Smart Irrigation System using Machine Learning and IoT

Bharath D.A.[1], S. Amith Nadig[2], Manjunath G.S.[3]

[1]*VIII Semester, Dept. of ISE, BNMIT*
[2] *VIII Semester, Dept. of ISE, BNMIT*
[3]*Asst. Professor, Dept. of ISE, BNMIT, Karnataka, India*

---***---

Abstract - *To realise IoT promise in commercial-scale applications, integrated Internet of Things (IoT) platforms are required. The key challenge is to make the solution flexible enough to fulfil the demands of specific applications. A IoT-based platform for smart irrigation with a flexible design is created so that it allows developers to quickly link IoT and machine learning (ML) components to create application solutions. The design allows for a variety of customised analytical methods to precision irrigation, allowing for the advancement of machine learning techniques. Impacts on many stakeholders may be predicted, including IoT specialists, who would benefit from easier system setup, and farmers, who will benefit from lower costs and safer crop yields.*

The typical irrigation procedure necessitates a large quantity of water use, which results in water waste. An intelligent irrigation system is desperately needed to decrease water waste during this tiresome process. Machine learning (ML) and the Internet of Things (IoT) have made it possible to develop an intelligent system that can accomplish this operation automatically and with minimum human intervention. An IoT-enabled ML-trained recommendation system is suggested in this paper for optimum water consumption with minimal farmer interaction. In the agriculture field, IoT sensors are used to capture exact ground and environmental data. The collected data is transferred and kept in a cloud-based server that uses machine learning to evaluate the data and provide irrigation recommendations.

Key Words: IoT, ML, cloud, irrigation, water.

1. INTRODUCTION

In India, where agriculture accounts for 60-70 percent of the GDP, there is a pressing need to modernise traditional agricultural techniques in order to increase output. The groundwater table is lowering day by day as a result of uncontrolled water usage; lack of rainfall and shortage of land water also contribute to a decrease in the amount of water on the planet. Water scarcity is currently one of the world's most pressing issues. Water is required in every sector. Water is also necessary in our daily lives.

Agriculture is one of the industries that need a lot of water. Water wastage is a serious issue in agriculture. Every time there is a surplus of water, it is distributed to the fields. Climate change and its consequences are widely explored in academic studies on water resources and agriculture. Because of the potential repercussions of global warming, water adaptation methods are being considered to assure water availability for food and human production as well as ecosystem sustainability. Additionally, the safety of water for human consumption and return to the environment must be maintained. Increased water shortages, poor quality of water, higher water and soil salinity, loss of biodiversity, increased irrigation needs, and the expense of emergency and corrective action are all possible risks from climate change. As a result of these factors, a rising number of research are focusing on creating creative water utilisation in irrigation. The Internet of Things (IoT) has now progressed from a concept to being implemented in real-world applications. Since then, the technological and application hurdles have been considerable.

IoT platforms enable complex real-time control systems by layering communication infrastructure, hardware, software, analytical approaches, and application knowledge. Recognizing the expected IoT consequences on systems is one of the most difficult technological issues, because IoT allows systems to become service mashups, combining items as services. System development will become dynamic plug-and-play interoperable service composition, and system logic will become service orchestration as a result.

An IoT-based smart irrigation system with an effective machine learning algorithm is developed to assist farmers in overcoming the uncertainty of rainfall and increasing production. This model provides a superior irrigation decision-making model. This research presents a Machine Learning (ML) strategy for successfully regulating irrigation and enhancing agricultural yield as a result.

2. LITERATURE REVIEW

Goldstein et al. (2017) [1] suggested a recommendation-based irrigation management system that combined machine learning with agronomic knowledge. According to the system, the best regression model with 93 percent accuracy, and the best classifier model with 95 percent accuracy, Gradient Boosted Regression Trees and Boosted Tree Classifier, provide superior irrigation prediction decisions than the linear regression model. To assist the agronomist in making better selections, the models were trained with eight separate sets of features. The Internet of

Multimedia Things (IoMT) was proposed by Al Zu'bi et al.[2] (2019), which focuses on the use of multimedia sensors in the field for irrigation optimization. Digital image processing is used to monitor crops and soil.

The multimedia sensor sends the collected photos of the crops to the image processing system, which makes the choice based on the proportion of soil cracks. This makes the Future No-Man irrigation management system possible.

