IOT BASED SOIL ANALYSIS AND CROP RECOMMENDATION SYSTEM USING MACHINE LEARNING

A PROJECT REPORT

Submitted by

MANISHMA S (113218205018)

RASHMI P (113218205035)

in the partial fulfillment of the award of the degree of

BACHELOR OF TECHNOLOGY IN

INFORMATION TECHNOLOGY

DEPARTMENT OF INFORMATION TECHNOLOGY

VELAMMAL ENGINEERING COLLEGE

(An Autonomous Institution)

CHENNAI

ANNA UNIVERSITY, CHENNAI –600 025

JUNE 2022

VELAMMAL ENGINEERING COLLEGE

DEPARTMENT OF INFORMATION TECHNOLOGY

BONAFIDE CERTIFICATE

Certified that this project report titled "**IOT BASED SOIL ANALYSIS AND CROP RECOMMENDATION SYSTEM USING MACHINE LEARNING**" is the bonafide work of **MANISHMA S (113218205018)** and **RASHMI P (113218205035)** who carried out the project work under my supervision. Certified further, that to the best of my knowledge the work reported herein does not form part of any other project report or dissertation on the basis of which a degree or award was conferred on an earlier occasion on this or any other candidate.

SIGNATURE	**SIGNATURE**
SUPERVISOR	**HEAD OF THE DEPARTMENT**
Mr. MOHAN RAJ	**Dr. JEEVAA KATIRAVAN**
Assistant Professor III	Professor
Department of Information Technology	Department of Information Technology
Velammal Engineering College	Velammal Engineering College
Chennai.	Chennai.

Certified that the candidate was examined in the Viva-Voce Examination held on

____

INTERNAL EXAMINER **EXTERNAL EXAMINER**

ACKNOWLEDGEMENT

Behind every achievement lies an unfathomable sea of gratitude to those who achieved it, without whom it would ever have come into existence. To them we express our words of gratitude.

We give all the glory and thanks to our almighty **GOD** for showering upon, the necessary wisdom and grace for accomplishing this project. We express our gratitude and thanks to our **Parents** first for giving health and sound mind for completing this project.

First of all we would like to express our deep gratitude to our beloved and respectable **Chairman Thiru M.V.Muthuramalingam** and our **Chief Executive Officer Thiru M.V.M.Velmurugan** for their kind encouragement.

We express our deep gratitude to **Dr.S. Satish kumar, Principal**, Velammal Engineering College for his helpful attitude towards this project. We wish to express our sincere thanks and gratitude to **Dr.Jeevaa Katiravan, Professor and Head of the Department**, Department of Information Technology for motivating and encouraging in every part of our project.

We express our sincere gratitude to the **Project Coordinators Mrs.G.Prethija, Assistant Professor, Mrs.V.Subbulakshmi, Assistant Professor-II** Department of Information Technology and the **Project Guide Mr. Mohan Raj Assistant Professor**, Department of Information Technology for their invaluable guidance in shaping of this project without them this project would not have been possible.

Our thanks to all other **Faculty and Non-Teaching staff members** of our department for their support and peers for having stood by me and helped me to complete this project.

ABSTRACT

India is a country that has Agriculture as the backbone of the economy. More than 58% of the population has Agriculture as the primary source of income. There are almost 100-150 million farmers in the mother nation. Seeking the dark part of this sector is that many face losses due to climatic and soil conditions. Many crops die before harvest due to scarcity of water and soil nutrients. Here is a project that guides the farmer on what crop to cultivate based on the forecast report of climate and the present soil conditions. The climatic condition is identified, moisture content present in the soil is detected and Ph level of water is also considered for crop recommendation. The sensed data from the sensors are fed into the machine learning algorithm which suggests a suitable crop for cultivation. The soil parameters are led into various machine learning for crop recommendation and the best model with the highest accuracy is considered for the result. Considering this, random forest algorithm is the most accurate model for crop recommendation. The developed project also predicts the crop yield that can be produced under given circumstances. The field's harvest strength can be either poor, good or excellent. Based on the same parameters ie, soil moisture, ph of water and temperature, the crop yield is estimated. The parameters are led into random forest algorithm and the data is processed to give the desired output. The algorithm of crop yield prediction undergoes the best feature extraction technique using some mathematical calculation in order to give a great accuracy of output. Both crop recommendation system and crop yield prediction system work together to give the farmer a clear picture of farming ahead. The hardware module takes the input from the soil whereas the software module process it to give the recommendation of crop and prediction of the crop yield. The developed project takes the uncertainty of weather and soil conditions and gives the farmer a clear picture of the crop field and the climate.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

IoT	Internet of Things
Py	Python
KNN	K-Nearest Neighbor
Np	Numpy
Pd	Pandas
Df	Dataframe
SVM	Support Vector machine
IDE	Integrated development environment
LCD	Liquid Crystal Display
ML	Machine Learning
A.P	Andhra Pradesh
U.P	Uttar Pradesh
M.P	Madhya Pradesh

CHAPTER 1

INTRODUCTION

1.1 OVERVIEW

Agriculture is the backbone of every economy. In a country like India, which has ever increasing demand of food due to rising population, advances in agriculture sector are required to meet the needs. From ancient period, agriculture is considered as the main and the foremost culture practiced in India. Ancient people cultivate the crops in their own land and so they have been accommodated to their needs. Therefore, the natural crops are cultivated and have been used by many creatures such as human beings, animals and birds. The greenish goods produced in the land which have been taken by the creature leads to a healthy and welfare life. Since the invention of new innovative technologies and techniques the agriculture field is slowly degrading. Due to these, abundant invention people are been concentrated on cultivating artificial products that is hybrid products where there leads to an unhealthy life. Nowadays, modern people don't have awareness about the cultivation of the crops in a right time and at a right place. Because of these cultivating techniques the seasonal climatic conditions are also being changed against the fundamental assets like soil, water and air which lead to in security of food. Many crops die before harvest due to scarcity of water and soil nutrients. The innovation developed guides the farmer on what crop to cultivate based on the forecast report of climate and the present soil conditions. The climatic condition is identified, moisture content present in the soil is detected and Ph level of water is also considered for crop recommendation. The sensed data from the sensors are fed into the machine learning algorithm which suggests a suitable crop for cultivation. We also predict the crop yield that can be produced under given circumstances. The system developed takes the uncertainty of weather and soil conditions and gives the farmer a clear picture of the crop field and the climate.

Today, India ranks second worldwide in farm output. India is an agrarian country and more than60% of population depends on agriculture for their livelihood. While residing in urban areas may be we would not realize much importance of agriculture but this fact is not new that agriculture is the main source of income for major part of our country's population. India's population is growing faster than its ability to produce rice and wheat. The required level of investment for the development of marketing, storage and cold storage infrastructure is estimated to be huge. The country produces innumerable crops ranging from medicinal to cereal crops. These commodities are used for various purposes from human consumption, in industries, for animal feed etc. Most of us do not know major categories of crops. We are unaware of its valuable contribution to our economy. Ignoring the negative approach of the citizens living in urban areas, the cropping activities are continue to go on all the year-round in India, provided water is available for crops.

Crops in India are broadly divided into three major categories, viz.

- **Kharif crops:** The crops that are sown in the rainy season are called kharif crops. Its season starts from July and ends in October, e.g. Maize, Sugarcane, Cotton, Jawar, Bajra soyabean, turmeric, paddy, moong, ground nuts , red chillies.

- **Rabi crops:** The crops that are sown in the winter season are called rabi crops. Its season is during October to March, e.g Wheat, oil seed, pulses, rubber beans.

- **Zaid crops:** Crops grown between March and June are known as Zaid e.g. muskmelon, watermelon and vegetables like guard, pumpkin etc.

In a short note let us know our main crops:

The following are some of the **Food Crops**

- **Rice:** Rice is the main grain crop of India. India ranks second in the world in production of rice. About 34% of the total cultivated area of the nation is under rice cultivation. Rice is cultivated in areas having annual average rainfall of 125 cm. Major rice cultivating areas are north east India, eastern and western coastal regions. West Bengal, Punjab and Uttar Pradesh are the major rice producing states.

