Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.777660

AN IOT BASED REAL TIME WATER QUALITY MONITORING SYSTEM USING ARTIFICIAL NEURAL NETWORK

Benitha V S
Department of Electronics and Communication Engineering
Mepco Schlenk Engineering College, Sivakasi.

Kousalya G
Department of Electronics and Communication Engineering
Mepco Schlenk Engineering College, Sivakasi.

Samrutha S K
Department of Electronics and Communication Engineering
Mepco Schlenk Engineering College, Sivakasi.

Abstract— The proposed IoT-based system is used for water quality monitoring utilizing an Artificial Neural Network (ANN). For the predictive maintenance of water, it needs to be monitored regularly to guarantee a safe supply of water from various resources. As a result, designing and developing a low-cost system for real-time water quality monitoring using the Internet of Things (IoT) is essential. The proposed IoT-based system along with the Wi-Fi module continuously monitors the quality of water. The built-in Wi-Fi module enables internet connectivity, which transfers sensor data to the Cloud. Water quality determining parameters like pH, turbidity, temperature, humidity, and dissolved solids are continuously monitored and its information is gathered in real-time. For the real-time water monitoring system, decision support is implemented using Artificial Neural Network (ANN) with forward propagation to classify water into two classes potable and non-potable. The system implementation provides an accuracy of 57.8%.
Keywords— IoT based system, Artificial Neural Network (ANN), Water quality management, Real Time Monitoring System (RTMS)

I. INTRODUCTION

Water is a global resource that is highly essential for agriculture, human life, and for industries. At the same time, the availability of drinking water is scarce and more new technology was needed for the continuous monitoring of portable drinking water. Due to the availability of contaminated water in large amounts due to urbanization and industrial evolution, water-borne diseases are spreading day by day, which increases the mortality rate [1] –[4]. Conventionally, water quality assessment is done manually, in laboratories. This process is a time-consuming process, which involves huge costs and human resources. Also, there is another system that gives an idea about the wastewater effluent treatment system using Long Short-Term Memory(LSTM) techniques by which wastewater can be monitored [5]. Another method to monitor the quality of water using a microwave sensor array with sensing elements operating at different frequencies [6]. Computer vision is also used in water quality monitoring systems with the help of biological sensors and deep learning techniques such as Recurrent Neural Networks (RNN) [7]. Among various techniques reported in the research community, IoT-based system is very promising for diverse types of applications for real-time water quality monitoring [8], [9].

Ange Josiane Uwayisenga et al. studied a paper on IoT-based floodwater detection and alarming systemin east African region [10]. Rufaizal Che Mamatet al. discussed Artificial Neural Networks and their applications for real-time monitoring [11]. Hina Afreen et al. discussed a system that automatically monitors and notifies the users about the condition quality of potatoes stating whether it is good, satisfactory, or alarming using ANN [12]. Shuangyinx Liuet al. studied a real-time water quality monitoring system fault diagnosis. It signifies whether the system has any fault in sensors using Support Vector Machines (SVM) and Rule-based decision Trees (RBDT) [13]. Abilio C. da Silva et al. discussed an online water monitoring system and managing water resources using the concept of, Internet of Water Things (IoWT) [14]. SanazImenetal., developed a model to assess the water quality where Decision Support System (DSS) is implemented [15]. Chuhan Qi et al., discussed water quality monitoring using remote sensing images, and LSTM is applied to predict the results [16].

From the above conclusions, it is inferred that there exists no real-time water quality monitoring system using Artificial Neural Network (ANN). In the proposed system, the quality of water is analyzed by using a powerful deep learning technique ANN. Forward Propagation Neural Network is implemented in this presented approach. The proposed work focuses on an IoT-based approach for real-time intelligent water monitoring where the water quality parameters like pH, turbidity, temperature, humidity, and total dissolved solids. Based on gauged environmental parameters, ANN model was used with forward propagation for water quality analysis and interpretation. The proposed system made use of low-cost hardware components, resulting in an affordable and cost-effective solution. To efficiently perform multi-classification about status, forward propagation neural network model with the "ReLu" activation function is used. Also, the system measured five important ambient environmental parameters pH, humidity, temperature, total dissolved solids, and turbidity.

