Automation in Agriculture Using IOT and Machine Learning

Abhishek L, Rishi Barath B

Abstract -The purpose of this project is to improve the efficiency of the agriculture sector. In India, agriculture plays a vital role for development in food production. Internet of Things (IoT) is a milestone in evolution of technology. IoT helps us in many fields among which agriculture is one of the primary ones. With the help of IoT along with Machine Learning in the field of agriculture, we can increase the efficiency of crop production.Different weather parameters are taken into consideration with which the best suitable crop to be grown are predicted with the help of supervised learning like Decision Tree Classifier, Regression. With help of different sensors, the soil and atmospheric conditions are determined and transferred through multi-hop communication to the server in which monitoring of crops' health and control of irrigation system takes place. TDMA is used for the above purpose.

Index terms - Internet of Things, Machine Learning, ZigBee, Wireless Sensor Networks, Smart Agriculture, Multi-hop communication, TDMA

I. INTRODUCTION

This is one of such projects that solely concentrates on making the farming process more efficient and accurate by analysing the different conditions of a farmland.This project not only helps in easing the farmers' jobs and making their life better but also helps in saving variety of environmental resources. Agriculture is not a very promising occupation, as there is no guarantee that crops will be healthy and farmers could make profit out of it, most of the farmers do cultivation alongside other occupation instead of taking farming as primary occupation. This leads to poor care of the crops as the time that these farmers could allocate to these crops are very limited in their busy schedule.So, this is where this project comes into play. The main motivation of this project is provide these farmers with an automation system that tries to resolve all the above mentioned issues thereby saving time, resources, money and man power. The main goal of this project is to act as a complete kit for the farmers by aiding them with the agricultural process from beginning to end.

II. METHODOLOGY

This project includes suggestion of crops that would be best suitable for cultivation in a particular farmland after the consideration of the weather and climatic conditions of that geo-location. Different supervised machine learning methodologies come into play for the above prediction.

This project also includes an automated irrigation [7] system for efficient water management after the continuous monitoring of the soil conditions during the entire growth of the crops. These include soil parameters like temperature and humidity, soil moisture content, soil water level, etc...

Arduino UNO or a custom-built MEGA is used to collect these sensor values and transmit them to the Raspberry Pi in which the Apache Web server is set up. Raspberry Pi also has SQL database for storage or container. The ZigBee module has been used to establish a communication link between the sensor arrays and the server. The farmer can access the server about the field condition anytime, anywhere thereby reducing the man power and time.

Multi-hop communication is implemented to extend the range of communication. The data from the sensor arrays are transmitted through their neighbouring sensor arrays which transmit them further to their neighbours. In this way, the data reaches the server at the end after travelling through multiple hops. The main purpose of this type of serial transmission involving consecutive sensor arrays is to increase the overall range and to ensure the reception of data from all sensor arrays at the server end in a fool proof way.

Since the role of a sensor array in the above multi-hop communication is to transmit its own data as well as transmit the data received from its neighbours, there is a possibility of ambiguity and alternation of data. However,TDMA [10] is used to eradicate the above mentioned ambiguity.

III. ARCHITECTURE DIAGRAM

The project architecture diagram in Fig.1 clearly depicts the entire architecture of this project. It indicates the flow of the data in the project.

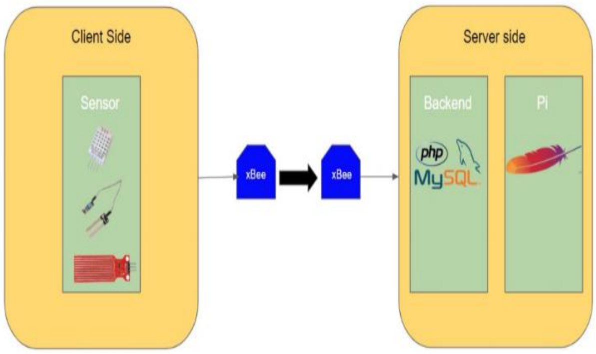

Fig.1. Architecture Diagram

Revised Manuscript Received on June 05, 2019
Abhishek L, School of Computing Science and Engineering, Vellore Institute of Technology Chennai, Chennai (Tamil Nadu), India
Rishi Barath B, School of Computing Science and Engineering, Vellore Institute of Technology Chennai, Chennai (Tamil Nadu), India

As the above picture explains, the sensor values are retrieved from the soil which are concatenated and sent to either the server or the next consecutive sensor array through ZigBee module. These sensors include temperature-humidity sensor, soil moisture content sensor and water level depth detector.