- **Wheat:** Wheat is the second major crop in India. Wheat is cultivated in areas with mean annual rainfall of 75 cm and fertile soil. Wheat has got an important role in 'Green Revolution'. The highest quantity of wheat in the country is in Uttar Pradesh. 35 % of wheat is produced only in Uttar Pradesh. This is produced by Punjab and Haryana where production of wheat is on a large scale.

- **Maize:** Maize is an important khaki crop of rainy season. Maize is cultivated in different areas and in different climates but it is suitable where temperature is 35° Celsius and rainfall is 75 centimetres. It is cultivated in hilly areas-of Jammu and Kashmir and Himachal Pradesh. Maize is cultivated throughout our country but it is cultivated more in Punjab, U.P., Bihar, M.P. and Rajasthan.

- **Pluses:** Pulses are grown in dry climate region in India. These crops provide nitrogen to the soil. Pulses are a source of proteins in the diet. Madhya Pradesh is the leading pulses producing state in India. It is also produce in Uttar Pradesh and Rajasthan.

- **Jowar:** This crop is grown where the climate is hot and dry. It is cultivated in Maharashtra, Karnataka, Andhra Pradesh, and Tamil Nadu.

In agriculture there are few crops which are grown for profit are called as **'Cash Crops or Commercial crops'.** Cash crop is a backbone of agriculture economy of India. It sets a strong base for Indian economy where country's trade and commerce flourish domestically and internationally. Cash crops are generally grown for money. In earlier days, cash crops were grown in a very small scale but today it forms a major contribution to our nation’s economy. Now it has grown at large scale for commercial purpose.

The following are some of **the Cash Crops**

- **Sugarcane:** Sugarcane is an important cash crop of India. Molasses, sugar and khandasari etc. are produced from the juice of sugarcane. Sugarcane cultivation needs temperature of 15° to 40° and rainfall of 100 to 150 centimetres and fertile loamy soil or hard soil. Sugarcane is cultivated from Kanyakumuri (southern part) to Punjab (north-west) but it is more cultivated in Uttar Pradesh.

- **Coffee:** There is a great demand of coffee in the world market. For reason, India exports coffee. Coffee cultivation needs hot and wet climate and fertile sloppy land. It is mainly produced in south Indian states of Karnataka, Kerala and Tamil Nadu. Coffee is produced on a large scale on the mountain ranges of Nilgiri.

- **Tea:** India is first in the cultivation of tea in the world. Tea cultivation needs hot climate, excess rainfall and sloppy soil. Tea is found more in Assam, Karnataka, Kerala and Himachal Pradesh, Dehradun, Ranchi of Bihar and Tripura.

- **Rubber:** Rubber is needed by different industries and transport industry in this modern age. It is cultivated in hot and wet climate garden in natural way. It is cultivated in the State of Kerala in India. Except Kerala, Andaman Nicobar Islands, Kurgan of Karnataka State and Chicmagalur district etc.

- **Oilseeds:** Groundnut, mustard, rapeseed, linseed and caster help us to get our edible oil. Oil is also extracted from coconut. India occupies the first position in the world in the production of groundnut. Groundnut cultivation needs temperature varying from 20° to 30° degree Celsius and needs 60 to 80 centimetres of rainfall. Groundnut cultivation needs sandy and light soil. Groundnuts are produced more in Tamil Nadu, Maharashtra, Gujarat, Andhra Pradesh and Uttar Pradesh.

The following are some of the **Fibre crops**

- **Cotton:** Cultivation of cotton needs 20° to 25° Celsius temperature and 50 to 75 centimetres of rainfall. Cotton plant needs wet climate at the time of growing and dry climate at the time of collecting seeds. It is known as Kharif crop. Cotton cultivation takes place in Gujarat, Maharashtra, Punjab and Haryana of our country. Except these States, cotton cultivation also takes place in Karnataka, Tamilnadu, M.P, Rajasthan, A.P. and U.P.

- **Jute:** India is a major producer of jute in the world. It is another type of fibre crop. Bags, ropes and a lot of other things are made out of jute. Jute cultivation needs hot and wet climate and fertile loamy land. Temperature of 24° to 35° Celsius and 90 to 150 centimetres of rainfall are suitable for jute cultivation. In our country, jute is cultivated in Orissa, West Bengal, eastern U.P., Bihar, Assam and Tripura.

Each crop will have some specification of soil nutrient and temperature in order to grow healthy. Considering this, the system developed analyses the soil and based on the climatic condition a suitable crop is recommended.

1.2 Motivation

The main aim of the developed system is to recommend the suitable crop and predict the yield predication of the field, in the different type of soil and weather condition. The system aims to recommend the most accurate crop based on weather, moisture content present in the soil and ph level of the water led for irrigation. The system the most accurate machine learning model to implement the innovation proposed.

CHAPTER 2

LITERATURE SURVEY

2.1 Soil Monitoring and Testing using IoT for Fertility level and Crop Prediction

Agribusiness is the foundation of India. In India 50 % of the remaining task at hand depends upon agribusiness. Commitment of agriculture part in Indian economy is higher than some other division in India. In any case, Farmers utilized customary strategy for developing harvests which comes to less profitability of yields. Additionally, a soil erosion and disintegration is likewise a principle motivation to less profitability of yields. This will impact in diminishes fruitfulness level. Loss of soil supplements through different courses is likewise motivation to diminish soil richness level. The supplements like potassium (K), nitrogen (N) and phosphorus (P) are basic for the development of a plant. The advancement in agriculture is important to tackle these issues in agribusiness part and shrewd cultivating is the appropriate response. This can be conceivable utilizing IoT gadgets. Farmers can get the necessary data just as the screen is yield. IoT associates the entire world with the assistance of sensors and other installed gadgets. The diverse soil tests will be taken from various fields and soil esteems will be determined to utilize a PH sensor where supplements worth will be separated from it. At that point, live information will be sent to the cloud database where information will be put away. At that point dependent on information or data it will be broken down lastly we show signs of improving crop yield. Here for information broke down information mining method is utilized. Information mining in agriculture assumes an imperative job in yield expectation, soil fruitfulness and plant illnesses and so on. Farmers can screen information on versatile by a website. What's more, farmers will likewise get crop list dependent on that information which harvests will be conceivable to yield in that soil.

2.2 Crop Recommendation based on Analyzing Soil using Machine Learning

The main challenge faced in agriculture sector is the lack of knowledge about the changing variations in climate. Each crop has its own suitable climatic features. This can be handled with the help of precise farming techniques. The precision farming not only maintains the productivity of crops but also increases the yield rate of production. In India it is very important to maintain sustainable agriculture to meet growing needs. Though many steps have been carried out to minimize the loss of crops, traditional methods has its own disadvantages. These disadvantages can be overcome with the help of precision farming. With the use of IOT and prediction system, precision farming makes decision. The data from the field area are collected using IOT system. The prediction framework is fed with the collected data from the sensors for obtaining suggestion. Crop selection and changing climatic conditions are the two major problems faced among the farmers. With the existing prediction and monitoring methods we can address this problem. Though these methods are useful, there is no optimum solution for the crop suggestion. Some of the drawbacks found in the existing system are the improper analysis, choosing effective algorithms, and efficient selection of attributes all these parameters may affect the crop yield. The system developed helps in overcoming the drawbacks found in the existing system. The methods in the proposed system includes increasing the yield of crops, real-time analysis of crops using IOT, selecting efficient parameters, making smarter decisions and getting better yield. Effective algorithms need to be used for early prediction of crops. By choosing correct factors most useful data can be generated using ML model by intake of various parameters and this generated data helps the farmers by suggesting a right crop to be sown. To suggest the precise crop for a particular field area is the main aim of crop suggestion model. By choosing the suitable crops for the field area we can minimize the loss of crops. Suitable algorithms with selective features has to be chosen since the accuracy level of suggesting crops may vary based on the type of algorithm chosen. Machine Learning is found be the most effective technology for predicting suitable crops and its yield.

