Hindawi
Journal of Sensors
Volume 2022, Article ID 9042382, 17 pages
https://doi.org/10.1155/2022/9042382

Research Article

Agriculture Field Automation and Digitization Using Internet of Things and Machine Learning

Sushabhan Choudhury,[1] Rajesh Singh,[2,3] Anita Gehlot,[2,3] Piyush Kuchhal,[1] Shaik Vaseem Akram,[2,4] Neeraj Priyadarshi,[5] and Baseem Khan[3,6]

[1]University of Petroleum and Energy studies, Dehradun, Uttarakhand, India
[2]Department of Research and Innovation, Uttaranchal Institute of Technology, Uttaranchal University, Dehradun 248007, India
[3]Department of Project Management, Universidad Internacional Iberoamericana, Campeche C.P. 24560, Mexico
[4]Law College Dehradun, Uttaranchal University, Dehradun 248007, India
[5]Department of Electrical Engineering, JIS College of Engineering, Kolkata 741235, India
[6]Department of Electrical and Computer Engineering, Hawassa University, Hawassa, Ethiopia

Correspondence should be addressed to Baseem Khan; basseemk@hu.edu.et

Received 21 May 2022; Revised 18 September 2022; Accepted 27 October 2022; Published 12 November 2022

Academic Editor: Qiang Wu

The real-time smart monitoring with intelligence highly gained significant attention for enhancing the productivity of the crop. Currently, IoT generates a lot of real-time data from the sensors, actuators, and identification technologies. However, extracting the meaningful insights from the data is necessary for realizing the intelligent ecosystem in agriculture. Based upon the previous studies, it is also identified that the limited studies have merely implemented machine learning (ML) on real-time data obtained through customized hardware with dedicated server. In this study, we have proposed a customized hand-held device that enables to deliver recommendations to the farmer on the basis of real-time data obtained through IoT hardware and ML. A three-layer structure is proposed in the study for realizing custom hardware with 2.4 GHz ZigBee and IoT sensors for the data acquisition, communication, and recommendation. As a part of real-time implementation, the calibration of the sensors is processed to form a real-time dataset with precision. The study evaluated four ML models and concluded that XGBoost has shown a better accuracy on the proposed dataset. The XGBoost recommended the crop based on selected parameters. The developed hand-held device can be customized with advance features with crop recommendations.

1. Introduction

Agriculture is the key to human survival, as it is the main source of grain and other basic resources. In addition, agriculture accounts for about 4% of the world's gross national product (GDP) [1]. Urbanization and population growth in 2050 conclude that food production must be sustainably doubled with minimal water resources [2]. About 97% of the water on the planet is salty, while the remaining 3 is freshwater [3]. Agriculture uses 70% of freshwater for irrigation in most developing countries [4]. Therefore, the efficient use of freshwater during irrigation is the most significant issue in terms of cost reduction and yield improvement. Using the traditional method, farmers manually check and regulate the availability of water resulting in a 50% water loss [5]. However, different irrigation techniques like drip irrigation, sprinkle irrigation, and furrow irrigation have minimized water wastage by 30-70% [6]. Yet, the optimal management of water content in the soil is not yet achieved with these irrigation techniques as overwater usage in the agricultural field leads to an overflow of nutrients from the soil [7].

Agriculture also requires adequate levels of fertilizers and pesticides, as farmers use fertilizers and pesticides while neglecting the optimal needs of the crop. Apart from that, water availability, nutrient levels, and soil moisture are some other factors that affect crop productivity as well. With the traditional approach, it is a challenging task for the farmer to determine water availability, nutrient levels, and soil moisture and also to identify which factor is fighting food production [8]. Here, smart and real-time systems help to monitor the various parameters of the agricultural field and effectively control the water level with other resources to increase productivity [9]. Nowadays, real-time monitoring and intelligent systems are possible with the Internet of Things (IoT), as IoT monitors agriculture with IoT-enabled sensors and communication protocols implemented in agriculture [10]. The above facts conclude that the IoT has played crucial role in multiple areas including agriculture with its sensing, communication, and real-time monitoring features through the IoT hardware. Moreover, it is concluded from the previous studies that the ML model delivered significant results on real-time data [11]. The study framed a research question "How and for what purpose may the ML technology be used in agriculture on IoT real-time data? On the basis of this research question, this study has carried out the literature review.

1.1. Literature Review. Large-scale agricultural monitoring applications require reliable WSN networks because the maximum number of sensors is operated over a long period. A wireless sensor network architecture for vegetable greenhouses is presented to achieve scientific cultivation and minimize management effort from an environmental monitoring perspective [12]. Smart greenhouse management systems and WSNs are used to control and monitor agricultural parameters and activities in greenhouses autonomously [13, 14]. A smart greenhouse information monitoring system records environmental factors with ZigBee wireless sensors [15].Wi-Fi-based smart WSN has been proposed to monitor the agricultural environment, and the system allows intelligent monitoring of agricultural conditions [16, 17]. WSN's efforts recommend that sensor data can be collected and sent to the main server [18]. A WPAN-based water quality monitoring system has been proposed to clean up and collect real-time sensor data on agricultural land with the LabVIEW data logger [19]. In addition to WSN, the advent of IoT technology has enabled farmers and technologists to solve the challenges farmers face, such as water shortages, cost control, and productivity issues [20].

A scalable IoT and WSN architecture is proposed for remote monitoring and control of agriculture [21]. For the same, an energy-saving ZigBee sensor network with bidirectional communication and end devices is integrated to deliver data from the sensors to the PC at variable times determined by the central node [22]. SiloSense is a unique architecture based on ZigBee to monitor the storage conditions of grain silos to protect them from spoilage and disease [23]. The most important parameters that are required to monitor while growing wheat and other vegetables are soil moisture, ambient temperature, air pressure, and sunlight intensity [24]. mIoT- and WSN-based agricultural system are for monitoring air, temperature, soil moisture, and humidity with RF modules [25]. The cloud-based and IoT-based smart irrigation system is designed to obtain data on soil moisture, soil health, and temperature to reduce water consumption [26, 27]. IoT-based greenhouse agriculture is implemented to monitor climatic conditions and to obtain data on a cloud server for analysis, while the ZigBee protocol and the Wi-Fi module are integrated [28, 29]. The IoT-based framework is designed to perform data analysis using real-time data to increase productivity on farms through temperature, soil moisture, and humidity sensor [30, 31].

IoT-enabled plant disease and pest prediction system is implemented to reduce the use of insecticides and fungicides, and additionally, an assessment of meteorological data is also carried out to identify the correlation between pest growth and climate [32]. ML helps to examine and analyze data from different fields of agriculture to improve crop yields and offers different analytical techniques to predict the yield of crop and plant disease [33, 34]. The ML-based predictive model helps farmers to get the right harvest with unconditional weather behavior [35]. ML algorithms such as neural network-based models are used for predictive analysis purposes [36]. From the literature review, it is identified that smart monitoring in the agricultural field needs to carry out with advanced technologies such as IoT. The literature also concludes that they are limited studies that analyzed the accuracy of the data obtained through sensors. In addition to this, customization of hardware on the basis of agricultural field requirement is limitedly explored by the previous researchers. ML technique is applied on different datasets for disease detection, environmental parameters monitoring, and automation in irrigation, but the previous studies have yet to explore the crop recommendation on the basis of real-time data. However, implementing ML on real-time data primarily requires a resource-constrained and dedicated server to fulfill the task of acquiring real-time agricultural data rather than using it for multiple purposes.

To overcome this research, the gap of this study implements customized IoT hardware for obtaining the real-time field data through wireless personal area network (WPAN) and wireless local area network (WLAN). WPAN enables to minimize the power consumption and also transmits the sensor data reliably and securely. WLAN is used to connect the customized hardware to the cloud server through IP protocol. To maintain accuracy in the data, the calibration methods are applied on the sensors. It is also identified from the previous study that IoT requires analytical techniques to provide intelligent decisions based on real-time sensor data obtained from IoT sensors [11]. The contributions of the study are as follows:

(i) Customized hardware for sensor node, master node, and hand-held device with ZigBee RF modem is designed for sensing the real-time data of agricultural field including temperature, humidity, soil pH, and water level

(ii) An interference test is implemented to verify that the ZigBee signal is not interfering with other signals on the same frequency band

(iii) To enhance the security of data transmission, the symmetric encryption approach with a private key is applied by leveraging XXTea encryption functions

(iv) A cloud server is developed to log the real-time sensor data of the agricultural field

(v) The pretrained machine learning model is applied to real-time data such as temperature, humidity, rainfall, and soil pH sensor on the cloud server for seasonal crop recommendations

The organization of the study is as follows: Section 1.1 covers the proposed system. Section 2 covers the circuit diagram for the development of the system. Section 3 covers simulation analysis and calibration. Section 4 covers the real-time implementation of developed nodes and the current consumption analysis. The article is concluded in the final section.

2. Proposed Architecture

To realize the main objective of crop recommendation, in real-time, machine learning is utilized. This system intends to leverage multiple sensors with real-time data collection such as temperature, humidity, rainfall, and soil pH sensor to improve the efficiency and the recommendation of crops, with the support of machine learning techniques. As shown in Figure 1, the architecture is divided into multiple modules, with both software and hardware parts, each with its purpose, from the hardware nodes, for data collection; the server, for data processing; and finally, the hand-held device to the user.