These data are received and transferred to the succeeding sensor arrays to enable multi-hop communication. The data are received by Raspberry Piin which the database resides and Apache Web server is hosted.

IV. TECHNICAL REQUIREMENTS

A. Hardware Requirements

Temperature and humidity sensor-DHT22 is a low cost digitalsensor that uses a thermistor to measure the air in the surrounding and also a capacitive humidity sensor to measure humidity. A capacitive humidity sensor measures relative humidity with the placement of a thin strip of metal oxide between two electrodes. The metal oxide's electrical capacity changes with atmosphere's relative humidity.This capacitive humidity sensor measures the relative humidity in the range from 0% to 100%. The ratio of moisture in the air to the highest amount of moisture at a particular air temperature is called relative humidity.Ideal input voltage range is 3 to 5V and maximum current permissible is 2.5mA. It is good for -40^{o}C to 150^{o}C temperature readings with +0.5 or -0.5^{o}C accuracy and 0-100% humidity readings with 2-5% accuracy.

Soil Moisture Sensor–It measures the volumetric water content in the soil.It uses capacitance to measure the water content of soil, which is by measuring the dielectric permittivity of the soil, which is a function of the water content. This makes it ideal for the analysis of soil.This sensor contains two probes that are used for measuring the volumetric content of the water. These two probes allow the current to pass through the soil. Based on the resistance value, the moisture value is measured. That is, when there is more water content in the soil, more electricity is said will be conducted as the resistance will be less. Thus, the moisture level become higher. Electricity will be conducted comparatively poorly in dry soils as there will not be enough water. Thus the resistance is aid to be high and hence the moisture level is low.Generally, these moisture values will be measured in percentage, that is, mapping the values from 0 – 1023 to 0 – 100 with which further processing is done effectively. But in our project, we mapped this 0 – 1023 values to 100 – 0 so that values near zero indicate that moisture content is very less and hence soil is a dry soil and values near 1023 indicate that the moisture content is very high and thus the soil is comparatively wet.

Water Level Depth Detector - This sensor is used to measure the water level depth in the soil. This sensor is helpful during the irrigation process. With help of this sensor it is easy to identify the water content level in the soil with which the amount of water need to be delivered to the fields during irrigation process could be determined. This ensures the delivery of no excess water and also helpful in ensuring absence of shortage of water to the crops in the fields.This water level depth detection sensor has a DC operating voltage of 3V to 5V and operating current less than 20mA. The operating temperature for this sensor is around 10^{o}C to 30^{o}C. With a detection area of 40mm x 16mm, it works well with a non-condensing humidity of 10% to 90%.

Arduino UNO -Arduino UNO is a microcontroller board based on the microchip Atmega328P.It has an operating voltage of 5V and input voltage can vary from 7V to 20V. For the 3.3V pin, the DC current is 50Ma whereas DC current per I/O pin is 20mA. It has a flash memory of 32KB in which 0.5KB is used bootloader itself.Arduino UNO has a static RAM (SRAM) of 2KB where SRAM is random access memory (RAM) that retains data bits in its memory as long as power is being supplied. The clock speed for the UNO is 16MHz.In this project, UNO is used to collect the data from the different sensors (used to collect the condition of soil) present in the sensor array. UNO also aggregates these data and transmits them using the wireless communication module named ZigBee.