2.3 Framework for Internet of Things for Remote Soil Monitoring

To meet the rising food demands, the technological implication in farming is emerging. Proper soil monitoring plays a vital role in agricultural production. IoT based soil monitoring system is used to maximize the yield of crops or plants by observing soil parameters and providing the necessary information to the farmers remotely. In general, the diagnosis of soil properties of the field is executed through manual laboratory testing. But this can be accomplished in real-time through IoT, where some agricultural soil sensors are engaged for remote sensing. Optimal utilization of resources is crucial in real-time sensors' data monitoring. After analyzing the acquired data, one can get useful inference about the recommendation of crops, fertilizer, and so on in the field. This research deals with several remote soil monitoring systems based upon IoT protocols developed by different researchers to increase crop yields and also provide an overview of sensors, technologies, benefits, challenges, and future aspects of systems. This research also proposes a general framework for IoT based soil monitoring system by using existing technologies.

2.4 Impact of Climate Change on Yields of Major Food Crops in India: Implications for Food Security

Suitable crop for particular land is predicted by considering parameter such as Nature of soil, Humidity, light, Soil moisture and market demand. Data from the sensor which are deployed in the field are collected via NodeMCU and sent to cloud for prediction of suitable crop for that particular land. Predicted crop list are sent to End User as SMS. End User can decide one of the crops from suitable list and how many acres he wish to cultivate and gives his input as a reply message to cloud. If the End User's selected crop supply satisfies the market demand threshold then threshold is updated in the server. else end User gets message to select alternative crop.

2.5 Sensors Driven AI-Based Agriculture Recommendation Model for Assessing Land Suitability

The rapid development of wireless sensor networks has triggered the design of low-cost and small sensor devices with the Internet of Things (IoT) empowered as a feasible tool for automating and decision-making in the domain of agriculture. This research proposes an expert system by integrating sensor networks with Artificial Intelligence systems such as neural networks and Multi-Layer Perceptron (MLP) for the assessment of agriculture land suitability. The developed system will help the farmers to assess the agriculture land for cultivation in terms of four decision classes, namely more suitable, suitable, moderately suitable, and unsuitable. This assessment is determined based on the input collected from the various sensor devices, which are used for training the system. The results obtained using MLP with four hidden layers is found to be effective for the multiclass classification system when compared to the other existing model. This trained model will be used for evaluating future assessments and classifying the land after every cultivation.

2.6 IoT-Based Soil Health Monitoring and Recommendation System

Soil is the base of agriculture. Soil provides nutrients that increase the growth of a crop. Some chemical and physical properties of soil, such as its moisture, temperature and its pH, heavily affect the yield of a crop. These properties can be sensed by the open-source hardware, and they can be used in the field. In this chapter, a soil health monitoring system is proposed in which farmer will be able to monitor soil moisture, soil temperature and soil pH in his android smart-phone. The farmer will also get the recommendations of lime and sulphur on the basis of pH of the soil. The area of study is Jaipur, Rajasthan (26.9124° N, 75.7873° E). The proposed system is tested with the results taken from authorized laboratories. The proposed system is validated with t-test having no significance difference between calibrated values and laboratory-recorded values. The proposed system is implemented on android smart-phone so it is useful for farmers, agriculture scientists, agriculture professionals and IoT experts.

2.7 Soil Nutrient measurement in paddy farming using IoT

Farmers in India still use primitive and old ways to grow their crops especially paddy which even though is not wrong but doesn't help farmers to get the maximum yield out of their paddy farms. To do that they need to utilize the modern technology that will not only help them increase their paddy yield but also reduce their dependency on weather conditions and also on manual labor. One such technology is discussed in this paper. This paper proposed a way to measure soil nutrients in real time and send that data to the farmer so that he can plan the type of seed for paddy crop he wants to grow for that season. This paper mainly measures nitrogen phosphorous potassium (NPK) in the soil which means nitrogen sodium and potassium which are major micronutrients in the soil. Farmers start using fertilizers which are not only harmful to the environment but also are damaging to the humans and animals who consume them. Not only do we provide a way to avoid this but also help the farmers which variety of seed to a farm to increase their yield. We also send a notification to the farmer's phone about the values in the soil and also the market values of the paddy crop they are farming in particular season.

2.8 Agricultural Crop Monitoring Sensors using IoT

The IOT is remodeling the agriculture enabling the farmers with the wide range of techniques. IOT technology helps in collecting information about conditions like weather, moisture, temperature and fertility of soil, Crop online monitoring enables detection of weed, level of water, pest detection, animal intrusion in to the field, crop growth, agriculture.Wireless sensor networks are used for monitoring the farm conditions and micro controllers are used to control and automate the farm processes. To view remotely the conditions in the form of image and video, wireless cameras have been used. IOT technology can reduce the cost and enhance the productivity of traditional farming.

2.9 Demand Based Crop Recommender System For Farmers

In today's world, IoT plays an important role in human life simply by allowing objects to be sensed or remotely controlled through other existing networks. Traditional farming methods with some recommendations will save time and effort. This paper planned to use technologies such as IoT, Machine Learning and Cloud computing to recommend to farmers the type of crop to be grown on the basis of available resources, climate, rainfall, temperature, market prices, area of land, past crop yield and other parameters, thereby reducing the effort and time required for different agricultural processes. In considering the above parameters, it is recommended that farmers grow and the type of pesticides and water supply be used from time to time. The system can help farmers produce suitable crops. As a result, they can improve their lifestyles and help society more. This paper present an attempt to predict the yield and price of crops that a farmer can obtain from his land by analyzing past data patterns. It gives them an overview of the recommendation to grow the plant until the crop is sold. Therefore, provide a comprehensive guide to help farmers in their agricultural work

2.10 Are regional climate models relevant for crop yield prediction in West Africa

Agriculture plays the major role in economics and survival of people in India. The purpose of this project is to provide embedded based system for soil monitoring and irrigation to reduce the manual monitoring of the field and get the information via mobile application. The system is proposed to help the farmers to increase the agricultural production. The soil is tested using various sensors such as pH sensor, temperature sensor, and humidity sensor. Based on the result, the farmers can cultivate the appropriate crop that suits the soil. The obtained sensor values are sent to the field manager through the Wi-Fi router and the crop suggestion is made through the mobile application. Automatic irrigation system is carried out when the soil temperature is high. Crop image is captured and it is sent to the field manager to suggest pesticides.

CHAPTER 3

PROPOSED SYSTEM

3.1 SYSTEM ARCHITECTURE DIAGRAM

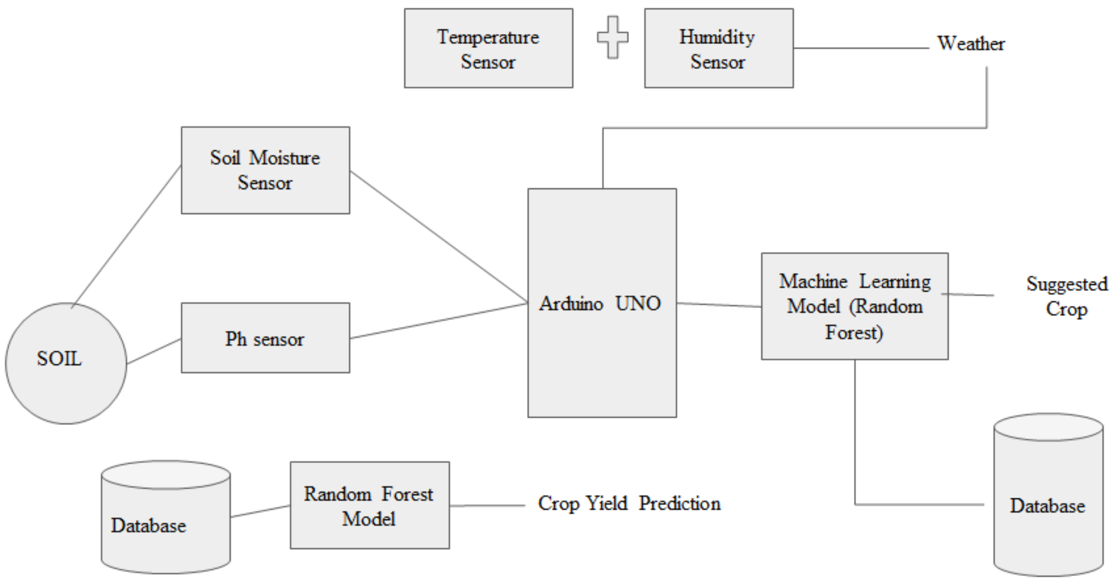

Fig 3.1 System Architecture Diagram

Fig 3.1 explains the system architecture diagram of the proposed work. The soil parameters are analyzed first to recommend a suitable crop. The soil moisture sensor senses the moisture content present in the soil. Ph sensor senses the Ph level of the water that irrigates the field. Temperature and Humidity sensor together gives the temperature. The Arduino UNO fetches the data from the sensors and displays it to the user. The sensor data (Soil Moisture, Water Ph level and temperature) is given into the machine learning model (Rain forest). The machine learning predicts the suitable crop based on the given parameters. The same sensors data are given into another machine learning model (Rain forest) for crop yield prediction by that field under given circumstances.

