iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL
FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 Issue: I Month of publication: January 2023

DOI: https://doi.org/10.22214/ijraset.2023.48841

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue I Jan 2023- Available at www.ijraset.com

Machine Learning Inspired Smart Agriculture System with Crop Prediction

Rajdeep Chatterjee[1], Oindrila Das[2], Ridam Kundu[3], Soumik Podder[4]
Department of Electronics and Communication Engineering, Guru Nanak Institute of Technology, 114/F, Nilgunj Road, Sodepur, Panihati, Kolkata-700114

Abstract: This paper presents Machine Learning Inspired Smart Agriculture System with Crop Prediction that will assist farmers in getting Live Data (Temperature, Soil Moisture, level of Nitrogen, Phosphorous and Potassium present in the soil, Alkalinity and Acidity in the soil etc.) for efficient environment monitoring which will enable them to increase their overall yield and quality of products. The Machine Learning Inspired Smart Agriculture System with Crop Prediction is proposed where Machine Learning Technology is hybridized with different Sensors and a Wi-Fi module that will yield live data feed using online software Localhost and PHPmyadmin. The product being proposed is tested on Live Agriculture Fields giving high accuracy over 98% in data feeds
Keywords: Machine Learning, Smart agriculture, Crop prediction, Sensor, Localhost, PHPmyadmin, environment monitoring

I. INTRODUCTION

After the green revolution, there has been an increasing trend in the use of machines in farm operations. This has led to the mechanization of Indian agriculture. Punjab, Haryana, Uttar Pradesh, River valleys of Andhra and Tamil Nadu are major agriculturally mechanized areas of India. Agriculture plays a vital role in the Indian economy. Over 70% of the rural households depend on agriculture. Soil quality measurement is essential to enhance the productivity. This project aims to design a system which is capable of tracking the soil resource level and monitoring pH values with quality parameters. We aim to predict the soil fertility and the plantation scenario accordingly by using Machine Learning on large scale real time IoT database. Internet of Things (IoT) technology has brought revolution to each and every field of common man's life by making everything smart and intelligent. IoT refers to a network of things which make a self-configuring network. The development of Intelligent Smart Farming IoT based devices is day by day turning the face of agriculture production by not only enhancing it but also making it cost-effective and reducing wastage. In the present work, Machine Learning Inspired Smart Agriculture System with Crop Prediction is developed that will assist farmers in getting Live Data (Temperature, Soil Moisture, level of Nitrogen, Phosphorous and Potassium present in the soil, Alkalinity and Acidity in the soil etc.) for efficient environment monitoring. This type of approach will enable them to increase their overall yield and quality of products. This system will be mixed with different Sensors and a Wi-Fi module that will produce live data feed using online software Localhost and PHPmyadmin. The product being proposed is tested on Live Agriculture Fields giving high accuracy over 98% in data feeds. For this purpose, it is aimed to design and develop an approach for soil quality prediction using IoT and Machine Learning approach. The IoT technology will be introduced to generate different soil parameters from different testing in order to know missing utilities present in the soil. It is also aimed to design and develop crop prediction possibility according to current soil quality and to study the effect of soil manure on soil properties such as salinity, pH, organic matter content and microbial and enzyme activity.

II. METHODOLOGY

The entire project is segmented into two parts viz. 1. Sensing of different parameters of the agricultural land, 2. Prediction of crop production from the land.

III. BLOCK DIAGRAM OF THE SYSTEM

The block diagram depicts the theme of the project where the sensing parameters from different sensors are fed to Node MCU Smart sensing starts from the firmware parts of the project such as network access point, gateway, rules engine and API backbone. The data will be then used for machine learning based prediction technology.