Raspberry Pi -The Raspberry Pi is a low cost, credit-card sized computer that plugs into a computer monitor or TV, and uses a standard keyboard and mouse.It also has very low power consumption between 0.5W and 1W.New Out Of Box Software (NOOBS) gives the user a choice of operating system from the standard distributions. Raspbian is the recommended operating system for normal use on a Raspberry Pi.In this project, Raspberry Pi is mainly need for two purposes. First, server and second for storage. Apache Web Server is hosted on this Pi [11] along with SQL database for storage of the data. Hence Raspberry Pi is placed at the centre of the field amidst all sensor arrays as shown in Fig.4

ZigBee -Zigbee is an IEEE 802.15.4-based specification for a suite of high-level communication protocols used to create personal area networkswith small, low-power digital radios, such as for home automation, medical device data collection, and other low-power low-bandwidth needs,. Zigbee is for low-data rate, low-power applications and is an open standard.Zigbee is built for control and sensor networks on the IEEE 802.15.4 wireless standard for wireless personal area networks (WPANs). The Zigbee WPANs operate on 2.4 GHz, 900 MHz and 868 MHz frequencies. There are several advantages of using the ZigBee protocol when comparing with other protocols for WSN [1][2]. ZigBee is standardized at all layers. In this project, the entire communication let it be among the sensor arrays or between the sensor array and server, is entire done wirelessly using ZigBee. ZigBee is responsible for the entire data transmission.

B. Software Requirements

Raspbian OS - Raspbian is a Debian-based computer operating system for Raspberry Pi. There are several versions of Raspbian including Raspbian Stretch and Raspbian Jessie. Though Raspbian is not intended to run the Pi like a desktop computer, it provides the users with LXDE desktop environment. The Pi does not have a great deal of processor speed or memory, but it does have enough resources to run LXDE and a handful of applications like the simple Epiphany web browser and many more.

Arduino Software -Arduino is an open-source electronics platform based on easy-to-use hardware and software. It's intended for anyone making interactive projects. By receiving inputs from sensors, Arduino senses the environment and affects or influences its surroundings with different actions and actuators. Arduino is needed to compile and upload the necessary codes to run the UNO.

Apache Web Server -Apache web server is free and open-source cross-platform web server software developed and maintained by Apache Software Foundation. It runs on 67% of all webservers in the world. It is fast, reliable, and secure. It can be highly customized to meet the needs of many different environments by using extensions and modules.

In this project, this server is needed to monitor the received data and perform the necessary tasks after the data interpretation and processing. The server is what is responsible for the monitoring the health of the crops and also to take control of the automatic water irrigation system. In this project, the server is hosted on the Raspberry Pi [11].

SQL Database -Structured Query Processing (SQL) is a domain-specific language used in programming and designed for managing data held in a relational database management system (RDBMS). SQL programming can be effectively used to insert, search, update, and delete database records. In fact it can do lot of things including, but not limited to, optimizing and maintenance of databases.

SQL is needed to store the data indicating the soil conditions (sensor values from the sensors) of the farmland safely for analysis and further processing.

Languages & Packages – The language PHP is used as the server scripting language in this project. Python is used to perform different machine learning algorithms for prediction of crop. Scikit-Learn[9], pandas, numpy, matplotlib are some necessary packages for the implementation of the same. HTML, JavaScript and CSS are combinedly used to provide an interactive UI to the users/farmers.

V. PROJECT MODULES

A. Data Processing

This module aids the farmers by suggesting with the best suitable crop to grow along with the production possible. To achieve the same, different predictive methodologies including both classification and regression are used. Attributes of the place in which the crop is intended to be grown are taken and fed into the already learned machine learning classifier with which the best suitable crop for growth in that farmland is predicted. These attributes include weather conditions, average temperature, average humidity, moisture content in the air, etc…