According to the data mining technique, Rushika Ghadge et al. (2018)[3] created a system that employs both supervised and unsupervised ML algorithms to forecast crop and soil quality and kind of land, as well as assess the nutrients present in the soil to boost agricultural yield. This effort assists farmers in cultivating healthier crops in the proper soil to increase yield, as well as serving as a conduit for providing timely information to farmers regarding crop quality and nutrient requirements. For soil moisture estimation, a learning model based on Support Vector Regression (SVR) and K-means clustering was created. Humidity, radiation, soil moisture, air, and soil temperature were all captured in the field and sent into the training system of SVR model.

To increase accuracy and reduce error rate, the SVR model output is sent to K-means clustering. For optimal yield management, the final output from k-means is employed to govern the water pump controller. However, the majority of the older prediction models had a large variance, which causes the machine learning model to perform poorly. Ensemble learning, as described by Zhao et al. (2018)[4], Catolino and Ferrucci[5] (2018), Joshi and Srivastava[6] (2014), and Ren et al.[7] (2016), may be utilised to deal with such significant variation. By combining different learning models to predict the output of a single system, ensemble approaches improve performance.

Some ensemble approaches, particularly bagging, have been shown to decrease the problem of overfitting and underfitting of training data. The Bootstrap Aggregation approach improves single regression trees by using many models, each of which is trained using randomly selected samples from the original dataset. The bagging approach has a smaller prediction error than the other single models. Gonzalez et al. (2014)[8] provide a bagging method for forecasting power price that is compared to the random forest approach for both classification and regression models.

A complete literature study was conducted, and the paper suggests some of the most efficient feasible technologies and algorithms for the creation of a Smart Farm Monitoring System based on the findings of the literature research and experiments. Ersin et al.[9] suggested a microcontroller-based irrigation system that is more efficient and cost-effective than other traditional techniques. Liu et al. described precision irrigation technologies. [10]. Agrawal et al. presented a smart irrigation system using Raspberry Pi and Arduino. [11]. Koprda et al. presented a microcontroller-based irrigation solution. Ahouandjinou et al. discuss farm pest detection using ultrasonic sensors in their paper [12]. [13]. Goap et al.[14] provided a full overall design for an IoT-based irrigation system.

Smith et al.[15] presented machine learning algorithms for soil categorization. Wu et al.[16] investigated a farm vehicle and smart dispatching strategy. Ryu et al. [17] presented an integrated method to smart farming. Kwok et al. [18] proposed utilising deep learning to recognise plants and then determining the optimal watering volume depending on plant type. Wang, Muzzammel, Raheel, and colleagues[19] explored deep learning and an altitude-based economical irrigation technique. A WSN technique for precision farming was presented by Martinell et al. [20]. Izquierdo et al. [21] suggested a smart farming solution based on cloud and edge computing.

Bacco [22] provided a detailed smart farming method in their study effort, including all constraints, enablers, and prospects. This paper emphasises and presents a comprehensive, precise picture of the feasible answer for agricultural demands following a thorough analysis of currently accessible literature that deals with contemporary farming challenges and their associated solutions. The paper describes a distributed sensor network field whose prototype was created for this research.

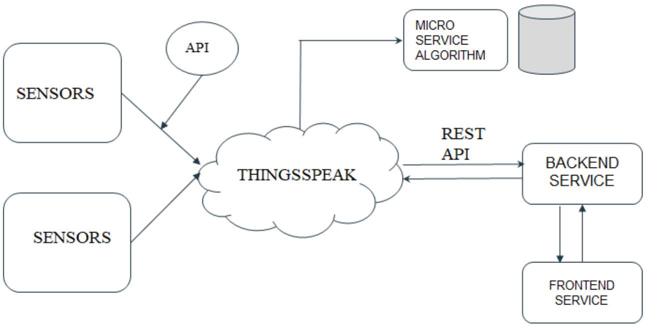
3. METHODOLOGY

3.1. Existing Methods

In general, there is no automated irrigation method that is being used all over the world. However, some study has been done in the topic of automating the watering process. In most existing studies, the following is the basic method for Automated Irrigation: To begin, data is collected from several sensors to determine the moisture content of the soil and the temperature of the surrounding environment. They're attached to a breadboard that's wired up to the Arduino board. The Arduino IDE receives the data from the board. The programming language employed executes instructions that extract data and reflect it, i.e., a decision is made whether to turn the water pump "On" or "Not" based on the extracted data.