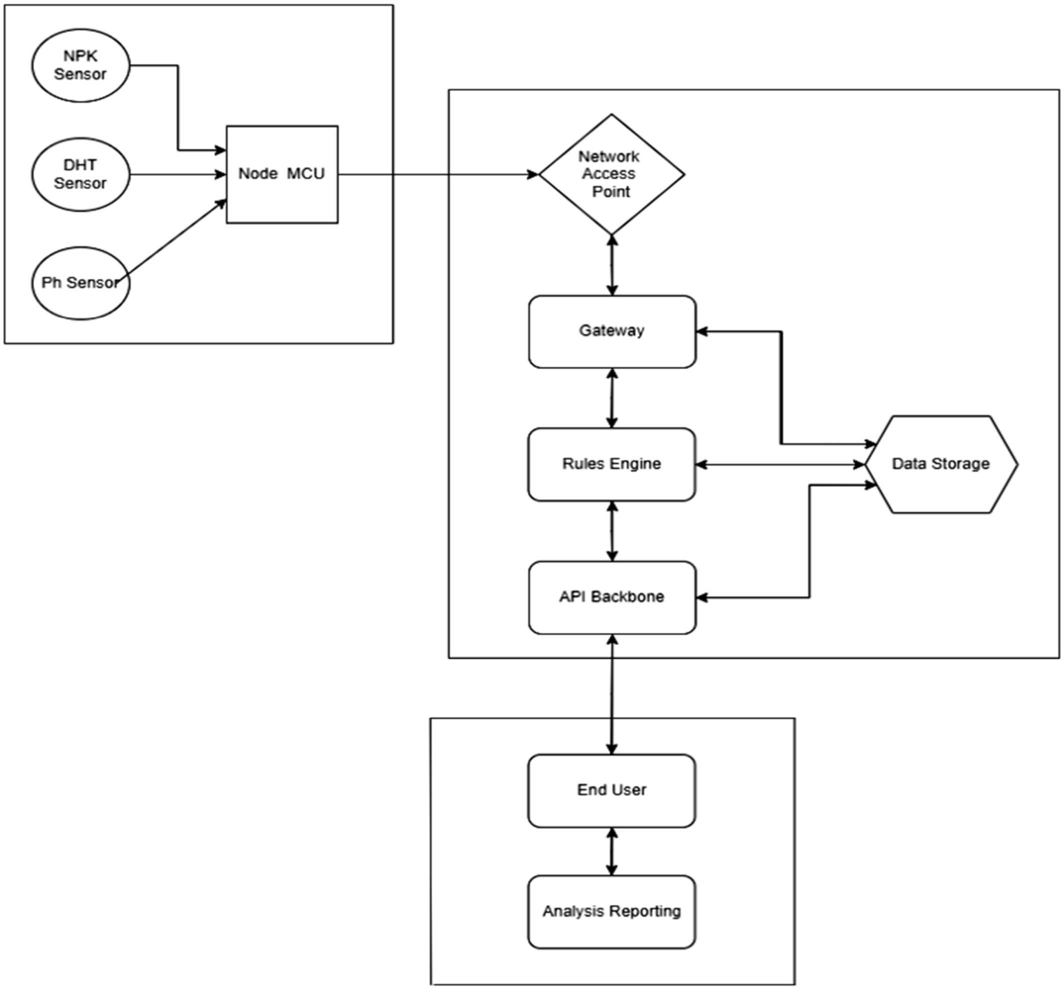


Figure 1: Block Diagram of the system

IV. COMPONENTS

A. Hardware Components

1) *Sensors:* A sensor is a device that detects and responds to some type of input from the physical environment. The specific input could be light, heat, motion, moisture, pressure, or any one of a great number of other environmental phenomena. We are using the Ph, DHT and NPK Sensors.

a) *pH Sensor:* A pH sensor is one of the most essential tools that typically used for water measurements. This type of sensor is able to measure the amount of alkalinity and acidity in water and in other solution.

