Smart IoT and Machine Learning - Based Framework for Water Quality and Crop Prediction Using Raspberry Pi

Malavika S[1], Dr. Chandrappa D N[2]

[1, 2]Scholar, M. Tech VLSI Design, and Embedded System, SJBIT, India

[1, 2]Professor, Department of Electronics & Communication Engineering, SJBIT, India

Abstract: ***The water's quality plays a significant role in our daily lives. Forecasting water quality will be beneficial for reducing water pollution and safeguarding human health. A clever process for monitoring the water's quality instantly alerts the water analyzer when there is a problem. This method automatically determines the water's condition by interpreting sensor data using the Internet of Things. The development of machine - to - machine communication makes data analysis and communication easy and efficient. As an outcome of this effort, an "intellectual Based iot water quality system" to lakes has indeed been developed and is currently implementing in rural areas. The structure uses sensors for pH, turbid, and temperature to determine the properties of water quality like hydrogen ions and the entire amount of dissolved solvents. The yield prediction remains a major issue that needs to be addressed with the available data. Machine learning methods are a superior choice in this case. In agriculture, a variety of machine learning techniques are used and evaluated in order to forecast crop yield for following year. In this, a method for predicting agricultural productivity with historical data is put forth and applied. To achieve this, farm data are used to apply machine learning algorithms, such as Support Vector Machine and Random Forest that recommend fertilizer suitable for each unique crop.***

Keywords: Internet of Things, Water Quality, Raspberry Pi, Adafruit, Thingspeak Cloud, Machine Learning, Support Vector Machine,

1. Introduction

All living organisms have a big impact on water quality. Water quality examinations are a key component of environmental monitoring. Water quality refers to biochemical, physical, biological, & radioactive properties of water. This is a measure of the condition of the water in respect to the environment. The main variables which are measured and monitored serve to identify the requirements of one or more biotic species and any human need. Among the sensors used are the pH, IR, and temperature sensors. The numbers were all calculated. By analyzing the parameters to a threshold value, the purity is identified.

The water that is so vital to our survival is only a small portion of the total water available on earth. The global hydrological cycle states that 97 percent of water is always retained in the ocean. A mass of ice caps and glaciers is where about 1.998 % of the other 3 percent are trapped.0.85% of rainfall falls only on the ocean, while 0.152 percent of precipitation on land appears to be available to humans as drinkable water. However, just 0.002 percent of the freshwater remains as underground water in lakes/rivers & groundwater reserves, which could be used to meet demand for a variety of authorized best uses. Instead, 0.15 percentage of the 0.152 % goes back to the sea as surface run - off. It is applicable in water management system, smart irrigation, aquaculture etc

Categories	Area of the Basin (km2)	Count of Basins	Total Drainage Area Percentage	Flow Contribution as Percentage	Number of people living as in basin as a percentage of total population
Major	More than 20000	14	83	85	80
Medium	Between 2000 and 20000	44	8	7	20
Minor	Less than 2k	55	9	8	20

Figure 1: River basins of India

The majority of the country is covered with rivers. There are 14 major river basins in the nation, which account for about 83 percent of all watershed, 85% of all surface flow, and 80% of the population. The Brahmaputra, Ganga, and Indus are the three rivers. The Brahmaputra and Sabarmati rivers, the Brahmani, Tapi, Mahi, Subarnarekha, Pennar, and the Brahmani, Tapi, Mahi, Subarnarekha, Pennar Based on their drainage basins, these rivers were divided into three groups.

2. Related Work

Benedict (2020) suggested IoT applications for cloud production, directed water quality or air quality monitoring, energy - conscious societal applications, and smart agricultural economics, all of which are built with a mix of high - end computing technologies like cloud, edge, and fog. Thousands of people, including developers, benefit from IoT technologies that are implemented in an automated/ decentralized approach with improved security controls, according to smart cities and regulatory authorities. Owing to the lack of appropriate technology, such as server less computing, existing IoT architectures are vulnerable to energy inefficiency or resource underutilization issues.

Olatinwo and Joubert (2019) looked at some of the newest wireless innovations that could be used in upcoming wireless sensor network applications for water quality control soft - ware. The innovations under consideration promise to solve long - standing problems with current wireless sensor network systems for tracking water quality parameters. Energy conservation and long - range water quality data exchange are examples of such problems. These flaws open the door for the use of newly emerging technologies found in this study to advance the field of water quality monitoring through a wireless sensor network. To do this, this paper recommends three main types of contact networks, namely architectural architecture and network implementation for water quality monitoring applications.

