**IOT AND WIRELESS SENSOR NETWORK (WSN) BASED DATA LOGGER SYSTEM WITH RAIN PREDICTION USING ML**

**[1] M V Narasimha Prasad, [2] Lavanya R, [3] Aakarsh Satyam, [4] Naveen N N M,**

**[5]** **Mr. Ramakrishna S**

[1][2][3][4] B.E. Students, [5] Associate Professor, Department of ECE,

Dayananda Sagar Academy of Technology and Management, Bengaluru-560082

[1] [mvnp.1999@gmail.com](mailto:mvnp.1999@gmail.com)  [2 ] lavanya4520@gmail.com

[3] aakarsh.satyam@gmail.com, [4] narendranmrl@gmail.com, [5] sramkrish2121@hotmail.com

**Abstract –** This paper brings out how agriculture has been the back bone of India as it is the case with many countries. But the use of modern technology for this sector’s development has been slow. To increase the availability of technology for the improvement of crop yield should be kept in mind and worked on. For ages the traditional method of farming has been the practice in India.

Here, farmers have very little knowledge on rain prediction or on the change of soil fertility and other important aspects of farming. Because of this, the current trend of increase in demand for food and agricultural yield is hard to be satisfied. Hence proper research is done on various technology that can be implemented in the field of agriculture which will aim at smart farming, sending the data to cloud using IoT and Firebase, using ML algorithms to predict rain or other desired features so we can manage with limited resources and get improved crop yield while managing the limitations.

***Keywords –*** *Wireless Sensor Network, Linear Regression, Naïve Bayes Algorithm, Smart farming, Rain prediction.*

1. **INTRODUCTION**

Agriculture has not changed much since the beginning of farming. With advancements in technology, we have come to such an era that we can improvise the methods and the way the farming has been done. Advancement in technology in every field has been on an increasing level from the past few decades. It is done in a few important sectors; which will help us with the issue of comprehending and predicting these climate occurrences that is caused by changes in global temperatures, natural disasters in the last three years, rising sea levels and shrinking polar region. Prediction is crucial, and they can be run and simulated as computer models to forecast climatic variables such as temperature, precipitation, and rainfall.

Using this for rain prediction in the area which has to be cultivated is a huge improvement in this field. Another technology called Wireless Sensor Network (WSN) can be used to collect the data from agricultural field in an efficient, cost-effective way. Implementing this will enable us to collect data from sensor nodes and transmit the same using the master node. Once this technology is deployed, we can send the data to cloud using Firebase and apply ML algorithms on it. This enables us to predict the feature we require based on the data collected. Hence, the data can not only be collected with minimal corruption, but it can also be logged and used to take decisions in such a way that it improves the quality of farming, increases the yield along with saving water as much as possible. Such a method is an efficient method for the current trend of increasing population and demand with limited resources which are being depleted. This data driven farming should help us achieve our goal.

1. **OUTLINE OF THE PROJECT**

Our proposal relies on the idea of a wireless sensor network (WSN), in which a sensor node gathers information from several sensors such as humidity, temperature, and soil moisture content.A sink node, also known as a base station or master node, gathers data from several similar sensor nodes and delivers it to a cloud database where it may be analyzed and displayed to make predictions or automate and manage real-time machinery.

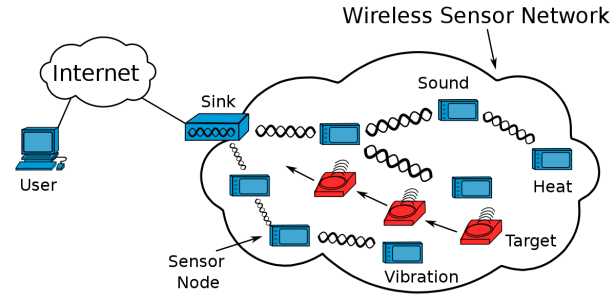
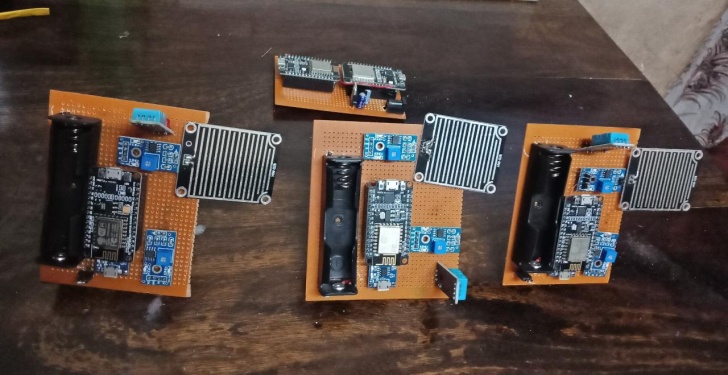