Figure 2. pH sensor

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue I Jan 2023- Available at www.ijraset.com

b) *DHT Sensor:* DHT is a temperature and humidity sensor. The DHT sensors are made of 2 parts, a capacitive humidity sensor and a thermistor. There is also a very basic chip inside that does some analog to digital conversion and spits out a digital signal with the temperature and humidity. DHT11 temperature range is from $0℃$ to $50℃$ with $\pm 2℃$ accuracy. This type of humidity sensor has the humidity range from 20 to 80% with 5% accuracy. DHT22 is the more expensive version which has better specifications such as the temperature measuring range is from $-40℃$ to $+125℃$ with $\pm 0.5℃$ accuracy. This type of sensor has better humidity measuring range, from 0 to 100% with 2-5% accuracy

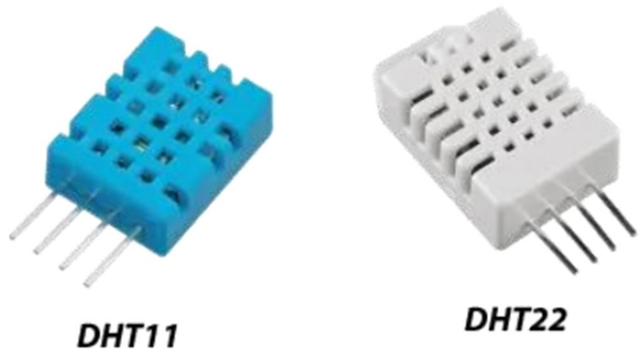


Figure 3. DHT Sensor

c) *NPK Sensor:* The soil NPK sensor is suitable for detecting the content of Nitrogen, Phosphorous and Potassium in the soil. It helps in determining the fertility of the soil thereby facilitating the systematic assessment of the soil condition.

Figure 4. NPK Sensor

2) *Node MCU:* NodeMCU is an open-source platform based on ESP8266 which can connect objects and let data transfer using the Wi-Fi protocol. In addition, by providing some of the most important features of microcontrollers such as GPIO, PWM, ADC, and etc, it can solve many of the project's needs alone. We use it as the microcontroller of our project.

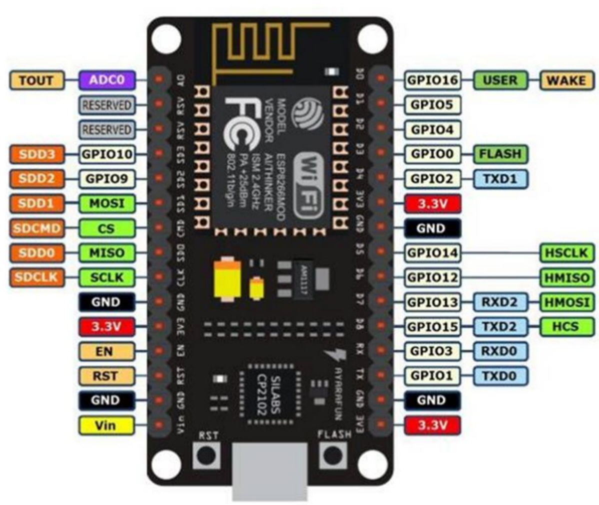


Figure 5. Node MCU

B. *Firmware Components*

1) *Network Access Point:* An access point is a device that creates a wireless local area network, or WLAN, usually in an office or large building. An access point connects to a wired router, switch, or hub via an Ethernet cable, and projects a Wi-Fi signal to a designated area. It helps to connect our Node MCU to Internet.
2) *Gateway:* A gateway is a network node used in telecommunications that connects two networks with different transmission protocols together. Gateways serve as an entry and exit point for a network as all data must pass through or communicate with the gateway prior to being routed.
3) *Rules Engine:* Rules engines are used to execute discrete logic that needs to have 100% precision. Machine learning on the other hand, is focused on taking a number of inputs and trying to predict an outcome. It's important to understand the strengths of both technologies so you can identify the right solution for the problem.
4) *Application Programming Interface (API):* An API is a software intermediary that allows two applications to talk to each other. In other words, an API is the messenger that delivers your request to the provider that you're requesting it from and then delivers the response back to you. End User and report analysis: It is essentially "the path to connected understanding" when interpreting data. It is that human element leveraging an intelligent technology component when attempting to make fact-based decisions.

C. *Software*

1) *Python:* Python is a computer programming language often used to build websites and software, automate tasks, and conduct data analysis. Python is a general-purpose language, meaning it can be used to create a variety of different programs and isn't specialized for any specific problems.