II. METHODOLOGY

This study presents a gainful IoT system for monitoring water quality, which automatically analyses real-time environmental factors such as temperature, pH, turbidity, humidity, and the number of dissolved solids present in the water. The real-time values are saved in the cloud. The collected data are retrieved from the cloud and an ANN classifier is used to predict the state of water into two categories: potable and non-potable. Python and the scikit-learn package are used to carry out the simulation work[17].

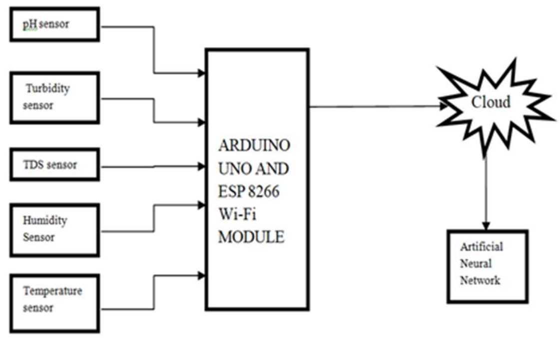

Fig. 1. Schematic representation of real-time water quality monitoring system

The proposed system consists of a sensing module and a status prediction module which is shown in the block diagram in Figure 1.

The suggested system includes five sensors: pH, turbidity, humidity, Total dissolved solids (TDS), and temperature, as well as a microcontroller unit as the main processing module and an ESP8266 Wi-Fi module for data transmission. The outputs of the sensors are directly connected to the Wi-Fi module unit's digital pins. Using the Wi-Fi data communication module ESP8266, all sensor data is processed by the module and updated to the ThingSpeak server.

The entire code is written in Embedded-C and emulated using the Arduino IDE. User login is provided to access the data using a user ID and password on the ThingSpeak server.

Using communication protocols, the RTMS sensing module predicts the real-time environmental data transmission from hardware to the cloud for storage. The data from the real-time environment is placed in a cloud Firebase database and then exported as a JSON file. The further processing of data is done by the status prediction module. Implementation of the ANN prediction model is done in Python with forward propagation.

A. ALGORITHM USED

An efficient water monitoring system for any real-time application is proposed in this work. The algorithms are developed based on a prediction model to make the suggested RTMS function. ANN is used to implement the proposed system. The first step describes the data acquisition which is done using the sensing system of RTMS.

After the data is read from the sensor, the collected data is positioned into the .csv file and it is updated in the cloud, and then further data processing is done.

B. HARDWARE

Embedded-based sensing devices are used in IoT-based systems to measure real-time environmental factors in water efficiently and cost-effectively [18], [19]. A sensor is a device that can detect changes in its immediate surroundings. To implement the proposed RTMS, various sensors and hardware components are used. Water's real-time environmental characteristics are measured using a sensor-based circuit that delivers the data to a cloud database, where a prediction model is applied to the gauge data to determine the water quality status. The devices communicate with the ESP- 8266 module, which is designed to accept data from the devices and send it to the cloud over the internet. The sensing system uses a microprocessor-based device called the ESP-8266, which has built-in Wi-Fi, Bluetooth, and Bluetooth Low Energy device.

ESP-8266 is appropriate for use in IoT devices due to its dual-core implementation. A temperature sensor DHT-22 is used in the sensing system to measure water temperature and humidity. It's a low-cost digital sensor that uses a capacitive humidity sensor and a thermostat to monitor the ambient air and then split off the digital signal on the data pin. Dissolved solids are another significant environmental parameter that must be assessed and has an impact on water quality. The concentration of dissolved solids was measured using a TDS sensor. Further, turbidity and pH measuring sensors are used to analyze the quality of water.