The system is composed of various environmental sensors that are deployed around the agricultural field to collect data on a variety of characteristics such as water level, temperature, soil pH, and humidity. This information is then sent to the master node, the main hub of our system, which is in charge of communicating with the developed cloud server and transferring the data obtained from the sensors. In the server, the information is stored and is also run through the ML algorithm to be studied. From there, based on the algorithm's analysis, the information is presented to the user on a handheld device, such as crop recommendation. The crop recommendation feature is activated from the hand-held device, and recommendations are received on the hand-held device based on the user's request.

2.1. Data Acquisition Layer. The data acquisition layer is the primary layer of the architecture for acquiring the environmental parameters of the agricultural field. This layer is specifically dedicated to continuously monitor the environmental parameters including water level, temperature, humidity, light intensity, and rain level. For this, IoT sensors are attached to the sensor node as shown in Figure 2 and through 2.4 GHz ZigBee communication [37], it transmits data to the master node. In addition, the sensor node is enabled with security, interference technique, and node mapping feature for secure and reliable communication.

2.2. Data Processing Layer. IoT devices and sensors record the environmental parameters of the agricultural field in real time. Sensory data processing is done with the data processing layer. The master node, which consists of the ZigBee RF module, receives the data from the data acquisition layer and transmits it to the cloud server via a Wi-Fi module (Figure 3). The master node is powered by the battery power supply. The data logger is also available in the data processing layer for visualizing the sensor data through Bluetooth. The master node with a Wi-Fi module connects to the Internet to log the sensory data in the cloud server.

2.3. Analytics and Visualization Layer. Figure 4 illustrates the system logic, where it explains the preprocessing and analysis of gathered data from the sensor node. The sensor data is received at the master node which is converted into a set of scripts and is logged into the cloud server. Here, a cloud server is also developed to log the sensor node values. In the preprocessing step, the sensor data is converted into a suitable format for performing machine learning analyses. The data is fed to a machine learning model and based on the data the model, it suggests recommended crop. The outcome of the model is stored in the cloud server, and the hand-held device that connects to the cloud server also visualizes the recommended crop and real-time sensor data. To provide the user with a way to see the data collected from the sensors and machine learning outcomes, a cloud server is developed.

A cloud server is developed with two different API protocols for data exchange between backend and frontend and backend and custom gateway. The interface between the backend and the frontend for the exchange of data is implemented with the REST API. REST APIs use HTTP requests to perform basic database activities within a resource, such as creating, reading, updating, and deleting records. REST APIs accept JSON for the request payload and send responses to JSON. JSON is the standard for data transfer. The interface between the backend and the gateway is implemented using the MQTT protocol. The cloud server assists in checking all the sensor values retrieved from the agricultural field in real time. Moreover, a hand-held device is connected to the cloud server for providing the updates of the crop based on request generated by the user.

3. Hardware Development

In this section, we present the schematic diagram of the sensor node that is primarily implemented for data acquisition. Moreover, the customization is carried out for the hardware of agricultural monitoring including data logger and hand-held device. A detailed description of the sensor node and the hand-held device is presented below.

3.1. Sensor Node. Figure 5 illustrates the connection of different electronic components of the sensor node. The node is composed of a 2.4 GHz RF modem that works in full-duplex mode to transmit the data from one node to another connected in the network. Principally, two sensors (gas sensor and humidity sensor) are connected to the node which are placed in the field. The function of the gas sensor is to measure the concentration of the gas (such as LPG and butane) in the present environment. The humidity of the environment is detected with the help of a humidity sensor. The humidity sensor gives serial data at the 9600 baud rate to the microcontroller. The output of these sensors is in analog form. Analog signal from the sensor is fed onto the analog to digital pin of the ATMEGA 8 [38].

The function of this pin is to convert the analog signal into a digital signal. After fetching the signal from a sensor

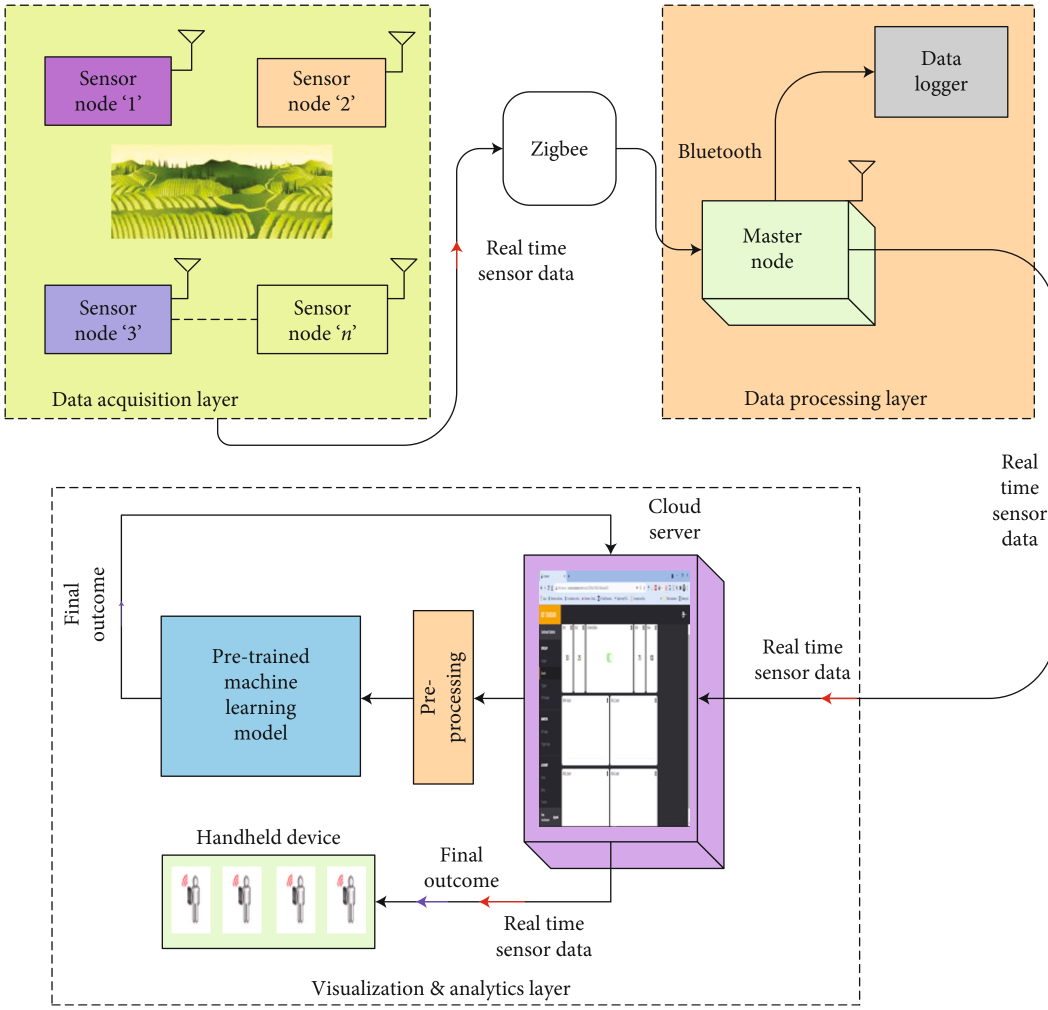


Figure 1: Proposed architecture.

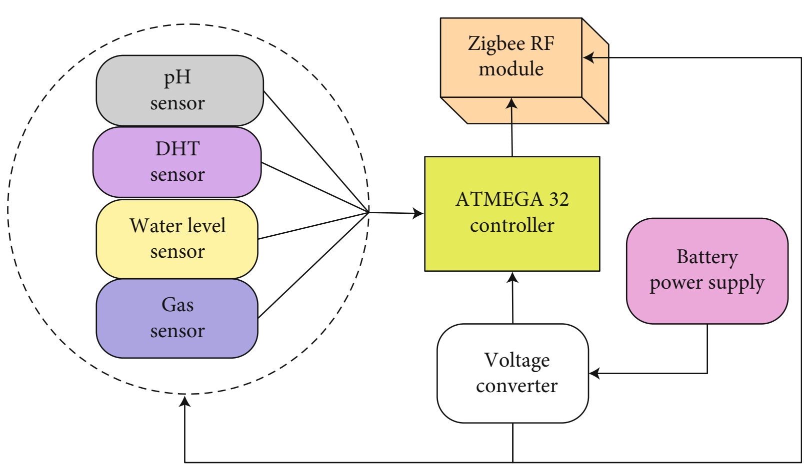


Figure 2: Sensor node.

attached, the microcontroller performs the logical operation and controls the overall operation. A display (16*2 LCD) is attached to each node. The function of this is to display the measured parameters. MAX 232 IC is also used to provide communication between the node and PC. With the help of this, the end-user can easily observe the parameters and execute the needful action. The whole circuit is working on the 5 V power supply. The RF module is attached to the

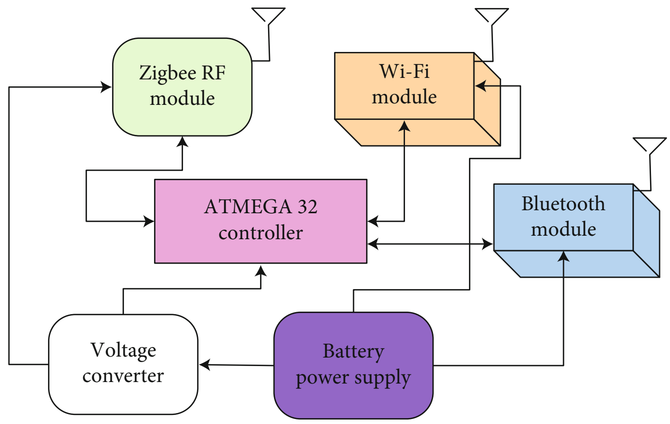


Figure 3: Master node.