3.2 LIST OF MODULES

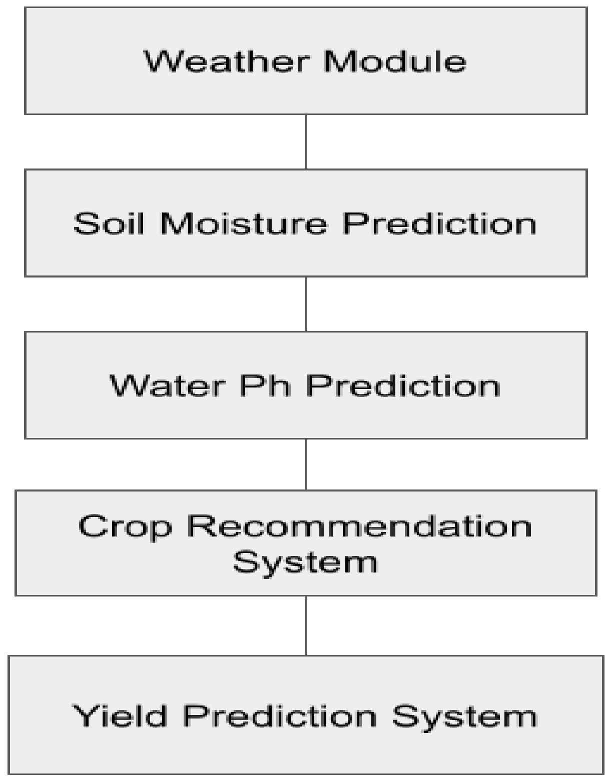

Fig 3.2 Flowchart Diagram for the list of modules

Fig 3.2 shows the flowchart diagram for the list of modules.

3.3 Module Description

3.3.1 Weather Module

Dht11 sensor is used here to monitor the humidity variation of the environment where the crops are cultivated. This is a digital sensor and measures the humidity value in percentage format. DHT11 sensor consists of a capacitive humidity sensing element and a thermistor for sensing temperature. The humidity sensing capacitor has two electrodes with a moisture holding substrate as a dielectric between them. Change in the capacitance value occurs with the change in humidity levels. The IC measure, process this changed resistance values and change them into digital form. For measuring temperature this sensor uses a Negative Temperature coefficient thermistor, which causes a decrease in its resistance value with increase in temperature. To get larger resistance value even for the smallest change in temperature, this sensor is usually made up of semiconductor ceramics or polymers.

3.3.1.1 DHT-11

This sensor is used here to monitor the humidity variation of the environment where the crops are cultivated. This is a digital sensor and measures the humidity value in percentage format.

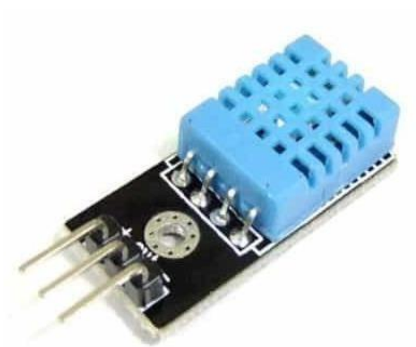

Fig 3.3 DHT11 Sensor

DHT11 humidity and temperature sensor is available as a sensor and as a module. The difference between this sensor and module is the pull-up resistor and a power-on LED. DHT11 is a relative humidity sensor. To measure the surrounding air this sensor uses a thermistor and a capacitive humidity sensor.

3.3.1.1.1 Working of DHT11

DHT11 sensor consists of a capacitive humidity sensing element and a thermistor for sensing temperature. The humidity sensing capacitor has two electrodes with a moisture holding substrate as a dielectric between them. Change in the capacitance value occurs with the change in humidity levels. The IC measure, process this changed resistance values and change them into digital form. For measuring temperature this sensor uses a Negative Temperature coefficient thermistor, which causes a decrease in its resistance value with increase in temperature. To get larger resistance value even for the smallest change in temperature, this sensor is usually made up of semiconductor ceramics or polymers.

The temperature range of DHT11 is from 0 to 50 degree Celsius with a 2- degree accuracy. Humidity range of this sensor is from 20 to 80% with 5% accuracy. The sampling rate of this sensor is 1Hz .i.e. it gives one reading for every second. DHT11 is small in size with operating voltage from 3 to 5 volts. The maximum current used while measuring is 2.5mA.

DHT11 sensor has four pins- VCC, GND, Data Pin and a not connected pin. A pull-up resistor of 5k to 10k ohms is provided for communication between sensor and micro-controller.

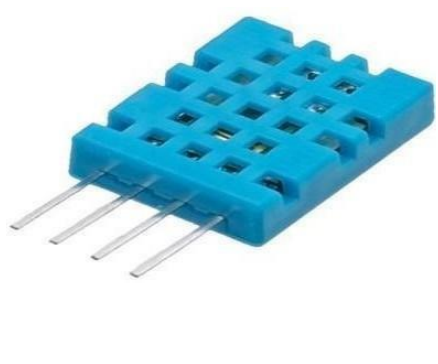

Fig 3.4 DHT11 Sensor pin diagram

3.3.1.1.2 Humidity (DHT11)

- DHT11 uses only one wire for communication. The voltage levels with certain time value defines the logic one or logic zero on this pin.
- The communication process is divided in three steps, first is to send request to DHT11 sensor then sensor will send response pulse and then it starts sending data of total 40 bits to the microcontroller.

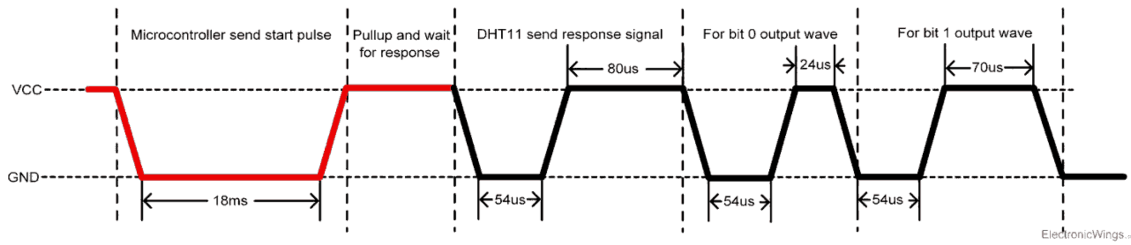

Fig 3.5 DHT11 Communication process

3.3.1.1.3 Start pulse (Request)

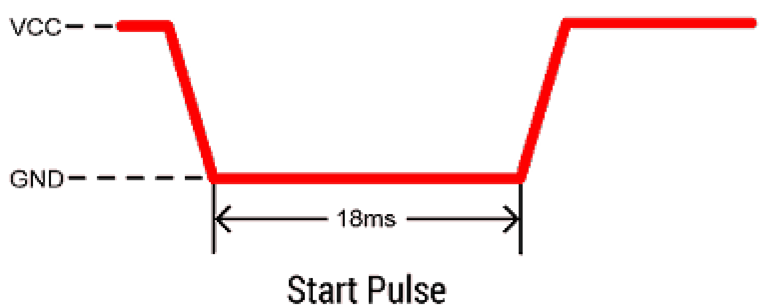

Fig 3.6 DHT11 Pulse diagram

- To start communication with DHT11, first we should send the start pulse to the DHT11 sensor.

- To provide start pulse, pull down (low) the data pin minimum 18ms and then

pull up, as shown in diag.