State_Name	District_Nan	Crop_Year	Season	Crop	Area	Production	Production/A	Min T	Max T	Min H	Max H
Andhra Pradesh	ANANTAPUR	2011	Whole Year	Banana	8000	505176	63.147	20	19	94	91
Andhra Pradesh	ANANTAPUR	2013	Whole Year	Banana	8661	483881	55.869	20	20	93	94
Andhra Pradesh	ANANTAPUR	2005	Whole Year	Banana	303	16231	53.568	17	16	94	94
Andhra Pradesh	ANANTAPUR	2009	Whole Year	Banana	5010	265981	53.09	17	16	91	95
Andhra Pradesh	ANANTAPUR	2002	Whole Year	Banana	19	1006	52.947	17	20	90	94
Andhra Pradesh	ANANTAPUR	2004	Whole Year	Banana	324	16897	52.151	18	19	91	92
Andhra Pradesh	ANANTAPUR	2010	Whole Year	Banana	4416	223825	50.685	17	19	91	90
Andhra Pradesh	ANANTAPUR	2006	Whole Year	Banana	289	14141	48.931	19	20	90	94
Andhra Pradesh	ANANTAPUR	2008	Whole Year	Banana	3603	160939	44.668	20	16	91	93
Andhra Pradesh	ANANTAPUR	2003	Whole Year	Banana	36	1302	36.167	18	17	92	92
Andhra Pradesh	ANANTAPUR	2000	Whole Year	Banana	769	11693	15.205	18	18	95	92
Andhra Pradesh	ANANTAPUR	2001	Whole Year	Banana	1582	462	0.292	18	17	93	94
Andhra Pradesh	ANANTAPUR	2013	Kharif	Cabbage	2	38	19	15	25	80	100
Andhra Pradesh	ANANTAPUR	2013	Rabi	Cabbage	11	207	18.818	19	22	91	90
Andhra Pradesh	ANANTAPUR	2014	Rabi	Cabbage	2	25	12.5	25	23	94	96
Andhra Pradesh	ANANTAPUR	2014	Kharif	Cabbage	10	123	12.3	23	18	89	88
Andhra Pradesh	ANANTAPUR	2013	Kharif	Cotton(lint)	37661	77316	2.053	21	37	90	100
Andhra Pradesh	ANANTAPUR	2013	Rabi	Cotton(lint)	62	127	2.048	22	33	100	100
Andhra Pradesh	ANANTAPUR	2014	Kharif	Cotton(lint)	73734	103779	1.407	23	34	95	97
Andhra Pradesh	ANANTAPUR	2014	Rabi	Cotton(lint)	884	1244	1.407	34	25	92	95
Andhra Pradesh	ANANTAPUR	2000	Kharif	Cotton(lint)	12025	16128	1.341	30	27	95	95

Fig.2. Dataset Sample

The above is a standard dataset retrieved from a government website [8]. A classifier model is trained with the above dataset which is further used for prediction. It is observed that the ID3 Decision Tree Classifier [9]gives a decent percentage of accuracy (around 94%) as it classifies based on the range of values that each attribute can fall (homogeneity) and that is what we ideally need here because each crop does not have a specified temperature, humidity or moisture content value but instead each crop can have a range of values for each of the above attributes.

Attributes of crops along with crop name are taken and with the help of regression, production per unit area is predicted. These attributes include the place of growth, growing season, growing duration and many more. Once the production per unit area is predicted, the total production possible could be determined with the knowledge of the total area of the field.

The graph plot in Fig.3depicts the different values of production per unit area observed from different farmlands. The same is plottedcrop wise in each state.Production per unit area is taken along Y-axis. Production per unit area value is found after real time plantation on different lands in different areas in a particular state. Number of data points is taken along X-axis. Such graphs are plotted for all states in India where the graph for Uttar Pradesh is picturized in Fig.3.

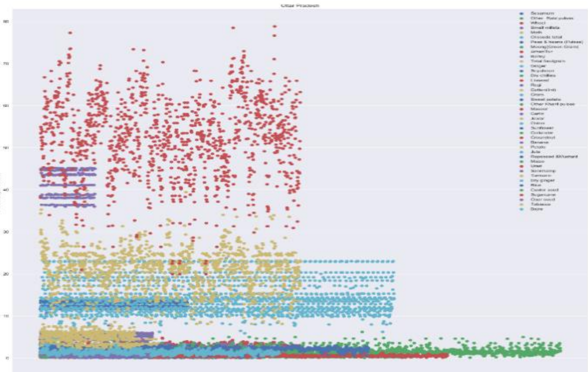

Fig.3. Production per unit area plot – Uttar Pradesh

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

B. *Data Collection*

Data collection plays an important role in this project as without data, the whole project loses its meaning. The data here refers to the data or values retrieved from different sensors used to analyse the soil conditions.