V. CIRCUIT DIAGRAM

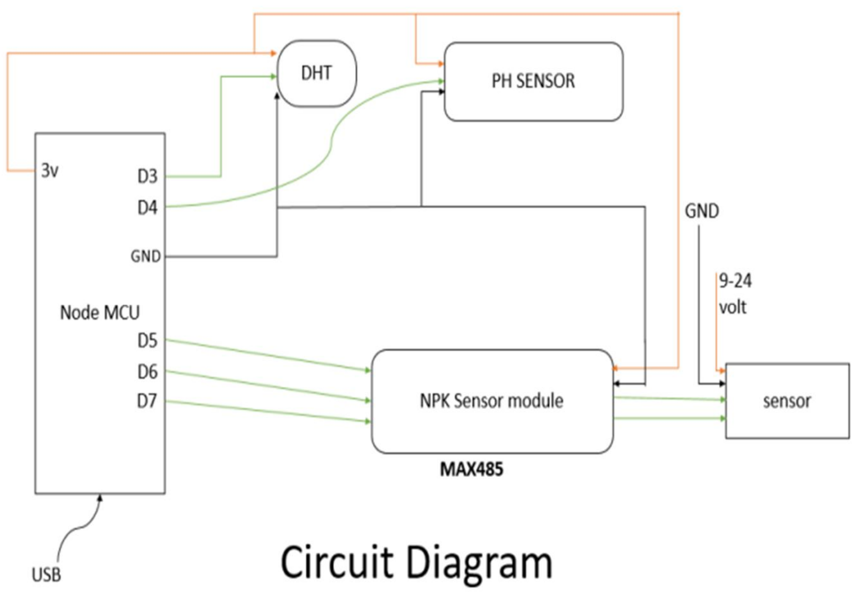


Figure 6. Circuit diagram of Smart agriculture system.

Firstly, a USB is added to Node MCU, and the dc power supply of 3V is connected to the circuit. Three different sensors are employed for smart monitoring the parameters of irrigation system and therefore they are connected to Node MCU. Input DC power is distributed to all of the three sensors with common point (GND). The D3 (GPIO0) port of Node MCU is connected to DHT sensor, D4 (GPIO2) port is attached to pH sensor. The other ports such as D5, D6, D7 are connected to NPK sensor module (Nitrogen, Phosphorous, Potassium). NPK sensor has 2 parts, i.e- NPK sensor (to be buried under soil) and 3 port NPK sensor module. 9-24 volts power is supplied to the underground NPK sensor. (Fig. 2)

VI. PREDICTION METHODOLOGY

In order to predict the crop production of the agriculture system, machine learning algorithm is employed. Machine learning is defined as a subset of artificial intelligence that is mainly concerned with the development of algorithms which allow a computer to learn from the data and past experiences on their own. There are three types of machine learning algorithm such as 1. Supervised learning 2. Unsupervised learning and 3. Reinforcement learning. In our present work, supervised learning is implemented. In supervised learning, machines are trained using well "labelled" training data, and on basis of that data, machines predict the output. The labelled data means some input data is already tagged with the correct output. In supervised learning, the training data provided to the machines work as the supervisor that teaches the machines to predict the output correctly. It applies the same concept as a student learns in the supervision of the teacher. Supervised learning is a process of providing input data as well as correct output data to the machine learning model. The aim of a supervised learning algorithm is to find a mapping function to map the input variable(x) with the output variable(y).

In this learning system, models are trained using labelled dataset, where the model learns about each type of data. Once the training process is completed, the model is tested on the basis of test data (a subset of the training set), and then it predicts the output. (Figure 3)