Fig.1: Overview of Wireless Sensor Network (WSN)

The operation of a wireless sensor network (WSN) is depicted in the previous figure ("Fig. 1"). To interact with users through the internet, a sensor node made up of several sensors is linked to a sink node.Our technology has the advantage that the sink node is not reliant on the quantity of sensor nodes, making it simple to expand the kind or number of sensors without altering the sink node. For the communication portion, there is a proprietary protocol called ESP-NOW [2] that is based on IEEE 802.11 vendor-specific action frames. The sensor node may become inactive once the data has been transferred from it. The data transfer operation will repeat once the user-specified time interval has passed. With this protocol, the sensor node uses just 170mA while transmitting data and 80mA when it is idle.

Fig 2: IoT based Agriculture system(overview)

The image of our functioning prototype of a WSN-based data logger, which comprises of sensor nodes and sink nodes, is displayed in "Fig. 2."

The sink node is made up of a master receiver (NodeMCU), which collects data from sensor nodes and transmits it via UART to an ESP8266-01 Wi-Fi module, which then sends the data to the internet. Having these two distinct systems has the benefit of keeping the system very modular. The master receiver may still receive data from the sensor node without any change if this WSN-based data logger has to be put in a remote location without access to the internet. We may send the information to a local server or Raspberry Pi, which can do the necessary calculation on-site and display the real-time data to the user over a local network or a physical dashboard.

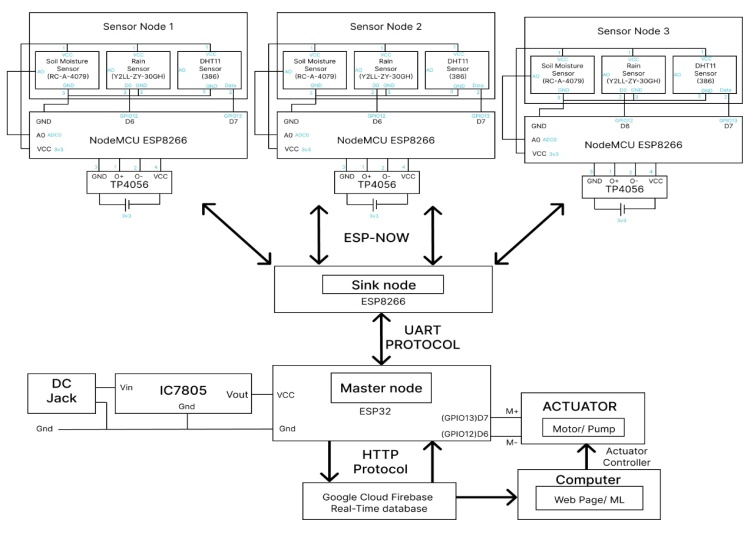
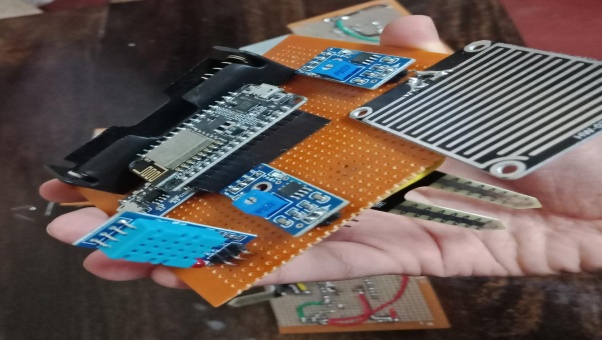


Fig 3: WSN based data logger system (circuit diagram)

As shown in figure 3, the overview of the circuit diagram of WSN based data logger system. Our sensor nodes consist of three sensors and a microcontroller. A single battery is used to power the circuit. Although sensor node can contain any sensor the user wishes, we have used DHT11 sensor, rain sensor and resistive moisture sensor**.**

# HARDWARE IMPLEMENTATION

**A. Sensor node**

For sensor node, as the name indicates, it contains various sensors that are required to measure various parameters of the agricultural field. As mentioned before, each piece of land is volatile. It is different; hence according to the need of the farmer, they can put the appropriate sensor to collect the parameters that are valuable to them. In our case, we have used DHT11 sensor, rain sensor and resistive moisture sensor. The DHT11 sensor is used to measure temperature and humidity of the environment.