These sensors are fused into one single rod to form a sensor array which is further connected to the Arduino UNO board. Many such sensor arrays are placed equidistantly in a matrix structure with the server at the centre of the field as shown in Fig.4.

The sensor values are retrieved from the above sensors with the help of Arduino UNO. UNO then aggregates these values into one unit block and make it ready for the transmission to the server. Data from these sensors clearly picturises the exact condition of soil at the time of data retrieval. After the data aggregation, the data is transmitted to the UNO present in the neighbouring sensor arrays with the help of ZigBee for the multi hop communication.

The Fig.4 shows the arrangement of the sensor arrays in the agricultural field. The sensor arrays are placed uniformly in a matrix like structure with Raspberry Pi at the centre of the matrix.

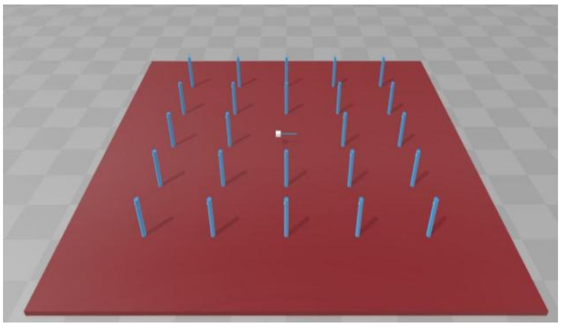

Fig.4. Sensor-Array arrangement in the farmland

C. *Data Transmission*

The transmission of data between the sensor arrays and the Raspberry Pi is one of the primary modules of this project. The Raspberry Pi has the Apache web server hosted on it [11] along with the SQL Database for storage of data.

The Pi is placed at the centre of the field to reduce the cost and risk of transmission. Placing the server at the middle of the field also reduces the total number of hops needed for all values from all sensor arrays to reach the server thus by also reducing the required power. In this way, placement of the server and storage at the middle of the field acts as a major advantage.

The distance between a sensor array and the server need not be lesser than the range of the ZigBee module and hence the concept of multi hop communication comes into play to extend the transmission range. The shortest path between a sensor array and the server through which the data has to be transmitted is found using the Dijkstra's Shortest Path Algorithm [12].

Once the path is computed, the data is transmitted along the path through the nodes defined by the path. Each sensor array or node in that computed path is activated to receive and transmit to the succeeding node in the path. In this way, the data could be transmitted from one end to other without any hassles.

As the sensor arrays plays both roles, that is, transmission of the received data as well as transmission of its own data, ambiguity and data alteration could occur due to interference of the data packets. To eradicate this ambiguity, concept of Time Division Multiple Access (TDMA) [10] is used.

A particular time slot is allotted for each end-to-end data transfer that is, from each sensor array to the server. During a particular time slot, the nodes participating in the path of transfer are only activated. This way of providing a dedicated time slot for data transfer for each sensory array eliminates data modification and ensures integrity of data.

D. Data Interpretation

Once the data is received on the server end, the server starts interpreting the received data. Server continuously analyses and monitors the received data for anomalies. Server always ensures whether the soil conditions are as desired and as required for the crop grown.

E. Task Execution

Different tasks are executed by the server depending upon the soil conditions. One of the main tasks by the server after the analysis of the data is to take care of the irrigation process. The server will decide and take the call on how much amount of water is needed for irrigation on that day after analysing the daily water requirement for the crop and the amount of water that is already present in the soil.

Server also plays the role of setting the valve of the automated irrigation system ON and OFF to irrigate the fields. The duration of irrigation is determined by the amount of water required to irrigate for the day which is done during the analysis of data.

Also, the server must raise an alarm and indicate the farmer if any of the interpreted data goes out of the desired conditions.

VI. RESULTS

The sensor data are collected for a time frame and the values from these sensors are visualized along with time. Fig.5 shows the plot of sensor values versus time.