3.2. Proposed Solution

Step 1: As demonstrated in Fig. 1, irrigation can be automated utilising sensors, microcontrollers, Wifi modules, and the ThingSpeak platform. A controller is necessary to maintain all of the sensors and to drive the motor as needed. We utilised NodeMcu to accomplish this. The NodeMcu can output a maximum voltage of 5 volts. The moisture sensors module and DTH11 sensor can both be powered by 5 volts, but not the motor. We need at least 7 volts to run a motor. To solve this issue, we utilised a 9v battery to power the motor. We'll need a switch to regulate the motor whenever it's needed. We utilised a relay module to accomplish this. It's a switch in the electrical system. We must deliver a strong pulse to the module in order to close the switch. The field is constantly monitored by the soil moisture sensor. Node MCU is attached to the sensors. The sensor data is sent to the user via wireless transmission so that he can manage irrigation.



Step2: Finally, all the data needs to be there in ThingSpeak for visualisation. by using write API keys we will send the sensor data to the server. In ThingSpeak we can visualise the data of every sensor over the time.

Step3: Fetching the data from ThingSpeak to our python script

1. Import all required libraries (json, urllib.request).
2. Create an API using READ_API_KEY and CHANNEL_ID of ThingSpeak.
3. Request the ThingSpeak website by using urllib.request module.
4. Store the json response from ThingSpeak.
5. Retrieve temperature, Humidity, Soil moisture values from json data.

Step4:

1. The weather data was obtained through the Kaggle platform.

2. The performance of rainfall prediction is benchmarked using a variety of learning methods in this work. These are the supervised learning methods NB, C4.5, SVM, ANN, and RF.

3. An ensemble of the above models is used to train a Voting Classifier, which predicts an output (class) based on the output's highest likelihood. It simply sums up the results of each classifier fed into the Voting Classifier and predicts the output class with the most votes. We propose a single model that trains on numerous models and predicts output based on the cumulative majority of votes for each output class, rather than building separate specialised models and determining their performance.

Step5: Taking the final decision

When the soil moisture falls below a certain threshold, a motor will switch on. Instead of turning on the motor immediately, we examine the probability of rain using the above-mentioned ensemble methodologies, and if rain is likely to occur during that time period, we wait a while. From crop to crop, the threshold level will differ. If a crop need more water, we will increase the high threshold level so that the crop receives more water. Alternatively, if the crop requires less water, we will specify a low value

4. IMPLEMENTATION

Hardware used:

1. NodeMcu ESP8266
2. Soil Moisture Sensor Module
3. Submersible DC motor
4. DTH11 sensor
5. Relay module

Software used:

1. Arduino IDE
2. ThingSpeak
3. Google Sites

NodeMcuEsp8266: It's an IoT gadget that's open-source. It's a 32-bit microcontroller that allows Wi-Fi-connected gadgets to send and receive data. It's a low-cost semiconductor with TCP/IP networking software built in. There are 17 GPIO pins on this board. It contains a Tensilica L 106 RISC CPU that uses very little electricity. It's compatible with ADCs, power amplifiers, and certain power management modules are all available. It contains 4KB of memory storage. Figure 1 depicts NodeMcu in its most basic form.

Fig - 1: NodeMcu

Soil Moisture Sensor Module: The purpose of a soil moisture sensor is to determine the amount of water in the soil. It comprises mostly of a pair of conducting probes. The change in resistance between those probes is used to calculate the moisture content. The quantity of moisture in the soil has an inverse relationship with resistance. It transmits analogue data.

The value will vary from 0 to 1023 after feeding this into ADC. As a result, if there is no water in the soil, the value decreases. 1023 will be the number. So for changing this value into percent we need map (0,1023) to (1,100) which can be done using map function.

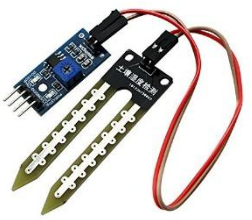


Fig - 2: Soil Moisture Sensor.

Submersible DC: motor is one that can be completely immersed in water. In order to prevent water from entering the motor, it is hermetically sealed. It converts rotational energy into kinetic energy, which is then converted into pressure energy, which pushes water to the surface. This engine will be submerged in water, with a conduit connecting it to the water's output.

Fig - 3: Submersible pump

DTH11 sensor: It's a multi-purpose sensor that measures temperature and humidity in the environment. It is made up of humidity detecting material and a temperature sensing thermistor. A humidity detecting material is a capacitor with humidity as a dielectric substance between them, causing the capacitance to alter as the humidity changes. We understand how thermistors function. The resistance value fluctuates as a function of temperature. It operates at 3-5 volts, which we can acquire from the NodeMcu.