Fig. 4. Sensor Node of the WSN data logger

As shown in Fig. 4, the sensor node contains afore mentioned sensors. The rain sensor measures the rain. Its output is binary that is 0 or 1 where 0 indicates no rain and 1 indicates that there is rainfall. This can be improved as there is a potentiometer provided with the sensor so we can adjust the sensitivity of the rainfall according to our need. The final sensor is the resistive moisture sensor. This sensor measures the soil moisture content. Again, the sensitivity of this can also be adjusted. Along with this, the sensor node contains a single cell rechargeable battery with TP4056 circuit to charge it and NodeMCU.

**B. Sink node**

The function of sink node is basically to collect the data from all the sensor nodes connected and send it to master node. The sink node is collecting all the data from sensor nodes using ESP-NOW protocol. Because the ESP NOW protocol only functions amongst the family of ESP created by Espressif, NodeMCU is utilised as a microcontroller [3]. The ESP8266-01 is used in conjunction with the microcontroller to transport the acquired data from the microcontroller to the firebase database. Although NodeMCU has integrated Wi-Fi, we still utilise an external Wi-Fi module since we can quickly convert to other communication types [GPS, LORA, ZIGBEE, etc.] because we employed UART connection between the microcontroller and the Wi-Fi module in the event that Wi-Fi communication is not available [4].

**C. Master node**

The master node contains ESP32 and its function is to collect data from Sink Node using UART protocol and to send the same data to a cloud storage using HTTP protocol. The sink node and the master node are physically connected using the Rx and Tx pins and hence, they are together.

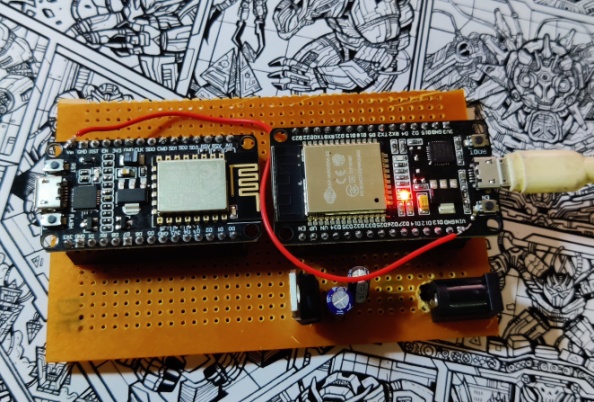


Fig. 5. Sink Node and Master Node

As shown in the figure, the sink node and master node are physically together and their function is to collect data from sensor node and to send it to cloud storage, Firebase from Google. A sensor node may be designed to provide data at any interval, and because the master is always powered, it can receive data from any sensor node at any time.

# FIRMWARE IMPLEMENTATION

This prototype's programming was carried out using an Arduino IDE that also has an installed ESP8266 core (the ESP8266 core is described in depth in [3]). Adafruit DHT sensor library for DHT sensor and a straightforward analogue read were used to read the sensor data from the sensor node for the moisture sensor. ESP NOW was used to transport data from the sensor node to the sink node [5].

Both the sensor node and the sink node have a structure (struct) in the code that specifies the precise data types and structure size. Data from the sensor node is transmitted along with this structure, together with the mac address and an ID to identify the sensor node from which the data was received. When the data is received, it is digested into a JSON object by Mobizt [6] before being transferred to Firebase using Firebase-ESP8266. Sink node handles the dynamic nature of adding more nodes since it can distinguish between various nodes, produces a new JSON Object with a distinct node name, and then pushes this data to Firebase.

# COULD DATABASE FOR DATA STORAGE

The data collected has to be stored to do any further analysis with it. For storage, we have used could storage. This has many advantages like it can be accessible all the time; it can be accessed from anywhere, etc. Firebase, a google cloud data storage provides data storage facilities along with further expansion into running AI/ML algorithms to it. It also provides real time viewing of data collected.