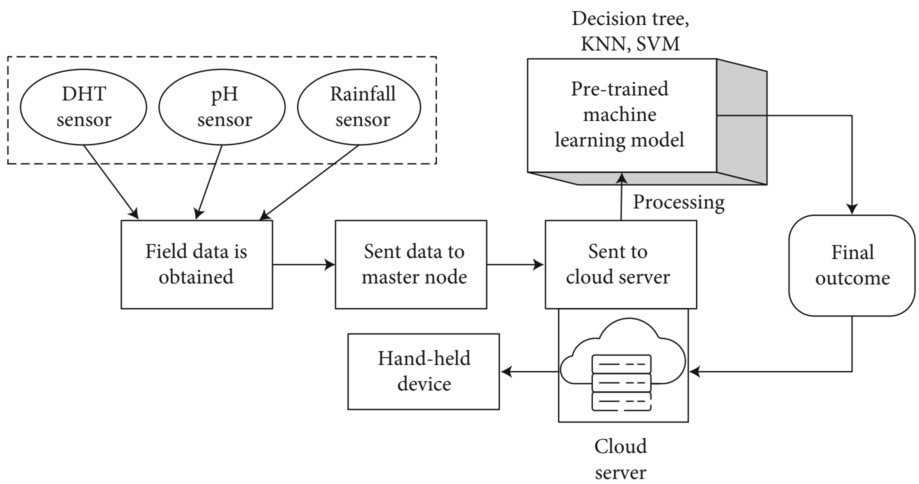


Figure 4: System logic.

transmitter and receiver pin of the microcontroller. A ZigBee module (CC2500) [37] is used in the system to provide wireless communication in the network. This module works in full-duplex mode and also can communicate with many devices at the same time. The topology of each node in the presented network is based on mesh topology. All the required components are integrated and developed hardware of the sensor node as shown in the Figure 6. The sensor node is integrated with LCD to visualize the sensor values.

3.2. Hand-Held Device. Figure 7 shows the circuit diagram of a hand-held device that consists of a Bluetooth modem. The main function of this modem is to receive the data transferred from the data logger node. This node is at the end-user and easily provides the collected data to the end-user. This node is connected with other nodes in a mesh topology. A display is used for showing the measured data. Based on this data, end-user can easily take the decision. Figure 8 illustrates the hardware of the hand-held device, and the hand-held device is also integrated with display to visualize the data like real-time sensor value and crop recommendation.

3.2.1. Current Consumption Analysis of Hardware. Current consumption analysis is performed based on the current required by each component used to develop the system. The power consumption by the ZigBee RF modem is maximum for each node, followed by consumption by the microcontroller. Even though power was not a design issue for the system, it is evident that the designed system requires less power than standard available devices like Mica2 and Micaz nodes. The current consumption analysis concludes that the designed hardware is consuming less amount of power during the data transmission. Specifically, the hand-held device is only consuming a current of 79 mA. The current consumption may vary when the components in the hardware are increased. From Table 1, it is concluded that the power consumption of node 1 is 85 mA, node 2 is 85 mA, node 3 is 264 mA, node 4 is 104 mA, and node 5 is 83 mA. To meet the power requirement of the sensor node, an energy harvesting system will be integrated into the sensor node in the future. In agriculture, solar panels will be used to implement solar energy harvesting systems on sensor nodes. Table 2 shows the current consumption of the hand-held device, and it is 79 mA. The power to the hand-held device is achieved with the battery.

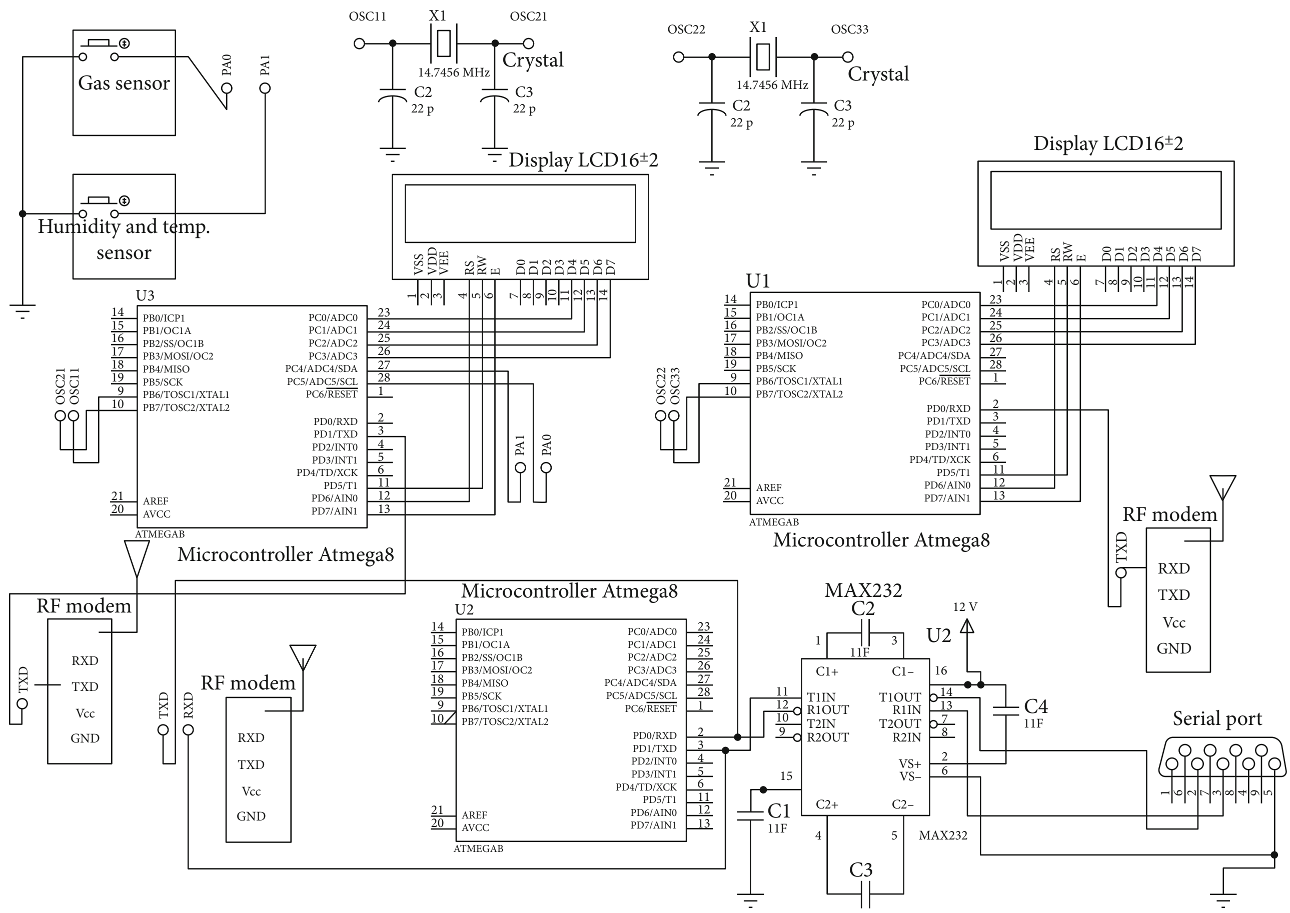


Figure 5: Circuit diagram for the sensor node.

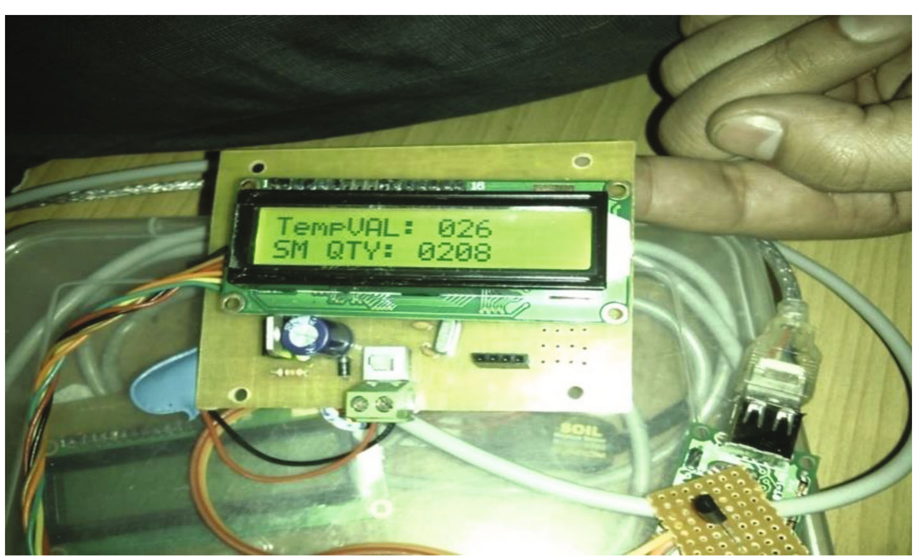


Figure 6: Hardware of sensor node.