Furthermore, the Arduino IDE is used to develop a program in C to link all sensors to the ESP-8266 and to connect the ESP-8266 to the cloud. Environmental factors influencing water quality including temperature, relative humidity, pH, turbidity, and total dissolved solids are measured and recorded as JSON nodes in a cloud database and used for further data processing.

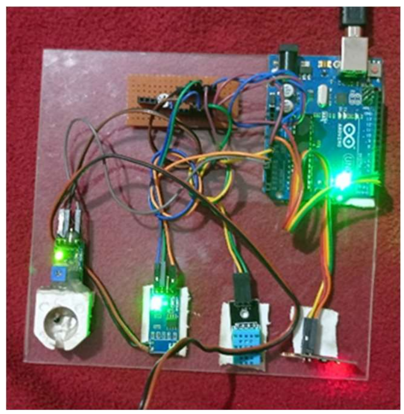

Fig. 2. Hardware Setup for water quality monitoring

C. ANN MODEL FOR DESIGNED IoT-BASED SYSTEM

The suggested module employs a forward propagation neural network to predict water quality based on ambient characteristics measured by the proposed RTMS's sensing system. Researchers used several classification techniques depending on the problem circumstances. The presented approach uses a forward propagation neural network, which is largely used in ANN architecture.

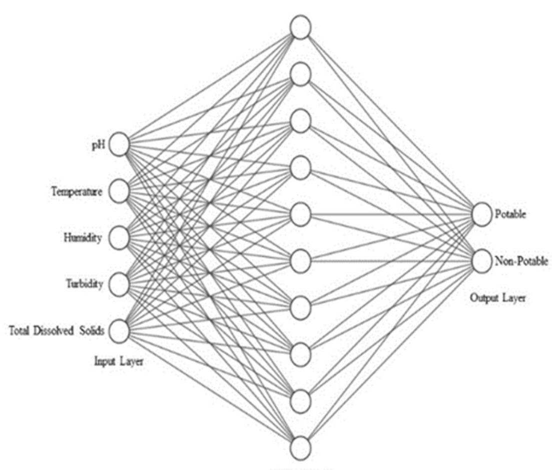

Fig. 3. ANN Architecture for the proposed work

ANN is chosen because of its ability to learn complex situations and flexibility. New features can be added by increasing the size of the training data, and no additional configuration is required in the event of a change. One of the most compelling reasons to use ANN is its ability to outperform other techniques because performance can be enhanced by increasing training data size. Figure 3 shows the complete ANN architecture for the proposed system

In ANN, several nonlinear processing units are present and they are as called artificial neurons. In the neuron, the inputs are received, like in a biological neuron. These neurons are linked together by weights. ANN may learn a task by altering weights. Every input (pH, turbidity, temperature, humidity, and TDS) is multiplied by the weight (w1, w2, w3...wb), and neuron j performs the summation of all weighted inputs and bias. The sum of previous weighted contributions and bias was then processed by an activation function, which produced the output Y.

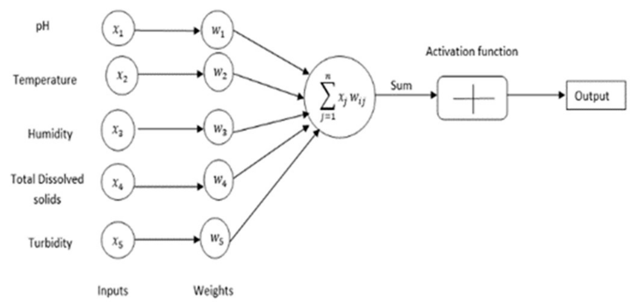

Fig. 4. Schematic of Artificial Neuron

From equation (1), it is understood that the mathematical description of the ANN is shown in Fig.4.

$$y = f\left(\sum_{j=1}^{n} I_j W_j + b\right) \qquad (1)$$

where Ij represents the input, Wj represents weight and b represented the bias value.