3.3.1.1.4 Response

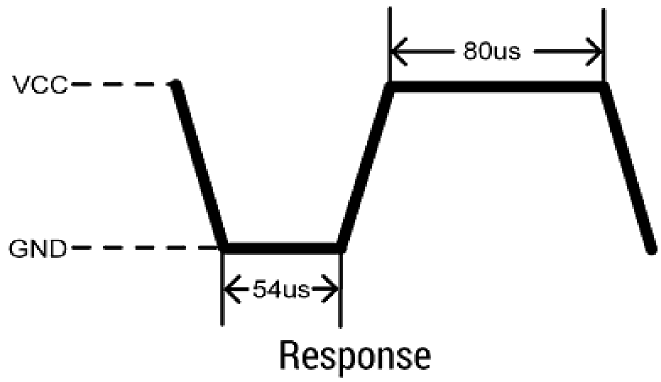

Fig 3.7 DHT11 Sensor Response diagram

- After getting start pulse from, DHT11 sensor sends the response pulse which indicates that DHT11 received start pulse.
- The response pulse is low for 54us and then goes high for 80us.

3.3.1.1.5 Data

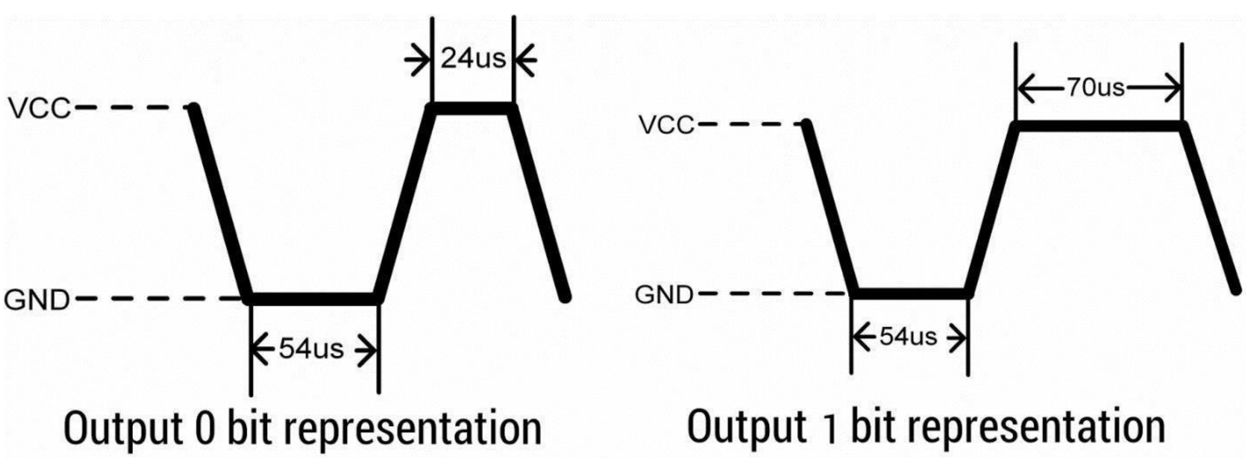

Fig 3.8 DHT11 Sensor Bit representation diagram

- After sending the response pulse, DHT11 sensor sends the data, which contains humidity and temperature value along with checksum.
- The data frame is of total 40 bits long, it contains 5 segments (byte) and each segment is 8-bit long.
- In these 5 segments, first two segments contain humidity value in decimal integer form. This value gives us Relative Percentage Humidity. 1st 8-bits are integer part and next 8 bits are fractional part.
- Next two segments contain temperature value in decimal integer form. This value gives us temperature in Celsius form.
- Last segment is the checksum which holds checksum of first four segments.
- Here checksum byte is direct addition of humidity and temperature value. And we can verify it, whether it is same as checksum value or not. If it is not equal, then there is some error in the received data.
- Once data received, DHT11 pin goes in low power consumption mode till next start pulse.

3.3.1.1.6 End of frame

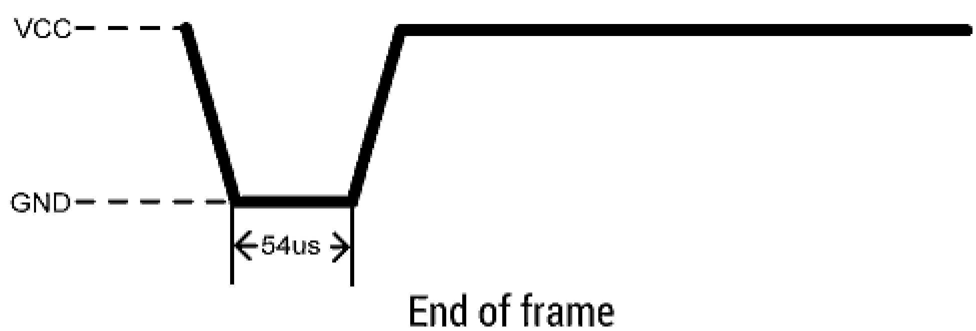

Fig 3.9 DHT11 Sensor End of frame

- After sending 40-bit data, DHT11 sensor sends 54us low level and then goes high. After this DHT11 goes in sleep mode.

3.3.2 Soil Moisture Prediction

The moisture sensor detects the wetness or dryness of the soil. Based on the dryness value of the soil the controller controls the pump. The Moisture sensor is used to measure the water content (moisture) of soil. When the soil is having water shortage, the module output is at high level, else the output is at low level. The Soil Moisture Sensor uses capacitance to measure dielectric permittivity of the surrounding medium. In soil, dielectric permittivity is a function of the water content. The sensor creates a voltage proportional to the dielectric permittivity, and therefore the water content of the soil. The sensor averages the water content over the entire length of the sensor. There is a 2 cm zone of influence with respect to the flat surface of the sensor, but it has little or no sensitivity at the extreme edges.

3.3.2.1 Moisture Sensor

The moisture sensor detects the wetness or dryness of the soil. Based on the dryness value of the soil the controller controls the pump. The Moisture sensor is used to measure the water content (moisture) of soil. When the soil is having water shortage, the module output is at high level, else the output is at low level. This sensor reminds the user to water their plants and also monitors the moisture content of soil. It has been widely used in agriculture, land irrigation and botanical gardening.

The Soil Moisture Sensor uses capacitance to measure dielectric permittivity of the surrounding medium. In soil, dielectric permittivity is a function of the water content. The sensor creates a voltage proportional to the dielectric permittivity, and therefore the water content of the soil. The sensor averages the water content over the entire length of the sensor. There is a 2 cm zone of influence with respect to the flat surface of the sensor, but it has little or no sensitivity at the extreme edges. The Soil Moisture Sensor is used to measure the loss of moisture over time due to evaporation and plant uptake, evaluate optimum soil moisture contents for various species of plants, monitor soil moisture content to control irrigation in greenhouses and enhance bottle biology experiments.

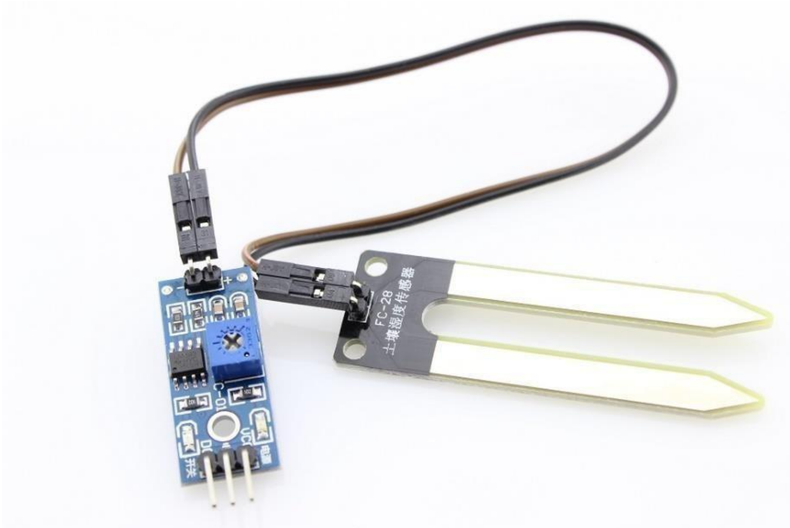

Fig 3.10 Moisture Sensor

3.3.2.1.1 Specifications

- Working Voltage:5V
- Working Current:<20mA
- Interface type: Analog
- Working Temperature:10°C~30°C

3.3.3 Water Ph Prediction

pH is the numeric representation of gram-equivalent per liter of hydrogen ion concentration in any solution. It varies between 0 to14. It is the logarithmic measurement of moles of hydrogen ions per liter of solution. The solutions having pH value between 0 to 7 are acidic solutions with large concentration of hydrogen ions whereas solutions having pH value between 8 to 14 are basic solutions with small hydrogen concentration. The solutions having pH value of 7 are neutral solutions. Measuring the pH gives the measure of alkalinity or acidity of a solution. pH meter basically works on the fact that interface of two liquids produces a electric potential which can be measured. In other words when a liquid inside an enclosure made of glass is placed inside a solution other than that liquid, there exists an electrochemical potential between the two liquids.