3.3. Calibration of Sensors. In this section, sensor calibration is performed to set the sensor to operate accurately and without error. This section describes the sensor calibrations (temperature/humidity, soil moisture, and ultrasonic) used to develop the system. Sensor calibration is an important step, before the actual implementation of the sensor in the system. For the developed system, each sensor is first calibrated with standard instruments and after checking its accuracy, sensors are used in the system. For calibrating soil moisture sensors, the oven method is used. For temperature/

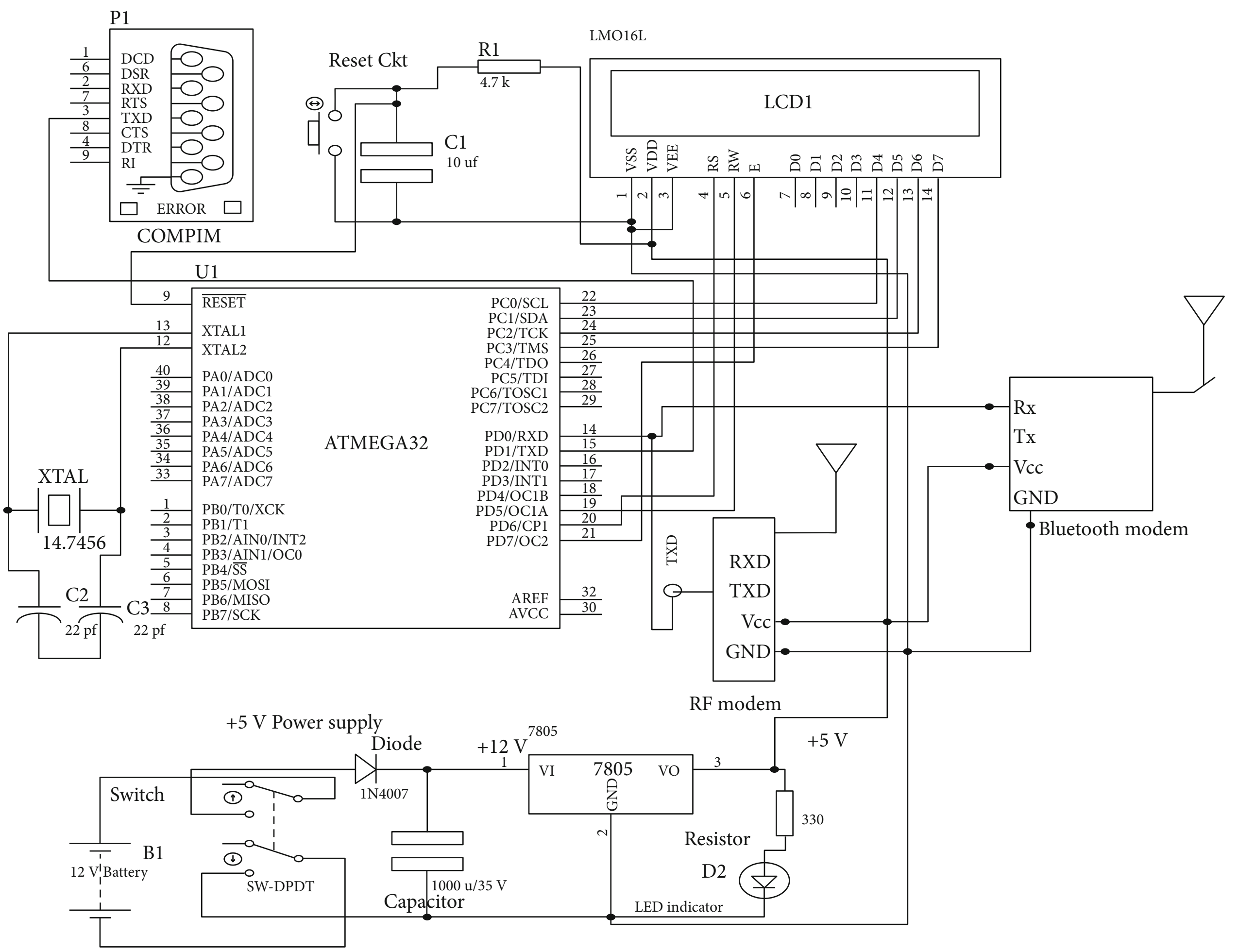


Figure 7: Circuit diagram of the hand-held device.

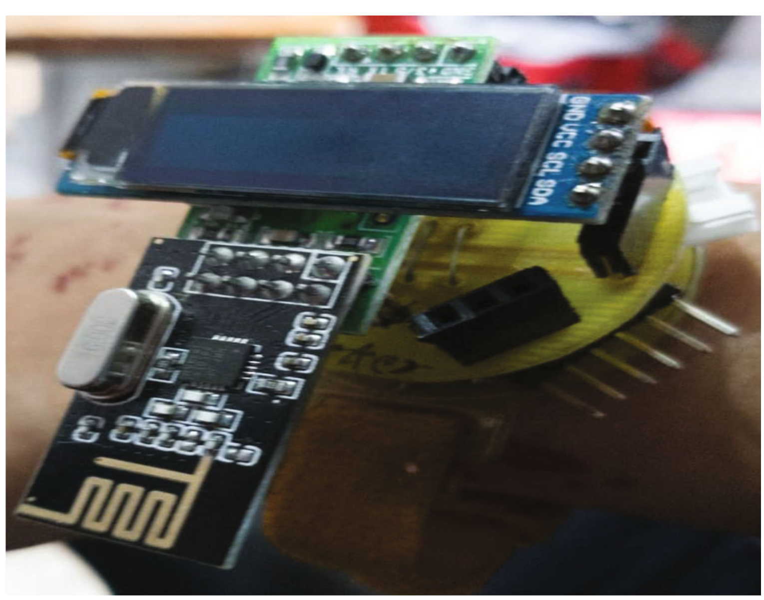


Figure 8: Hardware of hand-held device.

Table 1: Current consumption analysis of sensor nodes.

Components	Node 1 (mA)	Node 2 (mA)	Node 3 (mA)	Node 4 (mA)	Node 5 (mA)
Microcontroller	17	17	17	17	17
RF modem	58	58	58	58	58
Sensor 1	1	1	1	1	1
Sensor 2	5	5	180	20	3
LCD (16 × 2)	4	4	4	4	4
Total (mA)	85	85	264	104	83

Table 2: Current consumption analysis of hand-held device.

Component	Current (mA)
Microcontroller	17
RF modem	58
LCD (16 × 4)	4
Total (mA)	79

Table 3: Calibration of soil moisture sensor with the standard method.

Soil moisture reading	Standard value (%) (oven method)
100	39
120	46.8
150	58.5
200	78
250	97.5
255	100

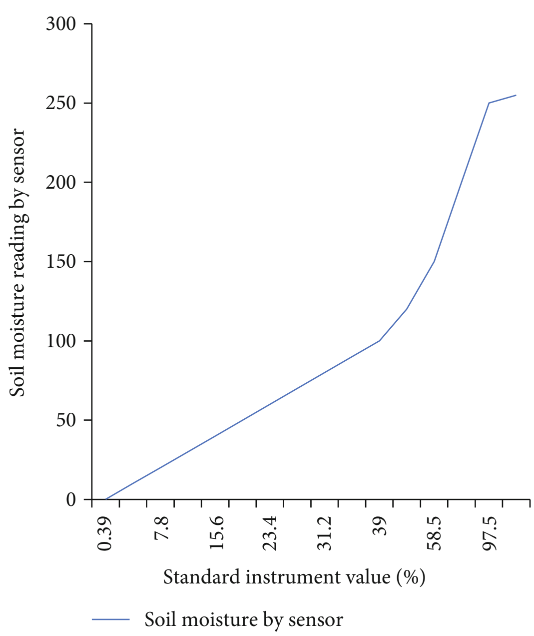


Figure 9: Calibration of soil moisture sensor w.r.t. standard instrument.

humidity sensor calibration, psychrometric is used. The ultrasonic sensor is calibrated with a standard measurement ruler. Tests were carried out for each sensor and compared with standard instruments, and it was observed that the sensors were calibrated appropriately as follows.

3.3.1. Soil Moisture Sensor. The soil moisture sensor is taken from the Sunrom model (http://sunrom.com/) no. 1282 [39]. The sensor gives a reading in terms of numerical values, as can be seen in Table 3. This sensor measures moisture based on the volumetric water content in the soil. On the other hand, the standard oven method for finding soil moisture gives a reading in percentage. Initially, the sensor readings are mapped by the oven method [40] to check the error between the sensor value and instrument value. Figure 9 shows the relation between soil moisture content shown by the sensor and by standard instrument. It provides calibration values of the sensor concerning the percentage value of the oven method. To validate the calibration, repeated experiments were conducted by taking soil samples from five different places with valid water content and calibrated them with the standard instruments. Table 4 shows the sample readings by soil moisture content (%) with sensor values. The sensor reading from the sensor is received by an analog-to-digital converter (ADC) and converted to a % value using equation (1). If the ADC value of the sensor is 255, then the value is considered 100% accurate concerning the standard oven method.

$$\text{Soil moisture}(\%) = (\text{wet soil (g)-dry soil (g)}) * \frac{100}{\text{dry soil (g)}}. \tag{1}$$

3.3.2. Ultrasonic Sensor. An ultrasonic sensor [41] is calibrated with a standard scale. Table 5 clearly shows the bias value between sensor value and scale which is 3 cm in each case. So, this biasing is managed by programming the microcontroller accordingly. It shows a constant bias value by the sensor, which is eliminated with the help of the microcontroller program. After calibration, the sensor shows accurate readings.

3.3.3. Temperature/Humidity Sensor. Temperature/humidity is calibrated with a psychrometer [42]. It is achieved by mapping the values of the sensor with that of the standard instrument. The values are adjusted using programming the controller. The readings are taken every half an hour in March 2020. As shown in Table 6, readings are taken, and

Table 4: Soil moisture sensor and reading by standard method for different samples.