In deep learning, the SoftMax activation function is employed for the prediction on the output layer. Our suggested solution has five input features and two outputs per target, the system is designed using the Rectified Linear Unit (ReLU) activation function for the last layer because of its ability to handle multi-classification problems rapidly The ReLU activation function is defined in equation (2),

$$y = \max(0, x) \qquad (2)$$

For any negative input, it returns 0 and for positive input x it returns that value. As a result, the output has a range of 0 to infinite.

To address the supervised learning challenge, the proposed solution used a forward propagation neural network. The data set contains all conceivable values of environmental parameters as well as their truth values as input data. The neural network learns from the pattern of the input data set and can offer the appropriate output class of the test dataset, which contains only input data and no truth values. The forward propagation process of ANN was used in the proposed approach. It has three layers: the input layer, the hidden layer, and the output layer. There is no feedback information and the information is merely passed forward between these levels utilizing forward propagation neural network architecture. It employs a tiny neural network with only a single layer of ten neurons since it provides accurate predictions of environmental parameter change. Furthermore, a tiny neural network has fewer weights to estimate from training data, which confirms its practical difficulty and improves the neural network's generalization capability. Temperature, humidity, total dissolved solids, turbidity, and pH are represented by five neurons in the input layer.

Each neuron of the input layer is connected to everyone neuron of the hidden layer, and each interconnection has a weight linked with it, as well as a subscript to identify it. The hidden layer has 10 neurons, while the output layer has two neurons equal to the number of status classes. The dot product of each facet and weight is calculated, and bias is introduced inside the dot product before performing an on the whole summation. This information is passed on to the activation function, which will provide an output for the first neuron and the first hidden layer. The entire procedure will be repeated until the final weight for the final input has been generated.

III. IMPLEMENTATION

The sensing module of the RTMS is used to collect data on temperature, relative humidity, turbidity, pH, and total dissolved solids. This section explains how RTMS works by demonstrating the experimental setup that was used to verify the proposed system's functionality. Figure 5, Shows the complete proposed work of the software module.

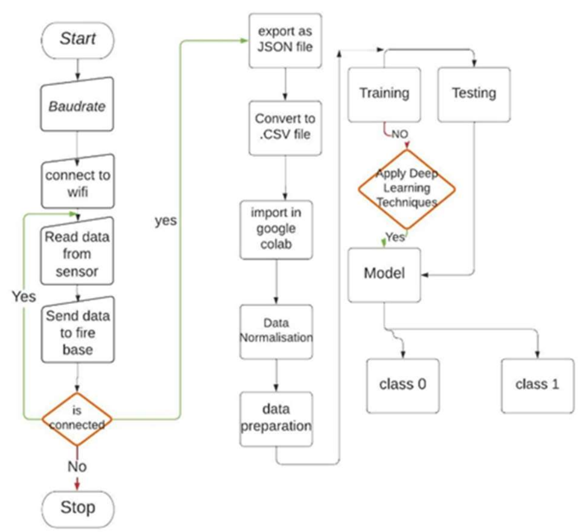

Fig. 5. Workflow of the Charted system

A. DATA ACQUISITION

The data is collected using an RTMS sensor system that measures real-time values of environmental parameters like temperature, humidity, turbidity, pH, and total dissolved solids. The ESP-8266 uses the internet to communicate real-time data to the cloud. The real-time database was exported as a json file, which was further converted to.xlsx. Figure.5 shows the complete workflow of the proposed idea. The data are collected at various intervals of time. The sensing module connects to Wi-Fi after starting at a baud rate of 115200. Sensing modules read real-time environmental parameters and communicate gauged data to the cloud if the connectivity is accessible. This procedure will be repeated until Wi-Fi connectivity is there. Then, using the panda's library method read_excel(), the file is imported into Google Colab (). The entire dataset is put on panda's data frame.

Furthermore, the gauged dataset's features are recorded as X, and status is targeted as y.