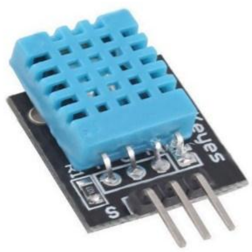


Fig - 4: DTH11 Sensor

Relay Module: It's a type of electrical switch that works by using magnetism. The Relay module's primary function is to control the motor. The NodeMcu's maximum output voltage is 5 volts, which is insufficient to drive the motor. So, to drive the motor, we'll connect the relay module to the NodeMcu, and power the module with a 9v battery. We will send a high to low pulse to the relay module anytime we want to turn on the motor, the switch will shut, and 9v will be sent to the motor. The single channel Relay module is shown in Fig 5 as an example

Fig - 5: Relay Module

5. ALGORITHMS

The categorization technique C4.5 is one of the most effective. C4.5 generates a decision tree, where each node divides the classes according to the information. The property with the highest normalised information gain is used to determine the splitting criteria. Humidity and temperature, for example, are included in our data collection. The C4.5 algorithm investigates these aspects first to determine which is optimal for data splitting (a feature with maximum information gain). After that, the feature is utilised to partition the dataset into the following feature till it reaches the final destination. The algorithm's output is shown in Table.

Naïve Bayes

Nave Bayes is a supervised machine learning model that belongs to the family of probabilistic classifiers that applies the Nave Bayes theory to the dataset's independence assumption. By computing the assumptions, Nave Bayes determines the probability of each feature in the dataset. Nave Bayes determines every attribute conditional probability on the class label for every known class label. The product rule is then used to calculate the joint conditional probability for the characteristics of a label. The Nave Bayes model is then used to derive the conditional probability for the class characteristics. The class with the highest probability is provided after performing this method for each class value. The results of the algorithm are shown in Table.

Support Vector Machine

The support vector machine is a machine learning model that uses a hyperplane to partition a dataset into two pieces. This partitioning procedure treats each class label separately, and it may be done by classifying the data into class A and not class B, where A and B are the two class labels. Calculating the Euclidean distance between each data point and the hyperplane's margin is used to classify the data. When data cannot be linearly separated in a lower level space, the Support Vector Machine model employs a kernel, which is a set of scientific functions, to allow for data categorization in a complicated dimensional space. In machine learning, many kernel functions, such as radial, are available to regulate the above.

Neural Networks

Exhibit machines that mimicked the brain's functions influenced the development of neural networks. Every brain unit is linked to a slew of others. In terms of the initial state effect of the linked neuronal units, links might be either enforcing or inhibitory. A summing function might be used to unite the input values of each individual brain unit. This model is utilised in regression and classification, as well as prediction and clustering. There are two primary factors that have a substantial impact on the performance of neural network classifiers. The number of hidden layers is the first, and the value of the learning rate is the second. The results of the algorithm are shown in Table IV.

Random Forest

Random forest is a machine learning model that may be used for prediction, regression, and classification, among other things. This algorithm is an ensemble of decision tree models that aims to produce a multiplicity of decision tree models from the same training data and generate the final class as the output. The number of characteristics to freely study (Num Features), maximum depth of the tree (Max Depth), and number of trees (Num Tree) parameters are modified in the random forest classifier. The findings of the research show shows the Random Forest classifier's classification performance improves as the number of features, trees, and depth grow. The results of the algorithm are shown in Table.

6. RESULTS AND DISCUSSIONS

A. C4.5 Algorithm

Accuracy: 0.77

	Precision	Recall	F1-Score	Support
0	0.80	0.88	0.84	274
1	0.64	0.50	0.56	118
Average	0.72	0.69	0.70	392

Table I: Results for the C4.5 Algorithm

B. Naïve Bayes

Accuracy: 0.81

	Precision	Recall	F1-Score	Support
0	0.84	0.91	0.88	280
1	0.72	0.56	0.63	112
Average	0.78	0.74	0.75	392

Table II: Results for the Support Vector Machine Algorithm

C. Support Vector Machine

Accuracy: 0.82

	Precision	Recall	F1-Score	Support
0	0.84	0.95	0.89	291
1	0.76	0.47	0.58	101
Average	0.80	0.71	0.73	392

