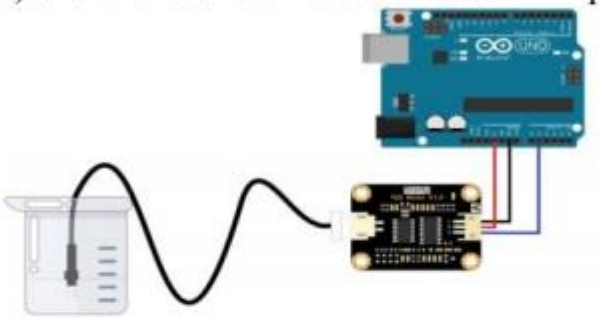

Figure 6 SKU: SEN0244 Connection

3.2.6 CONDUCTIVITY SENSOR

Conductivity indicates ionic strength of a chemical solution that is directly associated with salinity. This sensor will monitor the amount of nutrients, salts or impurities in the water.

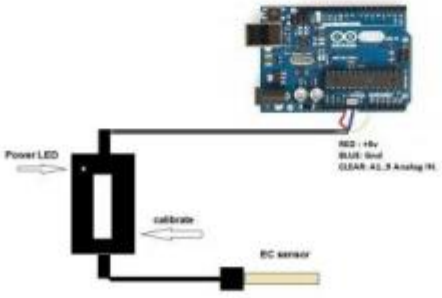

Figure 7 SKU: DFR0300 Connection

3.2.7 TURBIDITY SENSOR

The probe on this sensor works by sending a light beam into the water and the amount of light reflected is used to determine the particle density within the water.

Figure 8 SKU: SEN0189 Pinout

3.3 SOFTWARE

The following flowchart is shows the system working of this study. The important part of the flowchart are the sensors initializing and setting up the ThingSpeak platform to be used to read and record all the data of this system.

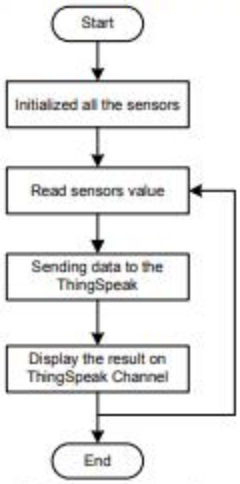

Figure 9 System Flowchart

GROUNDWATER LEVEL PREDICTION USING ML

DATA SET:

https://data.opencity.in/dataset/national-compilation-on-dynamic-ground-water-resources-of-india-2022

WATER POLLUTION MONITORING SYSTEM AND UNDERGROUND WATER LEVEL PREDICTION USING ML

ML MODELS USING FOR THIS PROJECT:

1. Random Forest Regression:

Advantages:

Handles non-linear relationships and interactions between features effectively.

Robust to outliers and noise in the data due to ensemble learning (combining multiple decision trees).

Can automatically handle feature importance and provide insights into which features are most influential in prediction.

2. Gradient Boosting Regression:

Advantages:

Builds powerful predictive models by sequentially improving weak learners (decision trees).

Particularly effective for complex relationships and capturing subtle patterns in the data.

Offers high predictive accuracy and can handle large datasets with many features.

3. Support Vector Regression (SVR):

Advantages:

Effective for high-dimensional datasets with complex relationships between features and target variable.

Uses kernel functions to map data into a higher-dimensional space, enabling it to capture non-linear relationships.

Robust to overfitting when proper regularization parameters are used.

Each of these models has unique strengths and suitability depending on the characteristics of your dataset and the specific requirements of your groundwater level prediction project. Experimenting with these models and tuning their hyperparameters can help identify which one performs best for your particular use case.