B. DATA NORMALISATION

Normalization and standardization are the two most common types of data scaling. The features of the dataset stored in variable X are normalized. Normalization is a rescaling technique that reduces data from its original range to a scale between 0 and 1. The fit () and transform () functions of the scikit-learn object Minmax Scaler are used to accomplish this. Equation 3 is used to represent the min-max normalization.

$$W_i = \frac{X_i - \min(X)}{\max(X) - \min(X)} \quad (3)$$

where denotes the number of features to normalise, and denotes the normalised features. This process will combine all of the features into a single scope with the same weights.

C. DATA PREPARATION

After normalization, data preparation is done, which involves dividing the dataset into training and testing sets. The training dataset accounts for 70% of the total dataset, while the remaining 30% is used for testing. For both the testing and training datasets, KNNcross-validation is used. A built-in function train test split is implemented to achieve this split.

D. BUILD PREDICTION MODEL

Following data normalization and preparation, the next step is to use the training dataset to build a prediction model. In our ANN model, there are three layers. To authenticate the hyperparameters of the ANN model, a grid search cross-validation approach is used. At the hidden and output layers, 'ReLU' (Rectified Linear Unit) and sigmoid activation functions were used respectively.

Since the proposed approach has a multiclass classification, the proposed prediction model is compiled using an Adam optimizer having a learning rate of 1e-3.

IV. RESULTS AND DISCUSSIONS

Here the outcomes of the software module are discussed. In using ANN the accuracy of the dataset is determined by the methods that have been implemented. The percentage of potability and non-potability is shown in Figure 6.

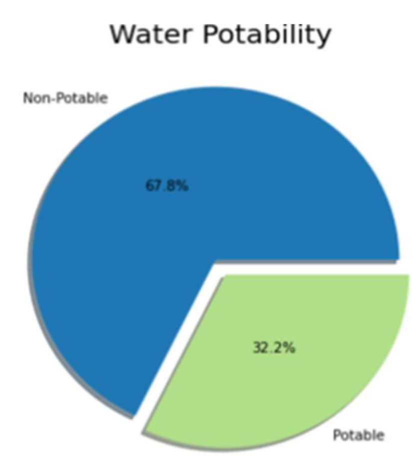

Fig. 6. Number of potable and non-potable water

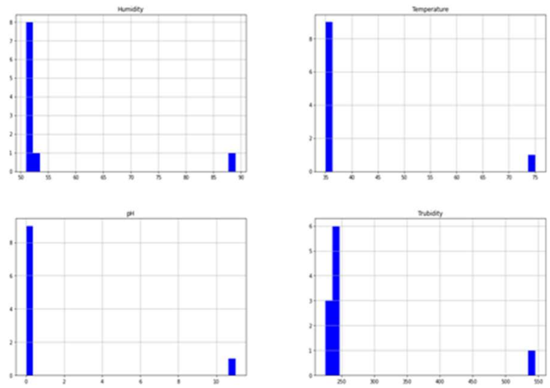

Fig. 7. Level of potability in each measured parameter

At the conclusion of each training period and plotted learning curves to highlight the performance of the ANN model over time.

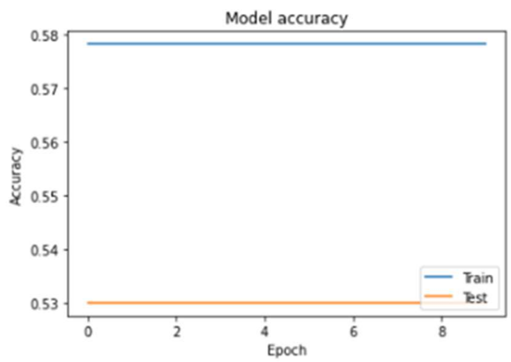

Fig. 8. Training and testing accuracy of ANN

The blue line reflects training accuracy, which starts at 57.8 percent at the first epoch and remains constant. Whereas the orange line shows the test accuracy that begins at 53% and remains constant. The ANN training and testing loss curves are shown in Figure 8. Findings reveal that the suggested ANN model outperforms with an accuracy of 57.8 % as shown in Figure 9.