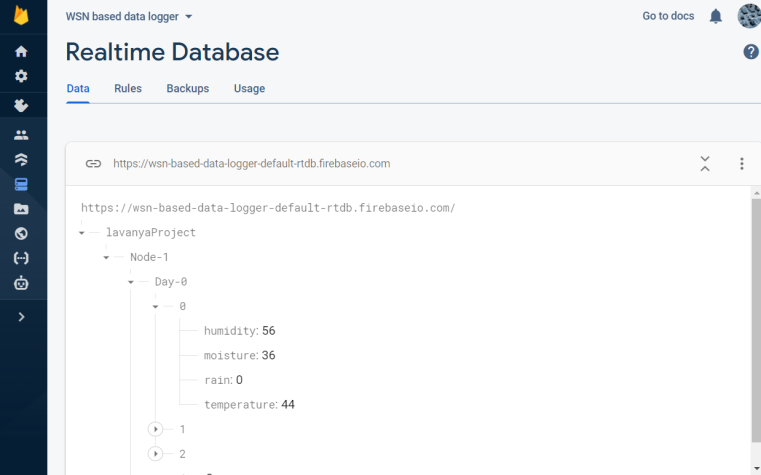


Fig. 6. A snapshot of could storage showing the data in realtime

As shown in the firuge (Fig. 6), the data is collected real time and stored in cloud storage. This is a service provided by Google to collect, store and monitor data. Google has created a platform called Firebase for building mobile and web applications[7]. Since Firebase makes it simpler to host websites and create mobile applications using an existing database, it was picked.

The user may easily and securely log in using their email to access their statistics thanks to Firebase Authentication. This is also readily expandable to support additional login ways, such as Facebook, phone, or any other.The data stream originating from the sensors is kept in a real-time database. Any change in the data will be immediately displayed on the dashboard because the database is real-time.

1. **ML DEVELOPMENT**

The Machine learning model for predicting rain at a particular time is going to be done by the linear regression model. Linear regression is the method of plotting data points onto a graph. Linear regression can be diverse which has multiple independent variables used as input features and simple linear regression which has only one independent or input feature. Both linear regressions have one dependent variable which can be forecasted or predicted based on the input features. This paper presented the multivariate linear regression because multiple environmental variables or features were used to predict the dependent variable called daily rainfall amount. Linear regression is a supervised machine learning technique used to predict the unknown daily rainfall amount using the known environmental variables. The multivariate linear regression used multiple explanatory or independent variables (X) and single dependent or output variable denoted by Y. Hence, the general equation of the multiple linear regression is given as:

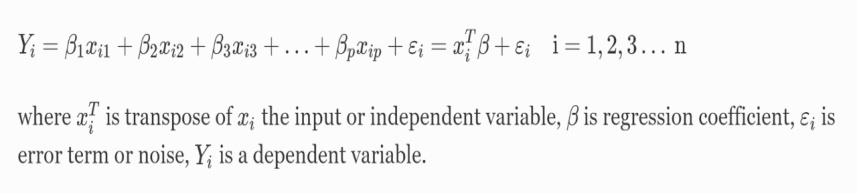


Fig 7: Linear Regression equation

where *xiT* is transpose of *xi* the input or independent variable, *β* is regression coefficient, *εi* is error term or noise, *Yi* is a dependent variable.

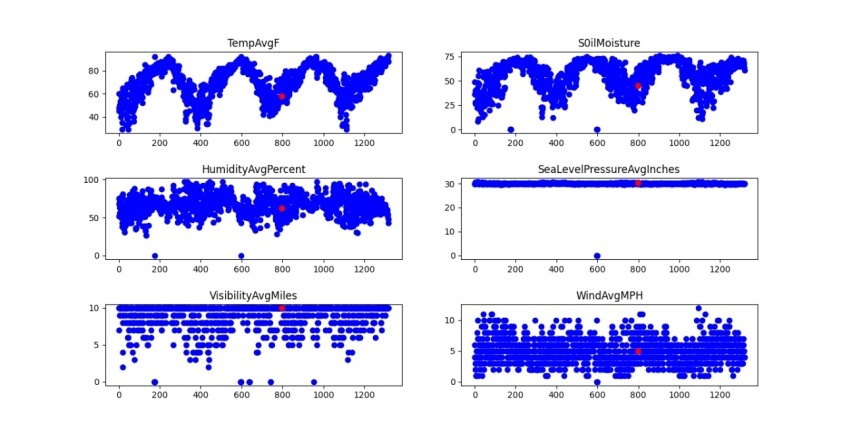
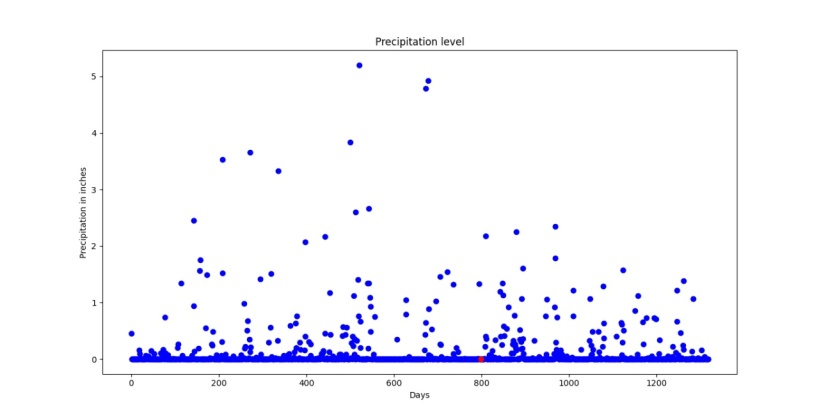