(a)

Samples	Sample 1 Wet soil (g)	Dry soil (g)	Soil moisture	Sample 2 Wet soil (g)	Dry soil (g)	Soil moisture
Standard instrument reading	114	98	16%	62	52	19.23%
Sensor reading			42			49

(b)

	Sample 3 Wet soil (g)	Dry soil (g)	Soil moisture	Sample 4 Wet soil (g)	Dry soil (g)	Soil moisture	Sample 5 Wet soil (g)	Dry soil (g)	Soil moisture
Standard instrument reading	68	50	36%	72	52	38.46%	78	53	47.16%
Sensor reading			93			95			123

Table 5: Calibration of ultrasonic sensor with standard instrument.

Distance by standard scale (cm)	Distance by ultrasonic sensor (cm)	Bias (cm)
10	13	3
20	23	3
30	33	3
40	43	3
50	53	3

Table 6: Calibration of temperature/humidity sensor with a standard instrument.

Dry temperature by standard instrument	Wet temperature by standard instrument	Relative humidity by a standard instrument in %	Temp. sensor (Sunrom 1211)	Humidity by the sensor in % (Sunrom 1211)
18	10.8	47.3	18	47
19	11.5	46.1	19	46
20	12.1	45	20	45
21	12.7	44	21	44
22	13.4	42	22	43
23	13.8	41	23	41
24	14.1	42	24	42

the sensor is calibrated accordingly. For calibration of the sensor, a psychrometer is used as per the calculation given in [43].

3.4. Security and Interference Test

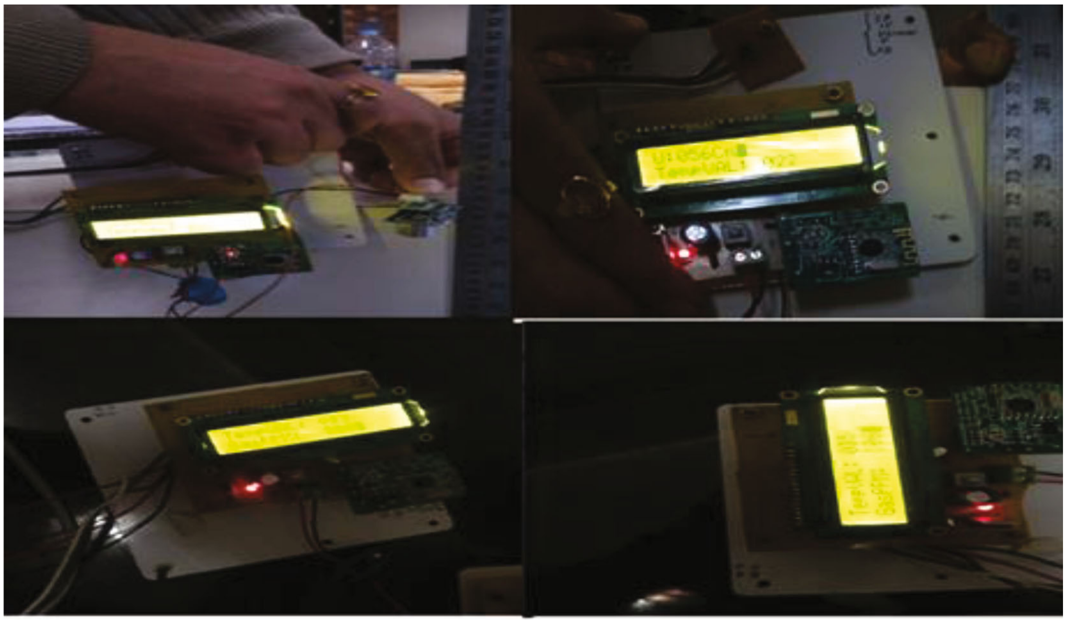
3.4.1. Security. As part of the security feature, during the implementation stage, we employ the AVR family of microcontrollers for applications comparable to the one being proposed. Complex algorithms are a curse in AVR utilizing time complex or memory because they frequently turn the executing code into a blocking one, which means that any other routine to be conducted by the controller is blocked and may create delays in executing the intended function. Keeping these factors in mind, we should take considerable sensitivity while employing cryptographic algorithms in these low-bit controllers. Cryptographic functions like XXTea and AVRCrypto can be used to implement the encryption method. In our case, we use the XXTea encryption functions to implement symmetric encryption with private keys that remain exclusive to another layer of network nodes to increase the security of data communication.

3.4.2. Interference. The basic parameter for detecting and handling interference includes bit error rate (BER), packet error rate (PER), received signal strength indicator (RSSI), signal-to-interference-to-noise ratio (SINR), throughput, and time delay. In this study, we apply an interference avoidance algorithm based on frequency agility [39]. This algorithm allows ZigBee to detect interference and flexibly move nodes to a secure channel with minimal power consumption and minimal latency while handling the interference. In this technique, the end devices measure PER with a transmission time of at least 20 packets, and if the PER goes beyond the 25% level, the interference information is transferred to the parent router to assess the link quality indicator (LQI). When the parent router determines that the LQI value is less than 100, it instructs the scanning energy detection to perform interference detection on accessible channels (ED).

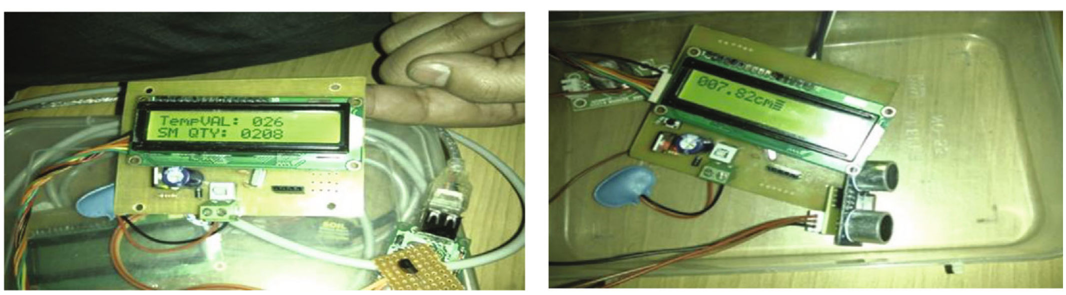
$$\mathrm{PER} = \left(\frac{\text{Number of failed messages}}{\text{Number of attempted measurements}}\right) * 100\%. \tag{2}$$

4. Machine Learning Models

As previously stated, the machine learning algorithm is used to recommend the crop based on sensor data. To achieve it, the algorithm must first be trained to understand the system and the environment to provide the outcome of crop



(a) Customized hardware



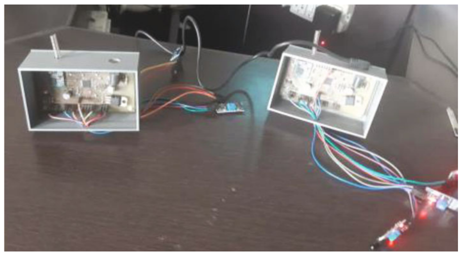
(b) Sensor node with sensors

(c) Blueterm data logger

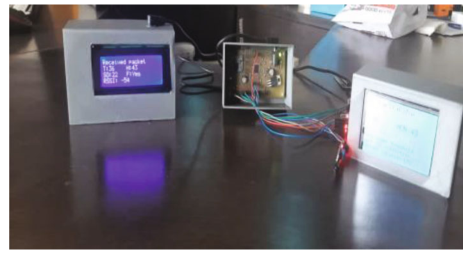
Figure 10: Hardware prototype.

recommendations. One of the goals of this study is that to choose which algorithm is the best algorithm that can be used not only in the system but in other situations with similar data. To do this, four different classification algorithms were tested to see which one had the highest accuracy.

(a) Decision trees (DTs): through hierarchical partitioning of training data, some functions are used to split the data, and this division is done iteratively until the leaf node contains the number record amount that can be used to classify data [44, 45]. However, as



(a) Sensor node



(b) Gateway

(c) Sensor node deployed in agricultural field

Figure 11: Final hardware and deployed sensor node in the agricultural field.

Table 7: Data splitting in training and testing.

Total samples	Ratio	Training sample	Testing sample
32,000	70 : 30	22,400	7,600

Table 8: Features of each data set.

Feature	Description
SensorID_i	ID of sensor
Value_i	Collected value
average_i	Average of the sensor values for last five observations
diff_sen_i	Difference from other sensor
$\text{hasRecommendation}_i$	Recommendation of crop

described in [46], this algorithm faces some limitations because a small change in the training data set can lead to a significant change in the tree and predicting the next value with accuracy becomes more difficult

(b) Support vector machine (SVM) is used mainly for classification, classifies data by building dimensions n between two classes, finds an optimal hyperplane to classify data, uses interval distance between neighboring points, and distinguishes between classes with minimum error margin [47]. According to a simpler interpretation, given the training data, the algorithm generates the best hyperplanes ranking new examples

(c) Random forest (RF) is best applied to classification problems and integrates the DT aggregate packing process by selecting a subset of features from the individual nodes in the tree, avoiding correlation on the bootstrap set [45] and working with a tree classifier, where one tree for each classifier

(d) XGBoost: with the same model as DT, the goal of this algorithm is, as the name suggests, to improve the performance of the model. It creates a sequence of models, and instead of training all the models individually, it models consecutively so that the new models try to correct the errors of the previous models [48]. The first model is built on the original

Table 9: Accuracy of the model.