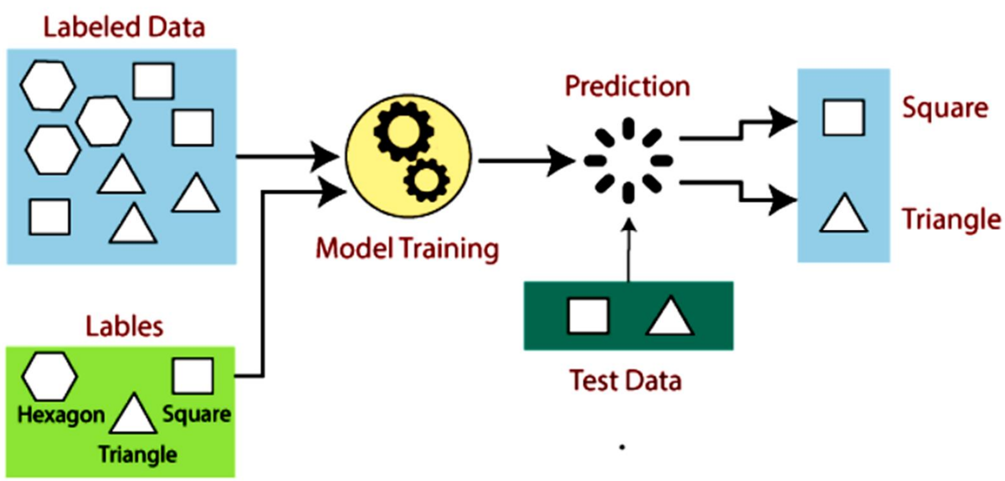


Figure 7. Schematic presentation of Supervised Learning

A. Crop Recommendation

Datasets Used: For Crop Recommendation module the dataset used are Crop recommender.csv, soil.csv, scientific_names.csv. All these datasets were obtained from Kaggle website. The Crop recommendation was used for training model since it contains attributes such as temperature, humidity, average rainfall, soil pH, nitrogen requirement ratio, potassium requirement ratio and phosphorous requirement ratio essential for predicting a crop. The datasets such as Soil names and Crop names are used after prediction to obtain the soil type and scientific name of the predicted crops.

Steps involved in Crop Recommendation module are as follows

1) *Step 1:* Importing Libraries and Dataset In order to utilize Machine Learning algorithms and preprocessing tools specific libraries needs to be imported. Using these libraries, the model building and prediction would be performed efficiently. The libraries such as NumPy, pandas, pickle, matplotlib, seaborn, Label Encoder, train_test_split were imported. The models such as Naïve bayes, Logistic Regression, SVM, Decision Tree Classifier, Bagging Classifier, Random Forest Classifier, AdaBoost Classifier, Gradient Boosting Classifier, XGBoost Classifier, LGBM Classifier and KNN was imported. The dataset called crop recommendation was used initially for training and testing the models. Each crop has a set of values for temperature, humidity, rainfall, Nitrogen, potassium, phosphorous.
2) *Step 2:* Descriptive Analysis: To obtain a best predictive model descriptive analytics is to be performed in prior. Descriptive analytics provides an idea of how the dataset looks like and helps to draw new insights. Once the dataset is imported, missing values per attribute is checked.

	N	P	K	temperature	humidity	ph	rainfall	label
0	90	42	43	20.879744	82.002744	6.502985	202.935536	rice
1	85	58	41	21.770462	80.319644	7.038096	226.655537	rice
2	60	55	44	23.004459	82.320763	7.840207	263.964248	rice
3	74	35	40	26.491096	80.158363	6.980401	242.864034	rice
4	78	42	42	20.130175	81.604873	7.628473	262.717340	rice
...	...	...	...	...	...	...	...	...
2195	107	34	32	26.774637	66.413269	6.780064	177.774507	coffee
2196	99	15	27	27.417112	56.636362	6.086922	127.924610	coffee
2197	118	33	30	24.131797	67.225123	6.362608	173.322839	coffee
2198	117	32	34	26.272418	52.127394	6.758793	127.175293	coffee
2199	104	18	30	23.603016	60.396475	6.779833	140.937041	coffee

Figure 8. Crop recommendation dataset sample values

For crop recommendation dataset the attributes are free of missing values. Once identifying that there are no missing values, the data type of the attributes is identified followed by listing the unique values in the dependent variable, i.e., Label attribute.