Machine learning is an emerging research field in crop yield analysis. Yield prediction is a very important issue in agriculture. Any farmer is interested in knowing how much yield he is about to expect. In the past, yield prediction was performed by considering farmer's experience on particular field and crop. The yield prediction is a major issue that remains to be solved based on available data. Machine learning techniques are the better choice for this purpose. Different Machine learning techniques are used and evaluated in agriculture for estimating the future year's crop production. This paper proposes and implements a system to predict crop yield from previous data. This is achieved by applying machine learning algorithms like Support Vector Machine and Random Forest on agriculture data and recommends fertilizer suitable for every particular crop. The paper focuses on creation of a prediction model which may be used for future prediction of crop yield. It presents a brief analysis of crop yield prediction using machine learning techniques.

3. Methodology

1) Software Methodology

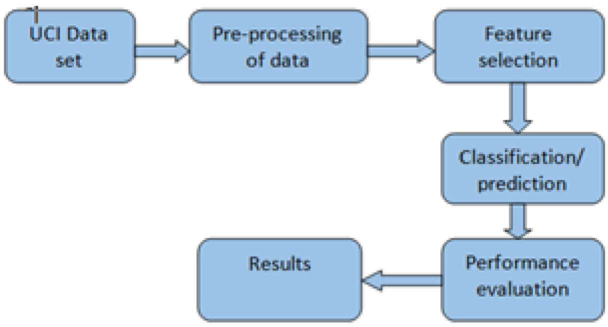

Figure 2: Block Diagram of Software Methodology

Artificial Neural Network

Multi - layered Preceptor artificial neural networks are defined as biologically inspired models that can represent incredibly complex non - linear functions. One of the key components of machine learning is the use of ANNs. These are neurologically systems that aim to mimic how humans learn, as the name "neural" suggests. Three layers—input, output, and hidden—make up a neural network. The units that make up a hidden layer often turn the input into patterns that the output manipulates.

Random Forest

The most well and potent artificial intelligence algorithms are Random Forest. Bootstrap Aggregation, often known as bagging, is a type of machine learning technique. The bootstrap is an extremely effective statistical method for estimating a value from a sample of data, such the mean. In order to provide a more accurate prediction of the actual mean value, many samples of data are

Simple Logistic Regressions

A machine learning method that draws from statistics is called logistic regression. This technique can be applied to binary classification, where two classes are used to distinguish between values. Similar to linear regression, the aim of logistic regression is to determine overall regression coefficients within each input variable. In contrast to linear regression, a non - linear function known as a logistic function is used in this case to forecast the outcome. Any value between 0 and 1 is transformed using the logistic function.

Support Vector Machine

SVM is a common Supervised Learning method. SVM selects the highest points/vectors that aid in the creation of the hyper plane. These extreme examples are referred to it as support vectors, and also the method is known as the Support Vector Machine.

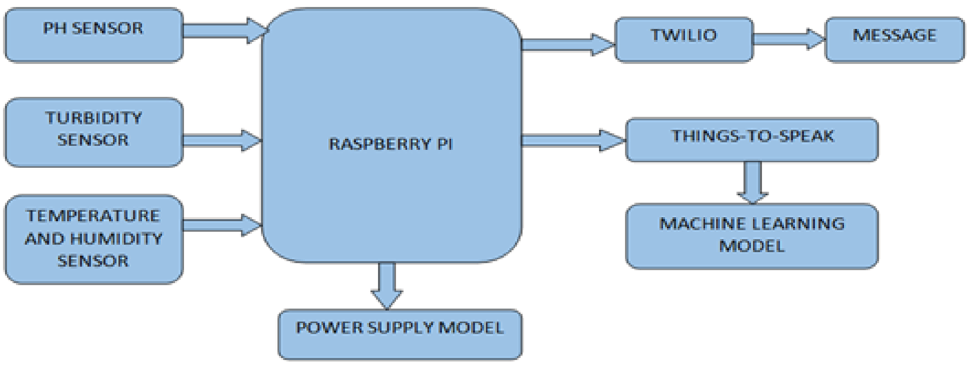

Figure 3: Methodology

The Wi - Fi module will send the collected data to the cloud, and the sensor information could be seen on ESP8266 Wi - Fi module Users can learn about water scale since the values supplied to them are arranged according to the immense level for each parameter. The sensor values checked via cloud will be revealed via a web application, with alerts given to the user if the value exceeds the threshold value. The pH of drinking water should be between 6.6 and 8.5, the turbid should be 1 - 5, and the temperature should be between 50 and 72 degrees.