Model	Accuracy (%) 1	2	3	4	5	6	7	8
DT	79.45	73.74	81.81	73.90	81.64	74.99	80.94	73.95
SVM	29.91	29.36	48.97	49.40	53.94	33.82	60.02	43.95
RF	76.20	69.85	80.35	73.53	80.35	73.42	80.17	74.49
XGBoost	80.45	75.20	75.74	78.82	85.06	80.36	86.71	83.64

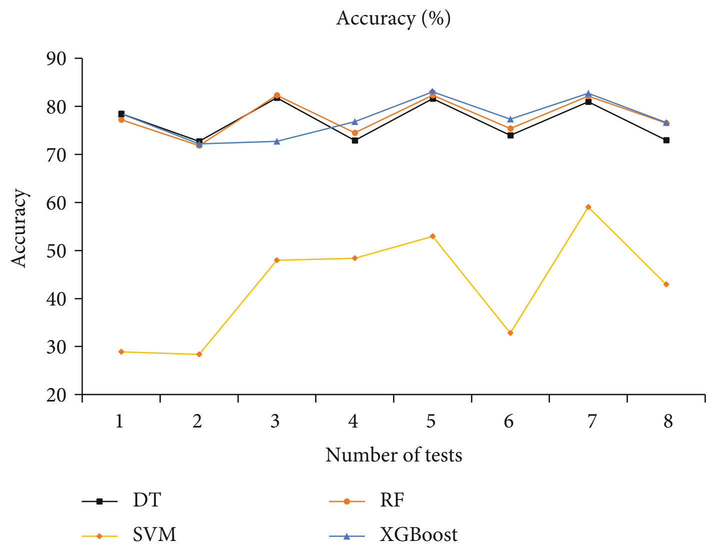


Figure 12: Accuracy of model.

data set, the second model improves the first model, the third model improves the second model, and so on. Models were added sequentially until no further improvements could be made

4.1. Data Creation. The complete hardware of the proposed system is shown in Figure 10(a). Figure 10(b) illustrates the sensor node based on the 2.4 GHz ZigBee module that is embedded with multiple sensors including water level, temperature, humidity, pH, and light intensity. Moreover, the sensor node is interfaced with a liquid crystal display for visualizing the sensor data. In this study, a custom sensor node with IoT sensors is deployed in the agricultural sector of the Maheru, Punjab region as shown in Figure 11. The sensor node is deployed in an outdoor environment to get data from the sensors in real time. As discussed earlier in this section, the sensor node is interfaced with multiple sensors to detect the environmental parameters of the farm field and communicate with the cloud server through a master node based on 2.4 GHz ZigBee communication and wireless fidelity (Wi-Fi).

The input parameters of the sensor data are temperature, humidity, water level, and soil pH. The 32,000 samples of IoT sensor nodes over 6 months are used to build a dataset for training. The manual data splitting methods have been adopted in which the total number of samples is divided into training and testing samples. The 70 : 30 ratio is chosen for the proposed dataset, in which 70% for training and 30% for testing as shown in Table 7.

During dataset generation, besides collected timestamps and sensor data, other features were added to each record based on calculations and preprocessed data from the values of the sensor. Table 8 shows the features of each data set. The difference between datasets is that the standard one only has one entry for each of the features while the clustered one has three entries for sensorID, average, value, and diff_sen, each one regarding the three sensors used in the test.

4.2. Model Analysis. A variety of ML algorithms were used to analyze the model to find the best one to use in our system. For our case, we performed a total of 8 tests for each algorithm, each with a different set of parameters, to train the algorithms to determine which one has the highest accuracy so that it can be applied in our system. Each test was run using Python, the scikit-learn library [37], and the Spyder platform. Scripts were developed for each algorithm using the appropriate library for scikit-learn classification and used the default configuration. As mentioned earlier, 70% of the dataset was used for training, and 30% was used for testing. Table 9 shows the results of each test, and Figure 12 shows the same results to improve the analysis.

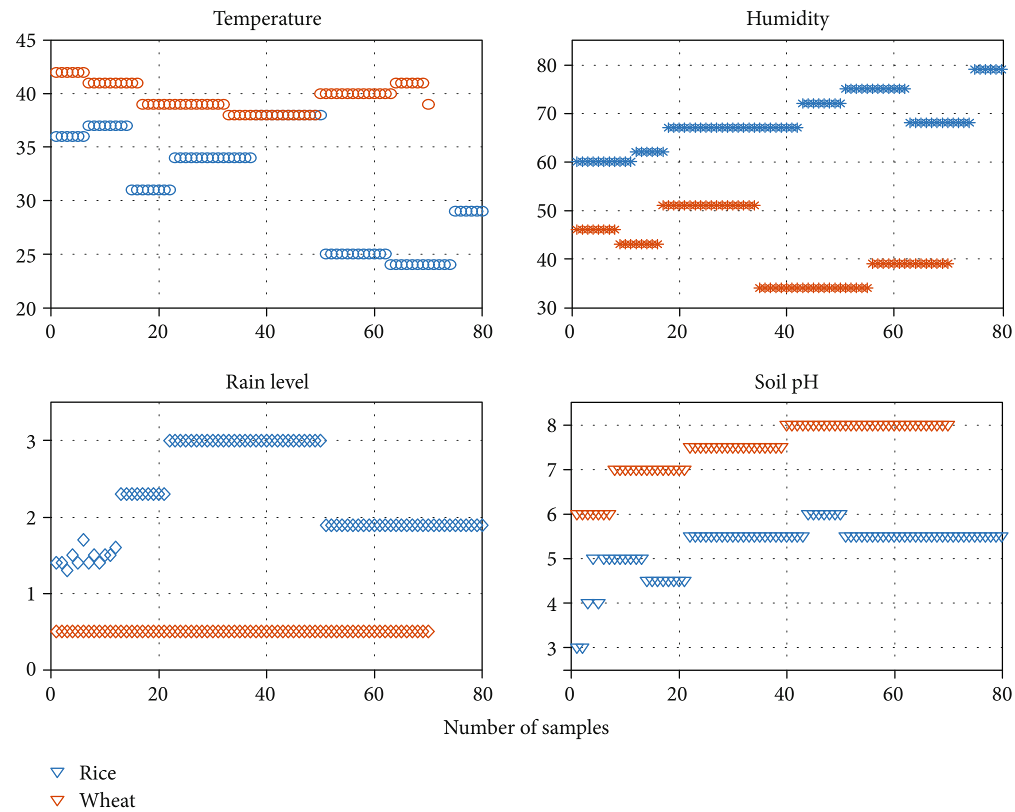


Figure 13: Crop recommendation based on temperature, humidity, rain level, and soil pH.

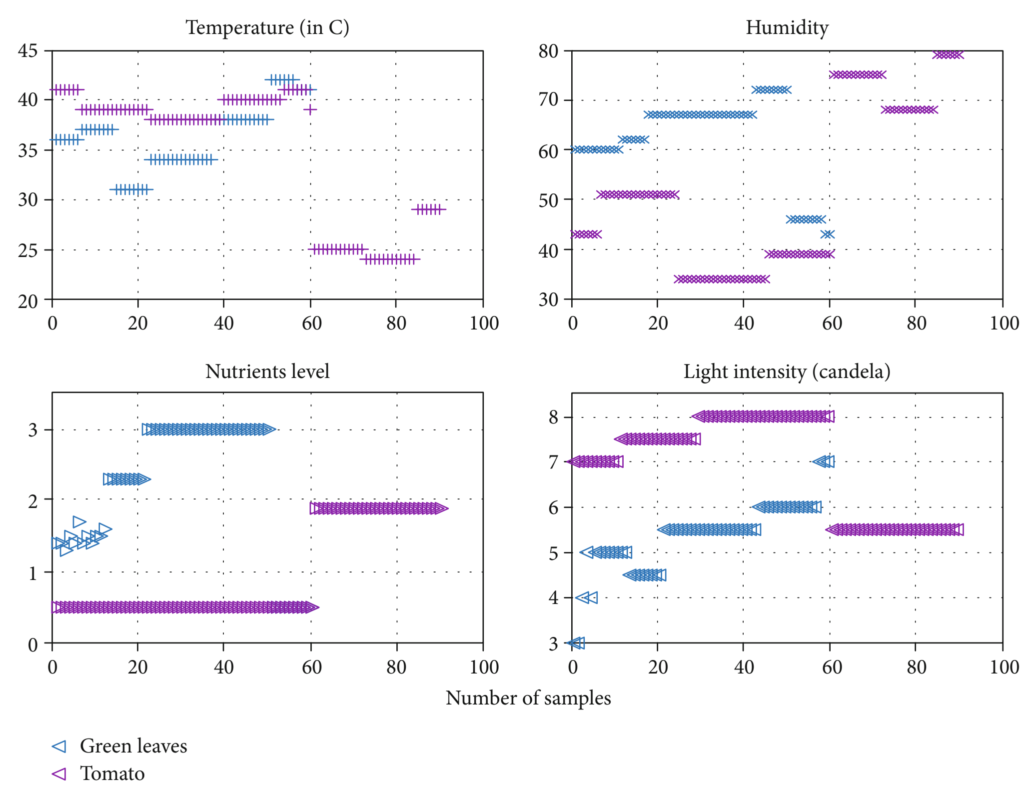


Figure 14: Crop recommendation based on temperature, humidity, nutrients, and light intensity.

Table 10: Comparisons with previous studies.