3) Step 3: Data Visualization

```
missing values
N              0
P              0
K              0
temperature    0
humidity       0
ph             0
rainfall       0
label          0
dtype: int64
```

(a)

```
dtypes
N                int64
P                int64
K                int64
temperature    float64
humidity       float64
ph             float64
rainfall       float64
label           object
dtype: object
```

(b)

```
unique crops
['rice' 'maize' 'chickpea' 'kidneybeans' 'pigeonpeas' 'mothbeans'
 'mungbean' 'blackgram' 'lentil' 'pomegranate' 'banana' 'mango' 'grapes'
 'watermelon' 'muskmelon' 'apple' 'orange' 'papaya' 'coconut' 'cotton'
 'jute' 'coffee']
```

(c)

```
rice           100
maize          100
jute           100
cotton         100
coconut        100
papaya         100
orange         100
apple          100
muskmelon      100
watermelon     100
grapes         100
mango          100
banana         100
pomegranate    100
lentil         100
blackgram      100
mungbean       100
mothbeans      100
pigeonpeas     100
kidneybeans    100
chickpea       100
coffee         100
Name: label, dtype: int64
```

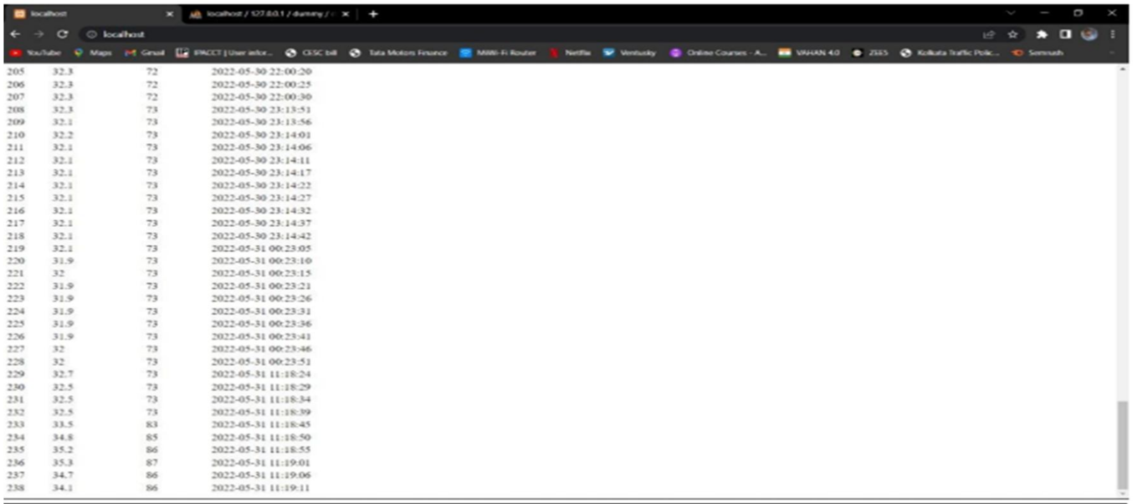
(d)

Figure 9. (a) Missing value details, (b) Datatypes of each column, (c) Crops present in the dataset, (d) Count of each Crop in the dataset.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue I Jan 2023- Available at www.ijraset.com

VII. RESULT AND DISCUSSION

Using original dataset, the PC was connected to a server using localhost, hence the Records are generated after connecting to the Node MCU USB port.