4. Results and Discussion

By running Code to read sensor data from the hardware, the findings that are displayed below were achieved. The software interfaces with Raspberry Pi model 3 and uses an external ADC to transport data. System libraries are smuggled for system timing data; in this case, they are being used to configure the general purpose input output pin connections. Analog to digital converters employ the serial peripheral interface, which has been initialised. In the DHT II humidity sensor, temperature and humidity sensor values are generated. These values are divided and recorded, and the results are achieved

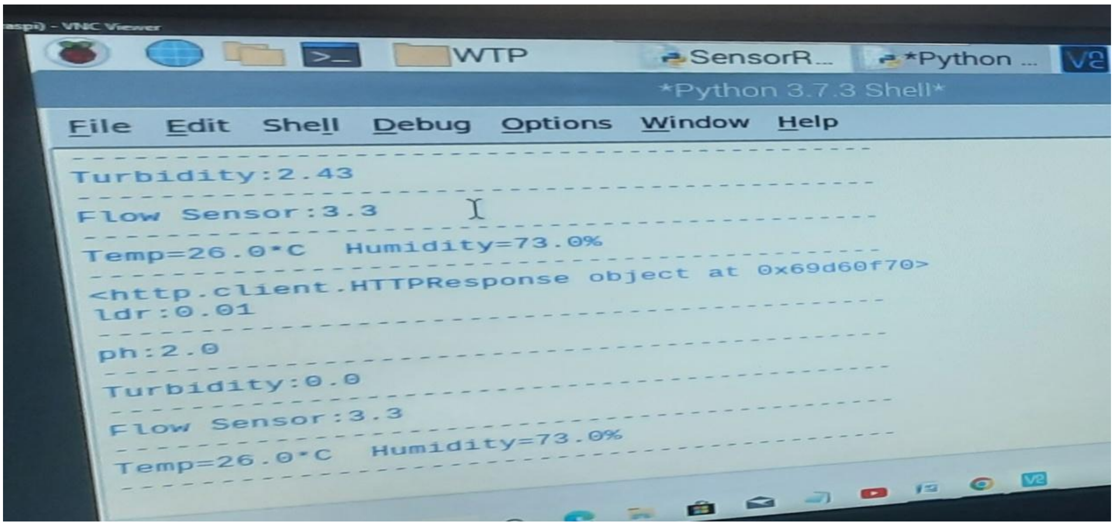

Figure 4: Sensor Results Using Python

A channel ID is generated in the things - to - speak cloud, and channel numbers (0, 1, and 2, n) are provided as pin numbers for the sensors used in the hardware. This is the read channel, once the cloud functions have read the sensor values, the function begins. The field number from thingspeak is read, provided to the Python code, and mentioned with an argument. This will be connected to the specific sensor location & area number and its values will be updated and executed, and results have been obtained in graphical form. SPI formula is then used to read values, verify it, & return the result and stored.