3.3.3.1 pH Sensor

3.3.3.1.1 Defining pH

pH is the numeric representation of gram-equivalent per liter of hydrogen ion concentration in any solution. It varies between 0 to14. It is the logarithmic measurement of moles of hydrogen ions per liter of solution. The solutions having pH value between 0 to 7 are acidic solutions with large concentration of hydrogen ions whereas solutions having pH value between 8 to 14 are basic solutions with small hydrogen concentration. The solutions having pH value of 7 are neutral solutions. Measuring the pH gives the measure of alkalinity or acidity of a solution.

3.3.3.1.2 Principle of pH Meter or pH Sensor

pH meter basically works on the fact that interface of two liquids produces a electric potential which can be measured. In other words when a liquid inside an enclosure made of glass is placed inside a solution other than that liquid, there exists an electrochemical potential between the two liquids.

3.3.3.1.3 PH measurement:

A very important measurement in many liquid chemical processes (industrial, pharmaceutical, manufacturing, food production, etc.) is that of pH: the measurement of hydrogen ion concentration in a liquid solution. A solution with a low pH value is called an "acid," while one with a high pH is called a "caustic." The common pH scale extends from 0 (strong acid) to 14 (strong caustic), with 7 in the middle representing pure water (neutral):

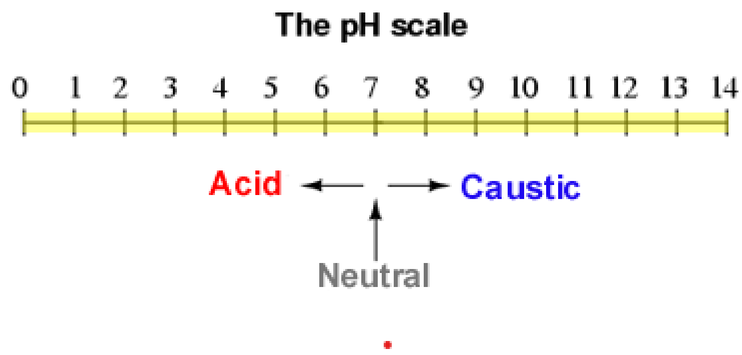

Fig 3.11 The pH scale

3.3.3.1.4 Working of pH Sensor

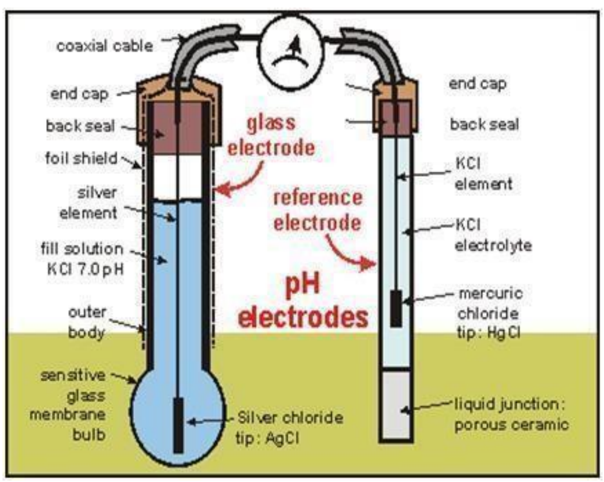

Fig 3.12 Working of pH Sensor

The electrode is placed inside the beaker filled with a solution whose pH is to be measure. The glass bulb welded at the end of the measurement electrode consists of lithium ions doped to it which makes it act as an ion selective barrier and allows the hydrogen ions from the unknown solution to migrate through the barrier and interacts with the glass, developing an electrochemical potential related to the hydrogen ion concentration. The measurement electrode potential thus changes with the hydrogen ion concentration. On the other hand, the reference electrode potential doesn't changes with the hydrogen ion concentration and provides a stable potential against which the measuring electrode is compared. It consists of a neutral solution which is allowed to exchange ions with the unknown solution through a porous separator, thus forming low resistance connection to complete the whole circuit. The potential difference between the two electrodes gives a direct measurement of the hydrogen ion concentration or pH of the system and is first preamplifier to strengthen it and then given to the voltmeter.

U = EpH – Eref

EpH – Voltage potential of measurement electrode

Eref – Voltage potential of reference electrode

The pH is calculated based on the Nerns't equation which states that change in total potential for every change in pH is

U = -kTpH

K- boltzman's constant, T- temperature.

3.3.4 Crop Recommendation System

The environmental conditions differ from region to region, a machine learning model is used to predict the best crop type for the selected land. To train the crop recommending model with the data collected from the Arduino sensors, machine learning algorithms are used to identify the best crop to cultivate with the highest probability of growing. From this model, it decided what type of crops that the farmer should grow. This is done by analyzing factors of humidity, temperature, soil moisture and pH level.

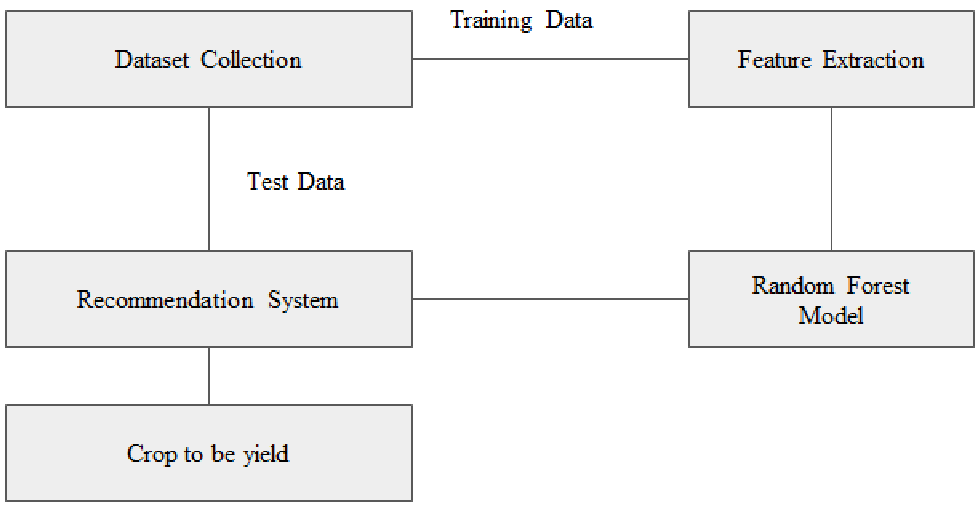

Fig 3.13 Crop Recommendation System

The dataset contains N, P, K, Temperature, Humidity, ph, Soil Moisture and the crop that can be sown.

	N	P	K	temperature	humidity	ph	Soil Moisture	label
0	94	35	35	20.879744	82.002744	6.502985	202.935536	rice
1	65	47	41	21.770462	80.319644	7.038096	226.655537	rice
2	91	48	35	23.004459	82.320763	7.840207	263.964248	rice
3	73	42	44	26.491096	80.158363	6.980401	242.864034	rice
4	63	57	36	20.130175	81.604873	7.628473	262.717340	rice

Fig 3.14 Dataset

In this dataset N, P, K, temperature, humidity, ph and Soil Moisture are considered as features. Label is considered as the target or the output value. The whole dataset is divided into train data and test data in the ratio of 0.2. The train data is fed into various machine learning models like Decision Tree, Gaussian Naïve Bayes, Support Vector

Machine, Randomforest,andK-Nearest Neighbor Algorithm. A comparative study is undergone between the accuracy of the listed algorithms.