Ref	Communication	Data processing	Analytics	Interference	Calibration	Hand-held device	Security
[24]	Wi-Fi	Raspberry Pi is used for data processing	Decision tree algorithm	Interference test is not carried out	Sensors are not calibrated	NA, an email has been used for alerts	Security for data transmission is not integrated into the hardware
[49]	LoRa + Wi-Fi	Gateway	Not implemented	Interference between the sensor node and gateway is not mentioned	Sensors are not calibrated	Yes	Hardware is not empowered with security for data transmission
[50]	Wi-Fi	Gateway	Not implemented	Wi-Fi based on 2.4 GHz interference is not implemented	Sensors are not calibrated	NA	Asymmetric encryption for data communication is not available
[51]	NA	Edge computing	Deep reinforcement learning	Not mentioned	Sensors are not calibrated	NA	Security for data transmission is not mentioned
[52]	NA	Data aggregator	Decision support system (DSS)	Not mentioned	Sensors are not calibrated	No	Hardware is not empowered with security for data transmission
[53]	LoRa	LoRa gateway	No analytics are carried out	Not discussed	Utilized PIR, DTH11 and soil moisture sensor, but calibration is not carried out	No	No
[54]	LoRa	Gateway	System is implemented in real-time for testing, but no analytics are carried out	Interference are not appeared during transmission	Sensor calibration is carried out but no discussion about the calibration is available	No visualizing device or hand-held device	Security to the system is implemented with RFID, but during communication, no security feature is considered
[55]	LoRa+ NB-IoT	Aggregation node	SVM is implemented for detecting the leaks based on sensor data	NA	Calibration	An android application is developed to visualize the sensor values in real time	Security between the node to node is not explored
Proposed study	2.4 GHz + Wi-Fi	Master node	Real-time analytics on sensor data with a pretrained model	Interference avoidance algorithm based on frequency agility	Sensor calibration with a standard instrument to confirm the error-free sensor value	Hand-held device to receive the real-time sensor data and crop recommendation based on the user request	XXTea encryption functions are utilized to implement symmetric encryption

5. Real-Time Implementation

The hardware prototype is deployed in real-time to evaluate the developed system for obtaining crop recommendations based on sensor values. As shown in figure 4, the real-time sensor values are fed to the pretrained model. Based on received real-time sensor data, the XGBoost model has achieved better accuracy among the other pretrained ML models as shown in Table 9. Furthermore, the outcome of the model is illustrated in Figures 13(a) and 13(b). In Figure 13(a), the crop recommendation is processed based on temperature, humidity, rain level, and soil pH. The x

-axis denotes the number of the data samples, and y-axis denotes the range of temperature, humidity, rain level, and soil pH. In the temperature plot, orange denotes the wheat crop, and blue color denotes the rice crop.

Based on the temperature, humidity, rain level, and soil pH, the optimal crops are recommended. For example, in the temperature plot, for the temperature range of 39°C-45°C, the recommended crop is wheat. In case of humidity, the minimum humidity level is recommended for the wheat crop. Rice crop is recommended for the rain level of 3 mm and the pH value of 8, and the wheat crop is recommended. The recommendation of the crop is delivered on the basis of trained data, and it may vary with other location, as the dataset is developed through the real-time data of the study location. In case, if the hardware needed to implement in other location, then the customized hardware with IoT sensors needs to be developed for a certain period of time for the development of dataset.

In Figure 14, the crop recommendation is processed based on temperature, humidity, nutrients, and light intensity. The x-axis denotes the number of the data samples, and y-axis denotes the range of temperature, humidity, nutrients, and light intensity. In the temperature plot, blue denotes the green leaves, and violent color denotes tomatoes. Based on the temperature, humidity, nutrients, and light intensity, the optimal crops are recommended. In the temperature plot, for the temperature range of 20°C-30°C, the recommended crop is tomato. At the temperature of 35°C, the plot concludes that green leaves and tomato crops are recommended.

In case of humidity, if the range of humidity level is between 60% and 70%, then both crops including green leaves and tomato are recommended. The tomato crop is recommended for the nutrients level in between 0 and 1%. If the light intensity is at 8 candelas, then tomato crop is highly recommended. The hand-held device based on the cloud server also receives the crop recommendation through the Internet based on demand. Table 10 presents the comparison of smart agriculture monitoring with previous studies. To validate the proposed study, the following evaluation parameters are communication, data processing, analytics, simulation, calibration, and hand-held device. The proposed study offers the advantages of a communication protocol that uses both 2.4 GHz ZigBee and Wi-Fi.

The integration of these two-communication platforms enables to obtain the data locally and also in the cloud. From the table, it is concluded that the proposed study is having beneficial in terms of providing the real-time data to the users on the hand-held device through stable and reliable communication. As seen in the previous studies, many researchers have implemented the advanced wireless communication technology like LoRa. The deployment of the LoRa-based sensor nodes increases the infrastructure cost as compared to ZigBee-based sensor nodes. However, LoRa and ZigBee can be utilized combinedly in the agricultural field monitoring in the following manner.

ZigBee-based sensor nodes can be deployed in the agricultural field to monitor environmental parameters, and the single LoRa-based node can act as supervisor node to all the ZigBee-based sensor nodes. This LoRa-based node connects to the gateway, as LoRa can transmit the data to a long range. From there, the information of ZigBee-based sensor node can be visualized on the cloud server. This approach can be implemented to enhance the connectivity and minimize infrastructure cost. This study have enhanced the hardware with node mapping feature, frequency agility interference avoidance, and XXTea encryption features. All these features are logged in the hardware during the programming.

6. Conclusions

Real-time smart monitoring with intelligence has gained significant attention for increasing crop productivity. At the moment, IoT generates a large amount of real-time data from sensors, actuators, and identification technologies. However, extracting meaningful insights from data is required for the intelligent ecosystem and portable device for monitoring of agriculture. The current study is focused on implementing the customized hand-held device for assisting the farmers with crop recommendation with ML. To realize it, first, the customized hardware is designed, and sensors are calibrated to obtain the error free data. In addition to this, security is also inbuilt in the customized hardware for secure transmission of data on the cloud server. As the real-time data is available in the cloud server, it is utilized for forming dataset to conclude to the optimal ML model for crop recommendation. After identifying the optimal ML model, it has been applied on the cloud server. Based on this, the ML model recommends the crop from the real-time data that is generated from the customized hardware.

Data Availability

Data will be available on request. For the data-related queries, kindly contact to Baseem Khan, basseemk@hu.edu.et.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] World - Agriculture, “Value added (% of GDP) -1995-2018 data,” 2021 Forecast Available online: https://tradingeconomics.com/world/agriculture-value-added-percent-of-gdp-wb-data.html/ (accessed on Apr 13, 2021).

[2] K. Pawlak and M. Kołodziejczak, “The role of agriculture in ensuring food security in developing countries: considerations in the context of the problem of sustainable food production,” *Sustainability*, vol. 12, no. 13, p. 5488, 2020.

[3] “Food production must double by 2050 to meet demand from world’s growing population innovative strategies needed to combat hunger, experts tell second committee,” Meetings Coverage and Press Releases Available online: https://press.un.org/en/2009/gaef3242.doc.htm (accessed on Apr 13, 2021).

[4] M. Ayaz, M. Ammad-Uddin, Z. Sharif, A. Mansour, and E.-H. M. Aggoune, “Internet-of-Things (IoT)-based smart

agriculture: toward making the fields talk," *IEEE Access*, vol. 7, pp. 129551–129583, 2019.

[5] M. Carlos-Mancilla, E. López-Mellado, and M. Siller, "Wireless sensor networks formation: approaches and techniques," *Journal of Sensors*, vol. 2016, Article ID 2081902, 18 pages, 2016.

[6] F. Edwards-Murphy, M. Magno, P. M. Whelan, J. O'Halloran, and E. M. Popovici, "b+WSN: smart beehive with preliminary decision tree analysis for agriculture and honey bee health monitoring," *Computers and Electronics in Agriculture*, vol. 124, pp. 211–219, 2016.

[7] U. Shafi, R. Mumtaz, J. García-Nieto, S. A. Hassan, S. A. R. Zaidi, and N. Iqbal, "Precision agriculture techniques and practices: from considerations to applications," *Sensors (Switzerland)*, vol. 19, no. 17, pp. 1–25, 2019.

[8] A. Pardossi and L. Incrocci, "Traditional and new approaches to irrigation scheduling in vegetable crops," *HortTechnology*, vol. 21, no. 3, pp. 309–313, 2011.

[9] V. P. Kour and S. Arora, "Recent developments of the Internet of Things in agriculture: a survey," *IEEE Access*, vol. 8, pp. 129924–129957, 2020.

[10] P. Majumdar and S. Mitra, "IoT and machine learning-based approaches for real time environment parameters monitoring in agriculture: an empirical review," in *Agricultural Informatics: Automation Using the IoT and Machine Learning*, pp. 89–115, Wiley, 2021.

[11] M. S. Mahdavinejad, M. Rezvan, M. Barekatain, P. Adibi, P. Barnaghi, and A. P. Sheth, "Machine learning for internet of things data analysis: a survey," *Digital Communications and Networks*, vol. 4, no. 3, pp. 161–175, 2018.

[12] B. S. Paul and S. Rimer, "Wireless sensor node placement due to power loss effects from surrounding vegetation," in *Proceedings of the International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness*, pp. 915–927, Greader Noida, India, 2013.

[13] D. V. Jose and G. Sadashivappa, "Mobile sink assisted energy efficient routing algorithm for wireless sensor networks," *World of Computer Science & Information Technology Journal*, vol. 5, 2015.

[14] P. Sanjeevi, S. Prasanna, B. Siva Kumar, G. Gunasekaran, I. Alagiri, and R. Vijay Anand, "Precision agriculture and farming using Internet of Things based on wireless sensor network," *Transactions on Emerging Telecommunications Technologies*, vol. 31, article e3978, 2020.

[15] L. I. U. Dan, C. Xin, H. Chongwei, and J. Liangliang, "Intelligent agriculture greenhouse environment monitoring system based on IOT technology," in *Proceedings of the 2015 International Conference on Intelligent Transportation, Big Data and Smart City*, pp. 487–490, Halong Bay, Vietnam, 2015.