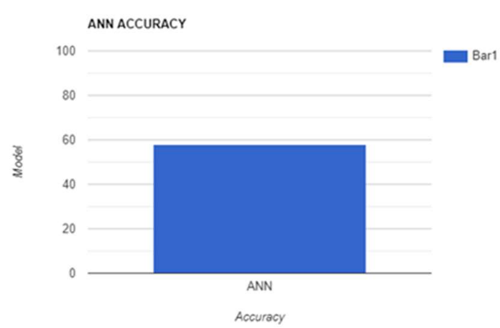

Fig. 9. Model Vs Accuracy graph

The obtained results from the hardware are shown. The data collected from the sensors are directly stored as a .csv file. Figure 10 depicts the values collected from sensors in the form of an excel file. The values obtained from the sensors are also viewed as a graph using the ThingSpeak server and are shown in Figure 11.

	A		B	C	D	E	
1	created_a	entry_id	field1	field2	field3	field4	field5
2	2022-04-0	1	51	35	0	246	82
3	2022-04-0	2	51	35	0	227	75
4	2022-04-0	3	51	35	0	247	82
5	2022-04-0	4	51	35	0	227	75
6	2022-04-0	5	51	35	0	226	75
7	2022-04-0	6	51	35	0	247	82
8	2022-04-0	7	51	35	0	246	82
9	2022-04-0	8	54	35	0	247	82
10	2022-04-0	9	53	35	0	245	81
11							
12							
13							
14							

Fig. 10. Values obtained from sensors

Fig. 11. Graphical representation of result from sensors

V. CONCLUSION

Water turbidity, pH, humidity, temperature, and total dissolved solids are monitored using a sensor that has a unique advantage and is already connected to a Wi-Fi-enabled system. The results are pretty accurate when a deep learning technique is incorporated while training data. Thus, it is concluded that ANN produces accurate results in water quality monitoring systems. In future work, the following parameters like detecting more parameters for the most secure application, and adding multiple sensors to the equation will increase the parameters, Other Deep learning techniques can also be implemented for results in new dimensions could be implemented to make the system a more beneficial one.

ACKNOWLEDGMENT

The authors would like to thank the Principal and the Management of Mepco Schlenk Engineering College for providing the necessary experimental facilities to carry out this project work.

REFERENCES

[1] S. Pasika and S. T. Gandla, "Smart water quality monitoring system with cost-effective using IoT," Heliyon, vol. 6, no. 7, p. e04096, Jul. 2020, doi: 10.1016/j.heliyon.2020.e04096.

[2] V. Lakshmikantha, A. Hiriyannagowda, A. Manjunath, A. Patted, J. Basavaiah, and A. A. Anthony, "IoT based smart water quality monitoring system," Global Transitions Proceedings, vol. 2, no. 2, pp. 181–186, Nov. 2021, doi: 10.1016/j.gltp.2021.08.062.

[3] A. Jerom B., R. Manimegalai, and R. Manimegalai, "An IoT Based Smart Water Quality Monitoring System using Cloud," in 2020 International Conference on Emerging