The sensor values are read from the sensors for a period of time and visualized to understand how the soil conditions change for that period of time. The same is shown in Fig.5. The sensor values used for the same include temperature and humidity of the atmosphere and the moisture content in the soil. In between the observed time, certain amount of water is added to the soil and hence a steep rise in soil moisture sensor values could be observed in the plot (Fig.5). Also, temperature of the air did not change much even after the addition of water in the soil. Humidity of the soil also remains constant without showing multiple major variations.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

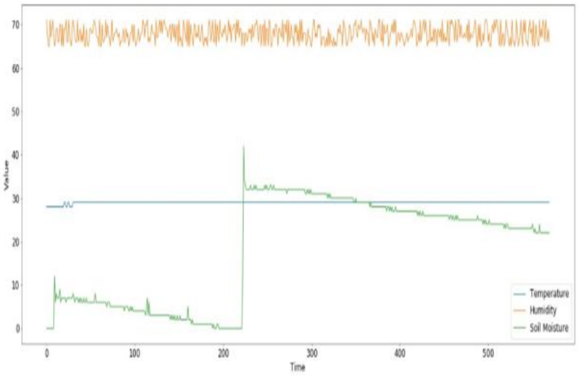

Fig.5. Sensor Data Visualization

VII. CONCLUSION

In this way, the projectstrives to bring efficiency and accuracy in the field of agriculture by automating the entire agricultural process with the help of Internet of Things and Machine Learning. It also tries to avoid the over exploitation of critical resources which may get extinct in near future. This acts as a complete package that every farmer would desire to possess. This project is completely oriented towards the farmer welfare and agricultural development. It aids the farmers with the complete process of farming from the start till end. With taking good care of these crops, it helps the farmers to come out of their poverty by providing with a good amount of yield at the end. It also ensures the health and nutrition of the crops.Solving all of the above issues is not just an advantage of this project, but also a necessary thing for the betterment of any nation's welfare. This project not only saves money and resources but also time and manpower. Since this project possess so many advantages, every farmer's dream would be to have this implemented in their fields. Since this project is also cost effective and affordable by most of the farmers in India, there is no doubt that this project would be a market hit.

'Go with Automated-Agriculture..!!
Make agriculture simple and effective..!!'

REFERENCES

1. G. Fellidis, V. Garrick, S. Pocknee, J.V. Stafford et al., "How wireless will change agriculture", Precision Agriculture '07 - Proceedings of the Sixth European Conference on Precision Agriculture (6ECPA), pp. 57-67.
2. Ning Wang, Naiqian Zhang, Maohua Wang, "Wireless sensors in agriculture and food industry-Recent development and future perspective", Computers and Electronics in Agriculture, vol. 50, pp. 1-14, 2006.
3. O. Vermesan, P. Friess, Internet of Things: Converging Technologies for Smart Environments and Integrated Ecosystems, 2013, [online] Available: http://www.internet-of-things-research.eu.
4. Z. Nakutis, "Remote Agriculture Automation Using Wireless Link and IoT Gateway" Infrastructure Conference Proceedings, 2015 26th International Workshop on Database and Expert Systems Applications (DEXA), 2015
5. Li Tan, "Cloud-based Decision Support and Automation for Precision Agriculture in Orchards", 2016
6. SanatSarangia, JayalakshmiUmadikar, Subrat Kara, "Automation of Agriculture Support Systems using Wisekar: Case study of a crop-disease advisory service", 2016
7. Dursun, M. and Ozden, S. (2019). A wireless application of drip irrigation automation supported by soil moisture sensors. [online] Academicjournals.org. Available at: https://academicjournals.org/journal/SRE/article-abstract/4644D6C20216
8. data.gov.in. (2019). data.gov.in. [online] Available at: https://data.gov.in/
9. scikit-learn, Scikit-learn: Machine Learning in Python, Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J. and Passos A., Cournapeau D, Brucher M, Perrot M, Duchesnay E. Journal of Machine Learning Research Volume 12 pages 2825--2830, year=2011
10. Vergados, D.J., Sgora, A., Vergados, D.D. et al. TelecommunSyst (2012) 50: 181. https://doi.org/10.1007/s11235-010-9397-9
11. https://www.raspberrypi.org/documentation/remote-access/web-server/apache.md
12. https://en.wikipedia.org/wiki/Dijkstra%27s_algorithm

Published By:
Blue Eyes Intelligence Engineering & Sciences Publication