Table III: Results for the Support Vector Machine Algorithm

D. Neural Networks

Accuracy: 0.77

	Precision	Recall	F1-Score	Support
0	0.81	0.88	0.83	274
1	0.76	0.64	0.72	118
Average	0.78	0.76	0.77	392

Table IV: Results for the ANN Algorithm

E. Random Forest

Accuracy: 0.75

	Precision	Recall	F1-Score	Support
0	0.81	0.85	0.83	278
1	0.58	0.50	0.54	114
Average	0.69	0.67	0.68	392

Table V: Results for the Random Forest Algorithm

Machine learning-based prediction performance varies amongst algorithms, with the artificial neural network approach having a modest performance edge over other categorization models. Each model has certain flaws, but the overall outcome is always better since the error rate is reduced. When one model fails, other models will step in to help. Because we're employing ensemble learning, we'll take into account a mix of model knowledge. a majority of people A voting model is constructed that trains on a mixture of the aforementioned models and predicts an output (class) based on the majority of the highest likelihood of each model's preferred class as the output. It simply aggregates the results of each machine learning model fed into Voting Classifier and predicts the output class based on the most votes.

The representation of detected data via sensors in ThingSpeak is shown in Figures 11,12,13. It aids in the interpretation of data. We may use this data to combine, transform, and create new data, and we can use built-in charting algorithms to graphically grasp the relationships between the data. In the future, we'll be able to combine data from numerous sources to provide a more complex study.