205	32.3	72	2022-05-30 22:00:20
206	32.3	72	2022-05-30 22:00:25
207	32.3	72	2022-05-30 22:00:30
208	32.3	73	2022-05-30 23:13:51
209	32.1	73	2022-05-30 23:13:56
210	32.2	73	2022-05-30 23:14:01
211	32.1	73	2022-05-30 23:14:06
212	32.1	73	2022-05-30 23:14:11
213	32.1	73	2022-05-30 23:14:17
214	32.1	73	2022-05-30 23:14:22
215	32.1	73	2022-05-30 23:14:27
216	32.1	73	2022-05-30 23:14:32
217	32.1	73	2022-05-30 23:14:37
218	32.1	73	2022-05-30 23:14:42
219	32.1	73	2022-05-31 00:23:05
220	31.9	73	2022-05-31 00:23:10
221	32	73	2022-05-31 00:23:15
222	31.9	73	2022-05-31 00:23:21
223	31.9	73	2022-05-31 00:23:26
224	31.9	73	2022-05-31 00:23:31
225	31.9	73	2022-05-31 00:23:36
226	31.9	73	2022-05-31 00:23:41
227	32	73	2022-05-31 00:23:46
228	32	73	2022-05-31 00:23:51
229	32.7	73	2022-05-31 11:18:24
230	32.5	73	2022-05-31 11:18:29
231	32.5	73	2022-05-31 11:18:34
232	32.5	73	2022-05-31 11:18:39
233	33.5	83	2022-05-31 11:18:45
234	34.8	85	2022-05-31 11:18:50
235	35.2	86	2022-05-31 11:18:55
236	35.3	87	2022-05-31 11:19:01
237	34.7	86	2022-05-31 11:19:06
238	34.1	86	2022-05-31 11:19:11

After generating the data in localhost, all those data were stored in the online MySQL database management software. This logistic regression based machine learning algorithm provides better accuracy and high efficiency in fetching the live data of Temperature, Soil Moisture, Alkalinity and Acidity present in the soil, Level of Nitrogen, Phosphorous and Potassium present in the soil and pH. Thus it will assist farmers in increasing the agriculture yield and take efficient care of food production as the system will always provide helping hand to farmers for getting accurate live feed of environmental temperature and soil moisture with more than 99% accurate results.

VIII. SCOPE OF THE PROJECT

This model will serve the purpose of agriculture and help the farmers to be independent.
By using this setup farmers will be able to do the direct business without the help of third party.
The present work will assist farmers in increasing the agriculture yield and take efficient care of food production.

IX. CONCLUSION

Thus the present work will give better accuracy in crop prediction depending on the soil parameters, humidity etc with high accuracy. Hence the work will definitely assist farmers in increasing agricultural production and provide food safety. Additionally farmers will be able to do the direct business without intervention of third party.

REFERENCES

[1] A. Goap, D. Sharma, A.K. Shukla, C. Rama Krishna, “An IoT based smart irrigation management system using Machine learning and open source technologies”, Computers and Electronics in Agriculture, vol. 155, pp. 41-49, December 2018.

[2] G. Singh; D. Sharma; A. Goap; S. Sehgal; A K Shukla; S. Kumar, “Machine Learning based soil moisture prediction for Internet of Things based Smart Irrigation System”, 2019 5th International Conference on Signal Processing, Computing and Control (ISPCC), DOI: https://doi.org/10.1109/ISPCC48220.2019.8988313, 2020

[3] Prakash Kanade, Jai Prakash Prasad, “Arduino Based Machine Learning and IoT Smart Irrigation System”, International Journal of Soft Computing and Engineering (IJSCE), vol. 10, pp. 1-5, 2021.

[4] R.Togneri, D. Felipe dos Santos, G. Camponogara, H. Nagano, G. Custódio, R. Prati, S. Fernandes, C. Kamienski, “Soil moisture forecast for smart irrigation: The primetime for machine learning”, vol. 207, pp. 117653, November 2022

[5] R. Togneri, C. Kamienski, R. Dantas, R. Prati, A. Toscano, J.-P. Soininen,T. Salmon, “Advancing IoT-Based Smart Irrigation”, IEEE Internet of Things Magazine, DOI: https://doi.org/10.1109/IOTM.0001.1900046, 2019.

[6] YounessTace, MohamedTabaa, SanaaElfilaliC, herkaouiLeghris, Hassna Bensag, EricRenault, “Smart irrigation system based on IoT and machine learning”, Energy Reports, vol. 8, pp. 1025-1036, November 2022.

[7] Anneketh Vij, Vijendra Singh, Abhishek Jain, Shivam Bajaj, Aashima Bassi, Aarushi Sharma, “IoT and Machine Learning Approaches for Automation of Farm Irrigation System”, Procedia Computer Science, vol. 167, pp.1250-1257, 2020.