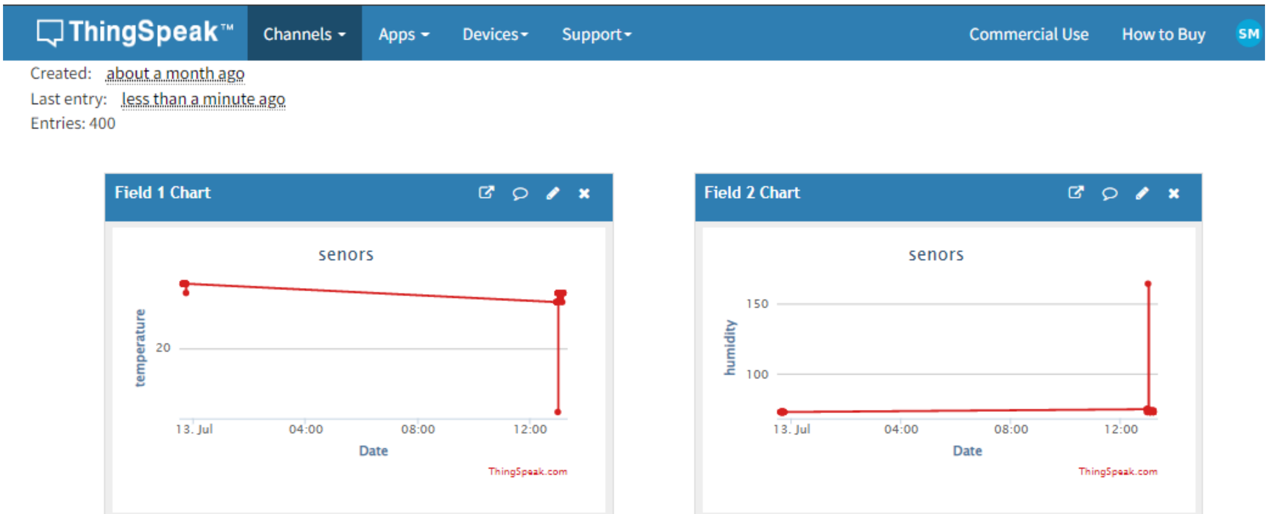

Figure 5: Graphical Representation of Temperature and Humidity

Figure 6: Graphical Representation of Turbidity and Flow

The results of crop predicted using machine learning is shown in below figures The outcome displayed below was produced by fusing machine learning with sensor information downloaded from the cloud. To use the pre - defined functions, pandas, numpy, matplotlib, and SK learn are imported. Here, pandas are used to read a CSV file for crop recommendation data handling, numpy is used for array handling, and matploplip is used for data analyzing. As the results show, the nitrogen, potassium, phosphorous, temperature, humidity, ph, and rainfall values are analysed and executed to give the types of crops recommended with precision and support values.

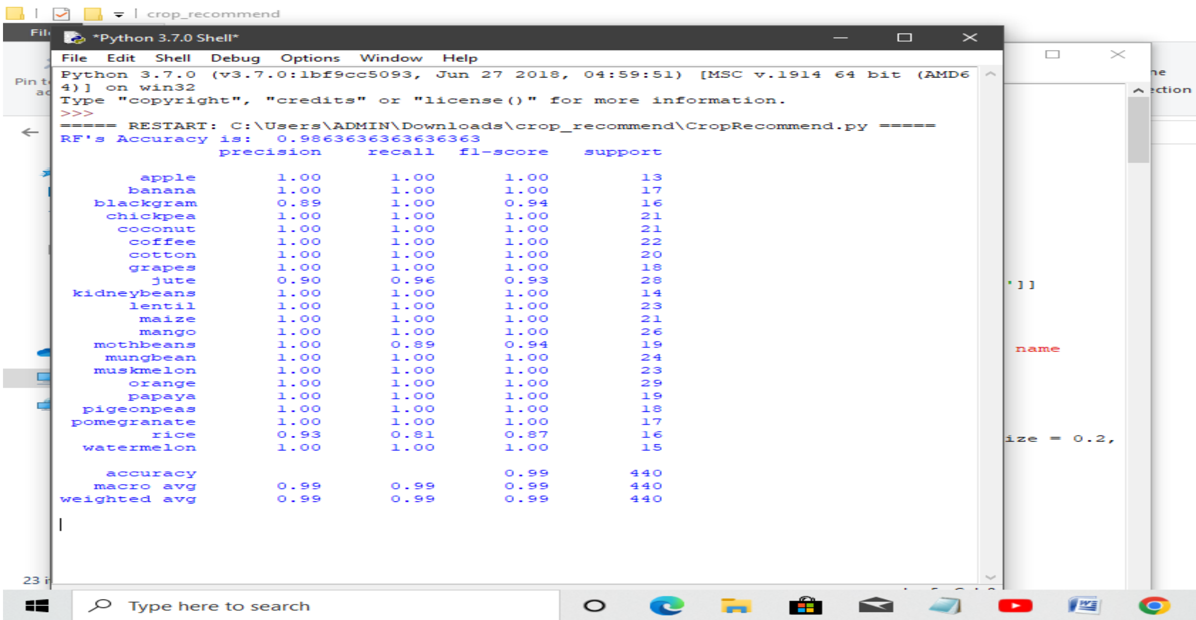

Figure 7: Results of RF's Accuracy

The accompanying figure demonstrates the obtained pH, temperature, and humidity data after integrating them with a machine learning algorithm. The crop anticipated from the data obtained is an apple, as seen by the findings below.