Fig 8: Rain prediction using ML

The size of the data set collected from the dataset for this paper was appropriate to use the machine learning algorithms called multivariate linear regression that can estimate the daily amount of rainfall in the region. This algorithm can show how strongly each environmental variable influences the intensity of the daily rainfall.

**VII. RESULT**

The result of the experiment performed showed good potential to be adapted in real world. The data obtained after performing the experiment is as tabulated below.

**TABLE I. TABULATION OF COLLECTED DATA ON TESTING**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Display of collected data | Node - 1 | | | |
| Day - 0 | | | |
| At 4:00pm | At 5:00pm | At 6:00pm | At 7:00pm |
| Temperature (in 0C) | 25.3 | 24.0 | 23.7 | 22.4 |
| Humidity (in %) | 43 | 45 | 56 | 67 |
| Moisture (in %) | 87 | 84 | 81 | 79 |
| Rain (0 or 1) | 1 | 0 | 0 | 0 |

In "Table I," the ambient temperature, humidity, soil moisture content and rain fall are all tabulated. On the first day, the information was gathered from sensor node 1 (day 0). To note the time at which this data was obtained, the sensor inserts a timestamp. Data can be received from the other nodes and forwarded to the master node in a manner similar to how the data from the first node is tabulated.

The designed firmware is extremely dynamic and can adapt to numerous nodes. The master node can manage an increase in the number of hardware nodes or the addition of new sensors to an existing sensor node with little to no modifications to the firmware. Due to the usage of Firebase Hosting, which is accessible from anywhere in the globe, the user has a lot of flexibility when accessing the data.

**VIII. CONCLUSION**

Agriculture plays an important role in every human’s life. It is one of those fields which require dire need of technology especially when India’s population is growing. Hence, our project aims to collect, monitor and store data accurately. It is further subjected to ML to predict rain to demonstrate how collected data can be used further. The project also aimed in being flexible to be able to adapt in real life agricultural field.

**IX. FUTURE SCOPE**

As demonstrated before, the collected data can be subjected to further AI/ML algorithms to implement useful strategies in agricultural field. Or it can also be used to predict rain, storm, etc. events. The project can also be developed further and additional physical layers can be added to shield the modules like sensor node and master node from environmental harshness.

**ACKNOWLEDGMENT**

We would like to thank our guide Mr. Ramakrishna S for guiding our project. We would also like to thank our families and friends for their support and encouragement.

# REFERENCES

[1] Bhatnagar R., Gohain G.B. (2020) Crop Yield Estimation Using Decision Trees and Random Forest Machine Learning Algorithms on Data from Terra (EOS AM-1) & Aqua (EOS PM-1) Satellite Data. In: Hassanien A., Darwish A., El-Askary H. (eds) Machine Learning and Data Mining in Aerospace Technology. Studies in Computational Intelligence, vol 836. Springer, Cham. <https://doi.org/10.1007/978-3-030-20212-5_6>

[2] Jiannong Xin, Fedro Zazueta. (2016) Technology Trends in ICT – Towards Data-Driven, Farmer-Centered and Knowledge-Based Hybrid Cloud Architectures for Smart Farming, Vol. 18 No. 4 (2016): CIGR Journal.

[3]<https://www.espressif.com/en/products/software/esp-now/resources>

[4]<https://arduino-esp8266.readthedocs.io/en/latest/>

[5]<https://docs.espressif.com/projects/espidf/en/latest/esp32/apireference/network/esp_now.html>

[6]<https://github.com/mobizt/Firebase-ESP8266>

[7]<https://en.wikipedia.org/wiki/Firebase>