Trends in Information Technology and Engineering (ic-ETITE), Feb. 2020, pp. 1–7. doi: 10.1109/ic-ETITE47903.2020.450.
[4] L. Wang, “Biological monitoring and application in environmental monitoring,” Technology Innovations, vol. Vol. 8.
[5] X. Liu, Q. Shi, Z. Liu, and J. Yuan, “Using LSTM Neural Network Based on Improved PSO and Attention Mechanism for Predicting the Effluent COD in a Wastewater Treatment Plant,” IEEE Access, vol. 9, pp. 146082–146096, 2021, doi: 10.1109/ACCESS.2021.3123225.
[6] K. Zhang, R. K. Amineh, Z. Dong, and D. Nadler, “Microwave Sensing of Water Quality,” IEEE Access, vol. 7, pp. 69481–69493, 2019, doi: 10.1109/ACCESS.2019.2918996.
[7] F. Yuan, Y. Huang, X. Chen, and E. Cheng, “A Biological Sensor System Using Computer Vision for Water Quality Monitoring,” IEEE Access, vol. 6, pp. 61535–61546, 2018, doi: 10.1109/ACCESS.2018.2876336.
[8] M. Aly, F. Khomh, Y.-G. Gueheneuc, H. Washizaki, and S. Yacout, “Is Fragmentation a Threat to the Success of the Internet of Things?,” IEEE Internet Things J, vol. 6, no. 1, pp. 472–487, Feb. 2019, doi: 10.1109/JIOT.2018.2863180.
[9] E. Hossain, I. Khan, F. Un-Noor, S. S. Sikander, and Md. S. H. Sunny, “Application of Big Data and Machine Learning in Smart Grid, and Associated Security Concerns: A Review,” IEEE Access, vol. 7, pp. 13960–13988, 2019, doi: 10.1109/ACCESS.2019.2894819.
[10] A. J. Uwayisenga, N. Mduma, and M. Ally, “IoT-based system for automated floodwater detection and early warning in the East African region; a case study of arusha and dar es salaam, Tanzania,” International Journal of Advanced Technology and Engineering Exploration, vol. 8, no. 79, pp. 705–716, Jun. 2021, doi: 10.19101/IJATEE.2021.874099.
[11] R. C. Mamat, A. Ramli, A. M. Samad, A. Kasa, S. F. M. Razali, and M. B. H. C. Omar, “Artificial neural networks in slope of road embankment stability applications: a review and future perspectives,” International Journal of Advanced Technology and Engineering Exploration, vol. 8, no. 75, pp. 304–319, Feb. 2021, doi: 10.19101/IJATEE.2020.762127.
[12] H. Afreen and I. S. Bajwa, “An IoT-Based Real-Time Intelligent Monitoring and Notification System of Cold Storage,” IEEE Access, vol. 9, pp. 38236–38253, 2021, doi: 10.1109/ACCESS.2021.3056672.
[13] S. Liu, L. Xu, Q. Li, X. Zhao, and D. Li, “Fault Diagnosis of Water Quality Monitoring Devices Based on Multiclass Support Vector Machines and Rule-Based Decision Trees,” IEEE Access, vol. 6, pp. 22184–22195, 2018, doi: 10.1109/ACCESS.2018.2800530.
[14] A. C. D. S. Junior, R. Munoz, M. D. L. A. Quezada, A. V. L. Neto, M. M. Hassan, and V. H. C. de Albuquerque, “Internet of Water Things: A Remote Raw Water Monitoring and Control System,” IEEE Access, vol. 9, pp. 35790–35800, 2021, doi: 10.1109/ACCESS.2021.3062094.
[15] S. Imen, N.-B. Chang, Y. J. Yang, and A. Golchubian, “Developing a Model-Based Drinking Water Decision Support System Featuring Remote Sensing and Fast Learning Techniques,” IEEE Syst J, vol. 12, no. 2, pp. 1358–1368, Jun. 2018, doi: 10.1109/JSYST.2016.2538082.
[16] C. Qi, S. Huang, and X. Wang, “Monitoring Water Quality Parameters of Taihu Lake Based on Remote Sensing Images and LSTM-RNN,” IEEE Access, vol. 8, pp. 188068–188081, 2020, doi: 10.1109/ACCESS.2020.3030878.

[17] A. Geron, Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. 2019.
[18] S. Kavi Priya, G. Shenbagalakshmi, and T. Revathi, "Architecture of Smart Sensors for Real Time Drinking Water Quality and Contamination Detection in Water Distributed Mains," 2019. [Online]. Available: https://www.researchgate.net/publication/337733237
[19] Ibn-, "Smart Device to monitor water quality to avoid pollution in IoT environment."