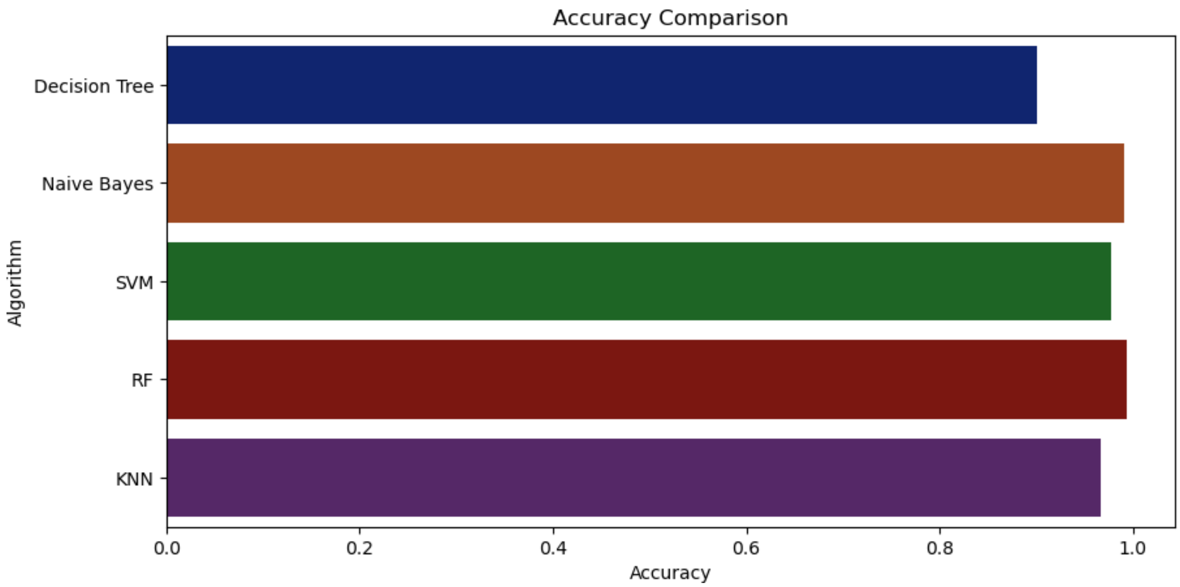

Fig 3.15 Comparing Accuracy of the various models

The Accuracy score of Random forest Model turns out to be high.

```
Decision Tree --> 0.9
Naive Bayes --> 0.990909090909091
SVM --> 0.9772727272727273
RF --> 0.9931818181818182
KNN --> 0.9659090909090909
```

Fig 3.16 Accuracy scores of the models

Thus the crop recommendation system is built using Random Forest Model with theaccuracy of 99.318%.

3.3.5 Crop Yield Prediction System

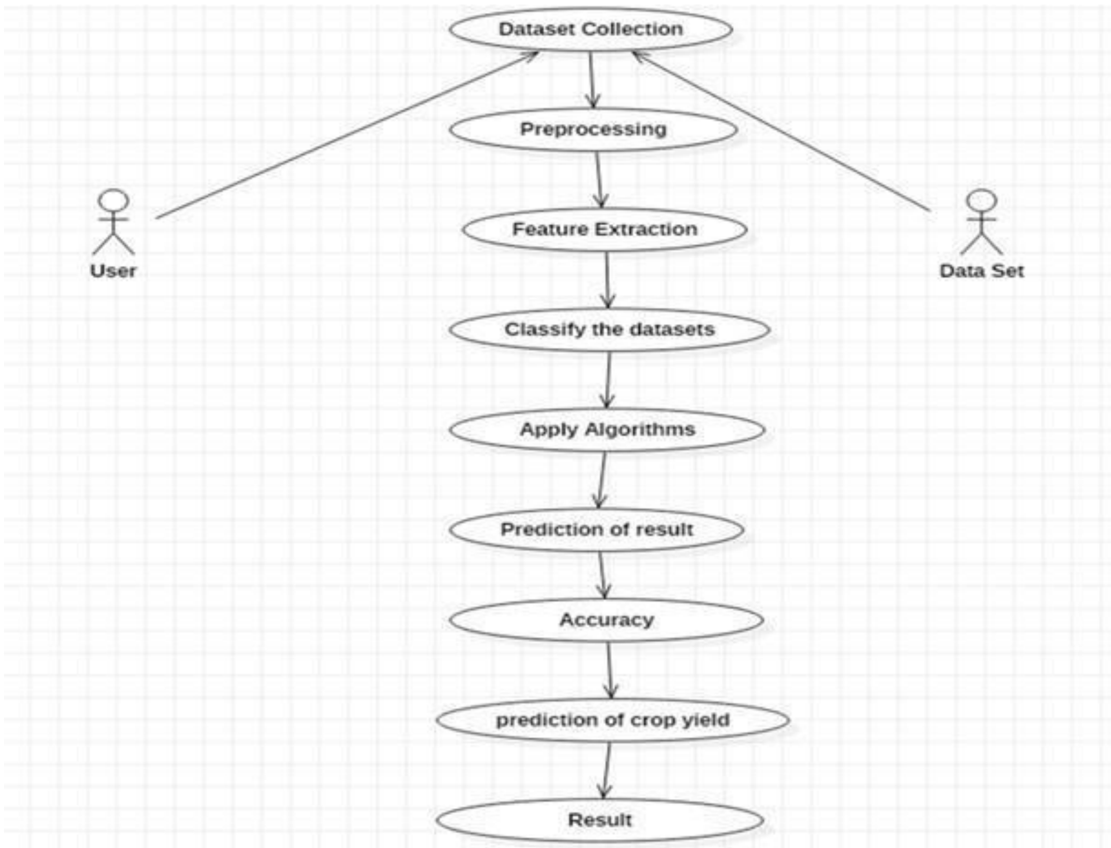

Fig 3.17 Crop Yield Prediction System

Crop yield prediction is an essential task for the decision-makers at national and regional levels (e.g., the EU level) for rapid decision-making. An accurate crop yield prediction model can help farmers to decide on what to grow and when to grow. Machine learning, a fast-growing approach that's spreading out and helping every sector in making viable decisions to create the foremost of its applications. Most devices nowadays are facilitated by models being analysed before deployment. The main concept is to increase the throughput of the agriculture sector with the Machine Learning models. Another factor that also affects the prediction is the amount of knowledge that's being given within the training period, as the number of parameters was higher comparatively. The core emphasis would be on precision agriculture, where quality is ensured over undesirable environmental factors. So as to perform accurate prediction and stand on the inconsistent trends in temperature and rainfall various machine learning classifiers like Logistic Regression, Naive Bayes, Random Forest etc. are applied to urge a pattern. By applying the above machine learning classifiers, a conclusion can be made that Random Forest algorithm provides the foremost accurate value. System predicts crop prediction from the gathering of past data. Using past information on weather, temperature and a number of other factors the information is given.

The dataset contains soil moisture, rainfall and temperature.

	Moisture	rainfall	Average Humidity	Mean Temp	max Temp	Min temp
0	12.801685	0.012360	57	62	71	52
1	12.851654	0.004172	57	58	73	43
2	12.776773	0.000000	56	58	69	46
3	12.942001	0.031747	62	56	70	43
4	12.984652	0.000000	65	56	70	42

Fig 3.18 Crop Yield Prediction Dataset

Add a column called soil with clay, alkaline, sandy and chalky as the soil type serially into the dataset. The dataset is then shuffled using the random feature of python.

	Moisture	rainfall	Average Humidity	Mean Temp	max Temp	Min temp	soil
0	12.801685	0.012360	57	62	71	52	alkaline
1	12.851654	0.004172	57	58	73	43	clay
2	12.776773	0.000000	56	58	69	46	sandy
3	12.942001	0.031747	62	56	70	43	clay
4	12.984652	0.000000	65	56	70	42	alkaline

Fig 3.19 Shuffled Dataset

The soil column is treated as dummy data and the values are changed into 0’s and 1’s. This is done because machines understand 0’s and 1’s better than alphabets. Another column called Bio is added at the end. The Bio column mentions the condition of the soil. Based on the soil type and its parameters, the soil condition is categorized as Excellent, good and poor. The Bio column is also changed into 0’s and 1’s.