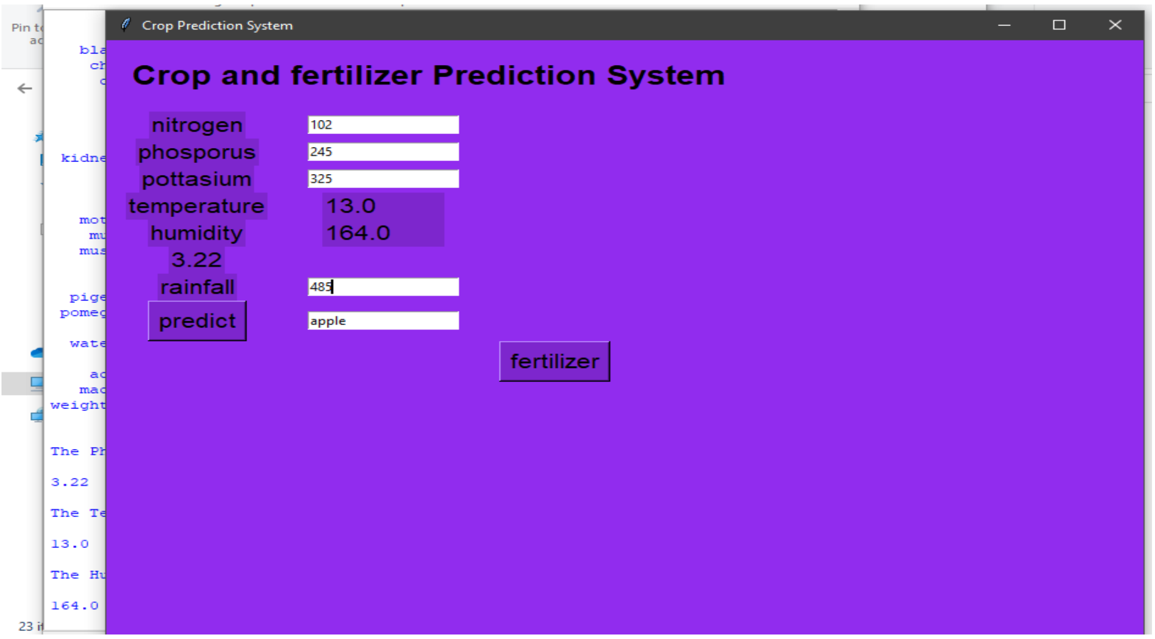

Figure 8: Crop Prediction using ML

Furthermore, total amount of nitrogen, phosphorus, and potassium needed to grow the sort of crop that was obtained is anticipated using the data from the sensor values.