[16] R. Singh, A. Gehlot, A. K. Thakur, M. Swain, and S. V. Akram, "Wireless sensor network with power management system for water level regulation in paddy fields," *International Journal of Innovative Technology and Exploring Engineering*, vol. 9, pp. 1243–1246, 2020.

[17] J. Gutiérrez, J. F. Villa-Medina, A. Nieto-Garibay, and M. Á. Porta-Gándara, "Automated irrigation system using a wireless sensor network and GPRS module," *IEEE Transactions on Instrumentation and Measurement*, vol. 63, pp. 166–176, 2014.

[18] X. Zhang, J. Zhang, L. Li, Y. Zhang, and G. Yang, "Monitoring citrus soil moisture and nutrients using an IoT based system," *Sensors*, vol. 17, no. 3, p. 447, 2017.

[19] H. Zhang and F. Zhang, "A kind of design of greenhouse environment monitoring system based on LabVIEW," in *Proceedings of the 2020 Chinese Automation Congress (CAC)*, pp. 6208–6211, Shanghai, China, 2020.

[20] K. Foughali, K. Fathallah, and A. Frihida, "Using cloud IOT for disease prevention in precision agriculture," *Procedia computer science*, vol. 130, pp. 575–582, 2018.

[21] M. Srbinovska, C. Gavrovski, V. Dimcev, A. Krkoleva, and V. Borozan, "Environmental parameters monitoring in precision agriculture using wireless sensor networks," *Journal of Cleaner Production*, vol. 88, pp. 297–307, 2015.

[22] X. Zhang, Q. Wen, D. Tian, and J. Hu, "PVIDSS: developing a WSN-based irrigation decision support system (IDSS) for viticulture in protected area, Northern China," *Applied Mathematics & Information Sciences*, vol. 9, no. 2, pp. 669–679, 2015.

[23] O. Ayurzana and S. Tsagaanchuluun, "Monitoring system of agriculture fields using ZigBee modules," *International journal of advanced smart convergence*, vol. 10, pp. 89–96, 2021.

[24] K. S. P. Reddy, Y. M. Roopa, and N. S. Nandan, "IoT based smart agriculture using machine learning," in *Proceedings of the 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, pp. 130–134, Coimbatore, India, 2020.

[25] G. Sushanth and S. Sujatha, "IOT based smart agriculture system," in *Proceedings of the 2018 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, pp. 1–4, Chennai, India, 2018.

[26] N. Sales, O. Remédios, and A. Arsenio, "Wireless sensor and actuator system for smart irrigation on the cloud," in *Proceedings of the 2015 IEEE 2nd World Forum on internet of things (WF-IoT)*, pp. 693–698, Milan, Italy, 2015.

[27] S. Rawal, "IOT based smart irrigation system," *International Journal of Computers and Applications*, vol. 159, no. 8, pp. 7–11, 2017.

[28] O. Elijah, T. A. Rahman, I. Orikumhi, C. Y. Leow, and M. N. Hindia, "An overview of Internet of Things (IoT) and data analytics in agriculture: benefits and challenges," *IEEE Internet of Things Journal*, vol. 5, no. 5, pp. 3758–3773, 2018.

[29] A. Khattab, A. Abdelgawad, and K. Yelmarthi, "Design and implementation of a cloud-based IoT scheme for precision agriculture," in *Proceedings of the 2016 28th International Conference on Microelectronics (ICM)*, pp. 201–204, Giza, Egypt, 2016.

[30] A. Kamilaris, F. Gao, F. X. Prenafeta-Boldu, and M. I. Ali, "Agri-IoT: a semantic framework for Internet of Things-enabled smart farming applications," in *Proceedings of the 2016 IEEE 3rd World Forum on Internet of Things (WF-IoT)*, pp. 442–447, Reston, VA, USA, 2016.

[31] F. Kiani and A. Seyyedabbasi, "Wireless sensor network and Internet of Things in precision agriculture," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 6, 2018.

[32] H. Lee, A. Moon, K. Moon, and Y. Lee, "Disease and pest prediction IoT system in orchard: a preliminary study," in *Proceedings of the 2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN)*, pp. 525–527, Milan, Italy, 2017.

[33] J. Treboux and D. Genoud, "Improved machine learning methodology for high precision agriculture," in *Proceedings of the 2018 Global Internet of Things Summit (GIoTS)*, pp. 1–6, Bilbao, Spain, 2018.

[34] A. Chlingaryan, S. Sukkarieh, and B. Whelan, “Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: a review,” *Computers and Electronics in Agriculture*, vol. 151, pp. 61–69, 2018.

[35] K. G. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, “Machine learning in agriculture: a review,” *Sensors*, vol. 18, no. 8, p. 2674, 2018.

[36] R. Sharma, S. S. Kamble, A. Gunasekaran, V. Kumar, and A. Kumar, “A systematic literature review on machine learning applications for sustainable agriculture supply chain performance,” *Computers and Operations Research*, vol. 119, article 104926, 2020.

[37] “ZigBee modules - everything RF,” Available online: https://www.everythingrf.com/search/rf-modules/filters?page=1&country=global&stechnology=;ZigBee/ (accessed on Oct 27, 2021).

[38] L. A. A. V. R. Microcontroller, I. Programming, O. B. Program, O. A. Comparator, F. Sleep, and M. Idle, “Reduction, A.D.C.N. – up to 16mips throughput at 16mhz – optional boot code section with independent lock bits – programming lock for software security – three PWM channels – 6-channel ADC in PDIP package – master / slave SPI serial interface – internal calibrated rc oscilla,” 2013.

[39] S. T. Electronics, “Soil moisture sensor - digital + analog output,” pp. 1–7, 2016, https://www.sunrom.com/.

[40] “Moisture content determination | Mettler Toledo | learn from experts now!,” Available online: https://www.mt.com/in/en/home/applications/Laboratory_weighing/moisture-content-determination.html/(accessed on Oct 27, 2021).

[41] S. Voltage, S. Current, and O. D. Format, “Ultrasonic distance sensor - serial out [1166]: Sunrom electronics/technologies”.

[42] “Psychrometer calibration |Sling Psychrometer Calibration,” Available online: https://www.alliancecalibration.com/psychrometer-calibration(accessed on Oct 27, 2021).

[43] “Relative humidity and dewpoint temperature from temperature and wet-bulb temperature formula - weather calculators and converters -4WX.COM? Weather Information Center!,” Available online: http://www.4wx.com/wxcalc/formulas/rhTdFromWetBulb.php (accessed on Oct 27, 2021).

[44] R. Saravanan and P. Sujatha, “A state of art techniques on machine learning algorithms: a perspective of supervised learning approaches in data classification,” in *Proceedings of the 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, pp. 945–949, Madurai, India, 2018.

[45] M. Mamdouh, M. A. I. Elrukhsi, and A. Khattab, “Securing the internet of things and wireless sensor networks via machine learning: a survey,” in *Proceedings of the 2018 International Conference on Computer and Applications (ICCA)*, pp. 215–218, Beirut, Lebanon, 2018.

[46] L. Li, W. Chou, W. Zhou, and M. Luo, “Design patterns and extensibility of REST API for networking applications,” *IEEE Transactions on Network and Service Management*, vol. 13, no. 1, pp. 154–167, 2016.

[47] Q. Tang, X. Ge, and Y.-C. Liu, “Performance analysis of two different SVM-based field-oriented control schemes for eight-switch three-phase inverter-fed induction motor drives,” in *Proceedings of the 2016 IEEE 8th International Power Electronics and Motion Control Conference (IPEMC-ECCE Asia)*, pp. 3374–3378, Hefei, China, 2016.

[48] “A beginner’s guide to XGBoost. This article will have trees…. lots of… | by George Seif | towards data science,” Available online: https://towardsdatascience.com/a-beginners-guide-to-xgboost-87f5d4c30ed7(accessed on Oct 26, 2021).

[49] M. Swain, D. Zimon, R. Singh, M. F. Hashmi, M. Rashid, and S. Hakak, “LoRa-LBO: an experimental analysis of LoRa link budget optimization in custom build IoT test bed for agriculture 4.0,” *Agronomy*, vol. 11, no. 5, 2021.

[50] T. Cao-Hoang and C. N. Duy, “Environment monitoring system for agricultural application based on wireless sensor network,” in *Proceedings of the 2017 Seventh International Conference on Information Science and Technology (ICIST)*, pp. 99–102, Da Nang, Vietnam, 2017.

[51] F. Bu and X. Wang, “A smart agriculture IoT system based on deep reinforcement learning,” *Future Generation Computer Systems*, vol. 99, pp. 500–507, 2019.

[52] P. K. Tripathy, A. K. Tripathy, A. Agarwal, and S. P. Mohanty, “MyGreen: An IoT-enabled smart greenhouse for sustainable agriculture,” *IEEE Consumer Electronics Magazine*, vol. 10, no. 4, pp. 57–62, 2021.

[53] H. Muthukrishnan, A. Jeevanantham, B. Sunita, S. Najeerabanu, and V. Yasuvanth, “Performance analysis of Wi-Fi and LoRa technology and its implementation in farm monitoring system,” *IOP Conference Series: Materials Science and Engineering*, vol. 1055, no. 1, article 012051, 2021.

[54] T. A. Khoa, M. M. Man, and T. Nguyen, “Smart agriculture using IoT multi-sensors: a novel watering management system,” *Journal of Sensor and Actuator Networks*, vol. 8, no. 3, 2019.

[55] J. Alves Coelho, A. Glória, and P. Sebastião, “Precise water leak detection using machine learning and real-time sensor data,” *IoT*, vol. 1, no. 2, pp. 474–493, 2020.