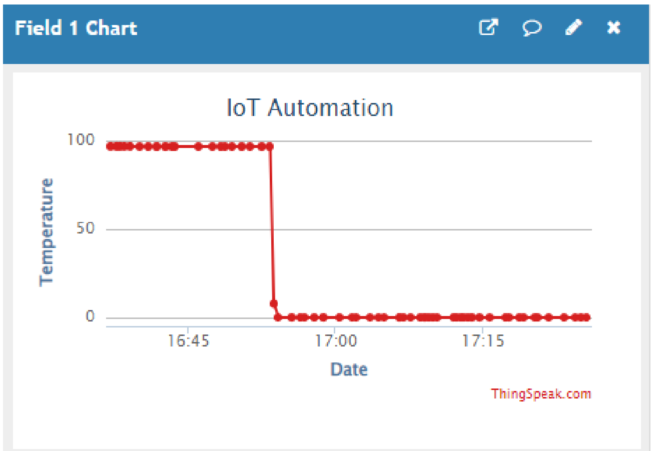


Fig - 6: Field Chart 1

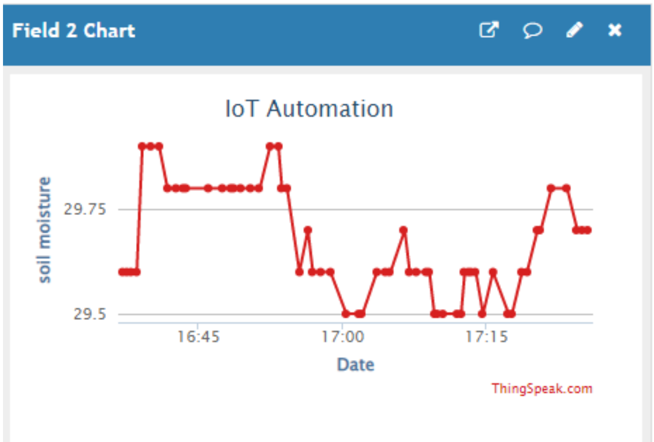


Fig - 7: Field Chart 2

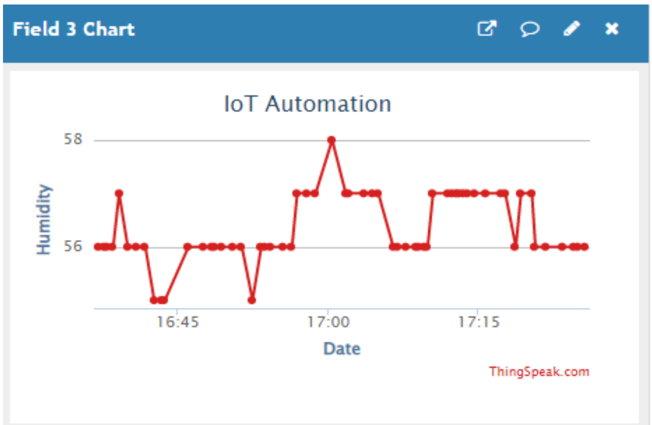


Fig - 8: Field Chart 3

7. CONCLUSION AND FUTURE ENHANCEMENT

Regular crop updates, such as moisture, humidity, and temperature, are critical in agriculture. Climate forecasting data accuracy has increased dramatically as a result of technological advancements, and weather forecasting data may now be utilised to estimate rainfall in a specific location. To estimate rainfall possibilities, this study suggests an Automated Irrigation System that uses the Internet of Things and Ensemble Learning techniques. The suggested technique predicts rainfall in the near future by combining sensor data from the recent past with weather projected data. We utilised the Ensemble learning approach to forecast the likelihood of rain on that particular day. Rather than constructing separate specialised models and calculating classification metrices for each of them, the main purpose of this technique is to develop a single model that trains numerous models and classifies the output based on their aggregate majority of votes for each output class. Forecasted rainfall possibilities are superior in terms of accuracy and mistake rate. A solo system prototype can also use the prediction method. The system prototype is low-cost because it is based on open-source technologies. We'd like to perform a water-saving study based on the suggested technique in the future, with more nodes and a lower system cost. The irrigation system automation we provided as part of our strategy performed wonderfully. It's also cost-effective. Using this technique, we can reduce the number of people needed in the fields for upkeep. This approach will not only irrigate the ground automatically based on the moisture level in the soil and the possibility of rain, but it will also send the data to the Thingspeak server, allowing the farmer to keep track of the land's status.

REFERENCES

[1]. Muhammad Ayaz, Mohammad Ammad-Uddin, Zubair Sharif, Ali Mansour And El-Hadi M. Aggoune (2019). “Internet-Of-Things (Iot)-Based Smart Agriculture: Toward Making The Fields Talk”. Special Section On New Technologies For Smart Farming 4.0.

[2]. HamzaBenyezz, Mounir Bouhedda, SamiaRebouh(2021). “Zoning irrigation smart system based on fuzzy control technology and IoT for water and energy saving”. Journal of Cleaner Production Volume 302, 15 June 2021, 127001

[3]. Vaishali S, Suraj S, Vignesh G, Dhivya S and Udhayakumar S (2017). “Mobile Integrated Smart Irrigation Management and Monitoring System Using IOT”. International Conference on Communication and Signal Processing, April 6-8, 2017, India.

[4]. Nurzaman Ahmed, Debashis De, and Md. Iftekhar Hussain (2018). “Internet of Things (IoT) for Smart Precision Agriculture and Farming in Rural Areas”. IEEE

INTERNET OF THINGS JOURNAL, VOL. 5, NO. 6, DECEMBER 2018

[5]. Jhonn Pablo Rodríguez, Ana IsabelMontoya-Munoz, Carlos Rodriguez-Pabon, Javier Hoyos, Juan Carlos Corrales (2021). "IoT- Agro: A smart farming system to Colombian coffee farms". Computers and Electronics in Agriculture Volume 190, November 2021, 106442

[6]. Md. Rezwan Hossain Naeem, Shadman Gawhar, Md. Belawal Hoque Adib, Sanjid Ahmed Sakib, Abir Ahmed and Nafiz Ahmed Chisty(2021). "An IoT Based Smart Irrigation System". 2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST)

[7]. H.G.C.R.Laksiri, H.A.C.Dharmagunawardhana, J.V.Wijayakulasooriya "Design and Optimization of loT Based Smart Irrigation System in Sri Lanka". 2019 IEEE 14th International Conference on Industrial and Information Systems (ICIIS), 18-20 Dec.,Peradeniya, Sri Lanka

[8]. Neha K.Nawandar, Vishal R.Satpute(2019). "IoT based low cost and intelligent module for smart irrigation system". Computers and Electronics in Agriculture Volume 162, July 2019, Pages 979-990

[9]. A. Anitha, Nithya Sampath, M.Asha Jerlin(2020). "Smart Irrigation system using Internet of Things". 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE) 978-1-7281-4142-8/20/$31.00 ©2020 IEEE 10.1109/ic-ETITE47903.2020.271

[10]. Thilina N. Balasooriya, Pranav Mantri, Piyumika Suriyampola (2020). "IoT-Based Smart Watering System Towards Improving the Efficiency of Agricultural Irrigation". 2020 IEEE Global Conference on Artificial Intelligence and Internet of Things (GCAIoT) | 978-1-7281-8420-3/20/$31.00 ©2020 IEEE | DOI: 10.1109 / GCAIOT51063.2020.9345902