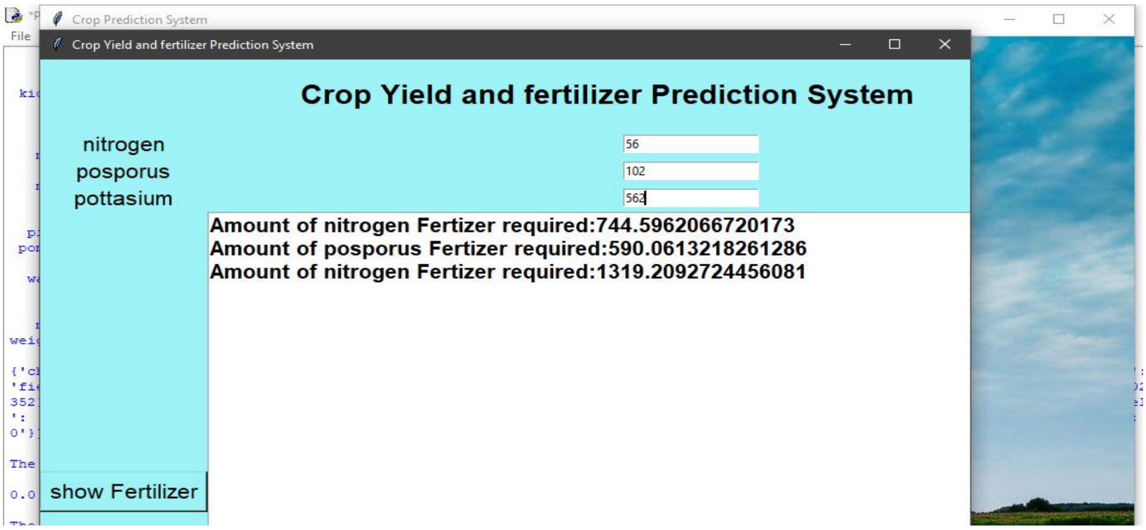

Figure 9: Fertilizer Prediction Using ML

5. Conclusion and Future Scope

This technology can automatically monitor water quality. The designed system, which uses a Raspberry Pi board, has successfully interfaced with a number of sensors. To track water quality, an effective algorithm is built in real time. The method is very adaptable. The gadget was designed to be low - cost since it employs hardware to measure four water quality metrics (temperature, pH, turbidity, and flow) utilising IoT sensors. As a result, water quality testing will most likely be more cost - effective, convenient, and quick. The above setup could be used to analyze different water quality metrics by simply replacing the matching sensors and altering the required software packages. The procedure is straightforward. The system can be expanded to track hydrologic, air pollution, industrial, and agricultural output, among other things. It has a large number of applications and extensions. It has a wide range of applications and extensions. Through the web server, a web - based application called Thing - to - Speak is used to check metrics including pH, turbidity of the water, and amount of liquid in the tank, temperature, and moisture of the surrounding atmosphere. Higher agricultural yields can be produced, as evidenced by the location - based crop yield prediction and effective algorithm implementation. According to the research mentioned above, Random Forest and Support Vector Machine are effective for classifying soil. Support Vector Machine is effective for crop yield prediction.

To get more accurate, dependable findings, the capabilities of the system for monitoring water quality might be increased. The system is expandable to monitor climatic conditions, environmental damage, industrial and agricultural output, and other variables. Additional functionality that could be added to the work is as follows.

To assist farmers, mobile applications that upload images of farms can be created. Employing image analysis to detect crop diseases and providing pesticides to users based on disease photos Increase agricultural output by implementing a smart irrigation system.

References

[1] R. Pitchai and Richard f. hill, Design of a Synoptic Water Quality Monitoring System for Narragansett Bay (1972), pp.17 - 24

[2] Pedro M. Ramos, J. M. Dias Pereira, Helena M. Geirinhas Ramos and A. Lopes Ribeiro, a Four - Terminal Water - Quality - Monitoring Conductivity Sensor (2008), pp.577 - 583

[3] N Vijayakumar and R Ramya, the Real Time Monitoring of Water Quality in IoT Environment (2015)

[4] Prashant Salunke and Jui Kate, Advanced Smart Sensor Interface in Internet of Things for Water Quality Monitoring (2017), pp.208 - 302

[5] Monira Mukta, Samia Islam, Surajit Das Barman, Ahmed Wasif Reza, M Saddam Hossain Khan, IoT based Smart Water Quality Monitoring System (2019), pp.669 - 673

[6] Sonali S. Lagu; Sanjay B. Deshmukh Raspberry Pi for Automation of Water Treatment Plant

[7] S. S. Lagu and S. B. Deshmukh, ―Raspberry Pi for Automation of Water Treatment Plant, ‖ Informatics, 2014 Int. Conf. Adv. Comput. Common. pp.1999–2003, 2019.

[8] Y. N. Santoso, H. Wicaksono, P. Santoso, and J. Siwalankerto, ―Build Automation Model By Using SCADA System Internet Based _System SCADA Berbasis Internet untuk Model Otomasi Bangunan, '‖ J. T. Elektro, U. K. Petra, J. Dimens Tek. Elektro, vol.1, No.1, pp.18–23, 2018.

[9] Irfan Jamil, ―Technical Communication of Automation Control System in Water Treatment Plant‖, International Journal of Innovation and Applied StudiesISSN 2028 - 932 2020

[10] S. Marsili - Libelli, G. M. Maietti, ―Energy - saving through remote control of a wastewater treatment plant, ‖ Proc. SIDISA Conference, Florence, June 24 - 28, 2018.