Based on some mathematical calculations and comparisons the final yield is set and this value is stored in the output column final yield.

	Moisture	rainfall	Average Humidity	Mean Temp	max Temp	Min temp	soil	alkaline	sandy	chalky	clay	bio	excellent	poor	good	yield	final yield
0	12.801685	0.012360	57	62	71	52	alkaline	1	0	0	0	poor_bio_cond	0	1	0	0	0.0
1	12.851654	0.004172	57	58	73	43	clay	0	0	0	1	excellent_bio_cond	1	0	0	0	1.0
2	12.776773	0.000000	56	58	69	46	sandy	0	1	0	0	poor_bio_cond	0	1	0	0	4.0
3	12.942001	0.031747	62	56	70	43	clay	0	0	0	1	excellent_bio_cond	1	0	0	0	1.0
4	12.984652	0.000000	65	56	70	42	alkaline	1	0	0	0	poor_bio_cond	0	1	0	0	4.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
3972	12.959730	0.000000	30	72	85	58	clay	0	0	0	1	excellent_bio_cond	1	0	0	0	4.0
3973	12.985416	0.000000	30	70	87	52	sandy	0	1	0	0	good_bio_cond	0	0	1	0	4.0
3974	12.947405	0.000300	33	70	86	53	chalky	0	0	1	0	good_bio_cond	0	0	1	0	0.0
3975	12.771689	0.000000	35	69	85	53	clay	0	0	0	1	poor_bio_cond	0	1	0	0	4.0
3976	12.845779	0.010131	36	70	86	53	alkaline	1	0	0	0	good_bio_cond	0	0	1	0	0.0

3977 rows × 17 columns

Fig 3.20 Final Dataset

The final yield is considered as the target value and the other columns are considered as the input features. The final dataset is splitted into test data and train data in the ratio of 0.3. The train data is passed into Random forest algorithm. The Random forest algorithm is a classification ML model that predicts the output based on possible outcomes. It picks the most frequent outcome from the possible ones and gives it as an output to the user. The model is built with accuracy of 99.58%.

CHAPTER 4

SNAPSHOTS AND RESULT DISCUSSION

The hardware module (soil analysis module) consists of LCD Display, Arduino UNO, temperature and humidity sensor, water ph sensor, soil moisture sensor and ground. All the sensors and LCD Display is connected to Arduino using connecting wires.

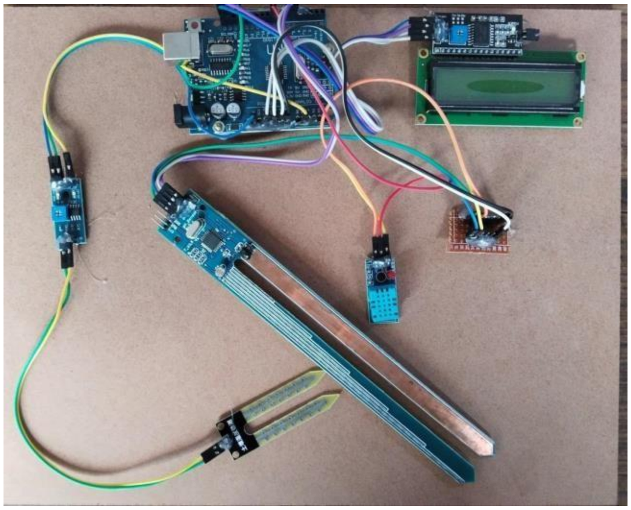

Fig 4.1 Hardware Module

The hardware module and the software module are connected through the serial port. Follow the steps below to connect software and hardware module.

- Open Arduino IDE with the module source code.
- Navigate into Tools and select ports.
- Select port in which you connected the hardware module. In this case we have connected through COM7 port.

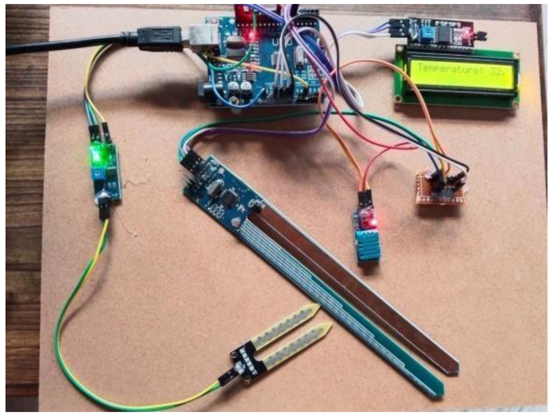

Fig 4.2 Hardware and software module connected

Immerse the soil moisture sensor in soil and ph sensor in water. The readings will be shown in the LCD display.

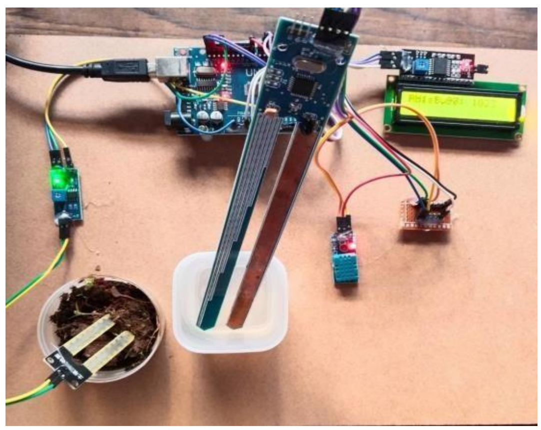

Fig 4.3 Placing sensors for capturing output

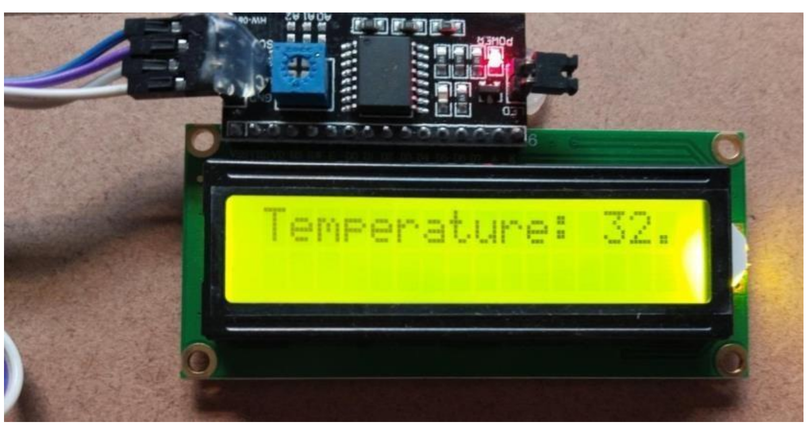

Fig 4.4 Temperature captured

Fig 4.5 Soil moisture captured

Fig 4.6 ph value captured

The output from Soil Moisture Sensor, Ph Sensor and Temperature Humidity Sensor is passed to Arduino IDE and displayed.

```
Humidity61.0
moisture: 1023
PH: 7.90
Temperature: 32.00
```

Fig 4.7 Output shown in Arduino IDE

The temperature, humidity, pH and soil moisture is given as input by the user. (The values inputted are the values captured by the sensor). The NPK value is taken as average for this project. A manual testing can be done to detect NPK using a NPK kit. The preferred one is the NPK kit because NPK sensors are costly.

```
temp=float(input("Enter the temperature:"))
hum=float(input("Enter the humidity:"))
ph=float(input("Enter the ph value:"))
mois=float(input("Enter the Soil Moisture"))
```

```
Enter the temperature:32
Enter the humidity:61
Enter the ph value:7.90
Enter the Soil Moisture102.3
```

```
#['N', 'P', 'K', 'temperature', 'humidity', 'ph', 'soil moisture']
data = np.array([[57,49,42,temp,hum,ph,mois]])
prediction = RF.predict(data)
print(prediction)
```

```
['jute']
```

Fig 4.8 Recommended crop

```
test_vector = np.reshape(np.asarray([12.737998,0.026821,61,56,70,42,1.0,0.0,0.0,0.0,0.0]),(1,11))
p = int(rf.predict(test_vector)[0])
yield_list = ['20-30%','30-40%','50%','60-70%','80-100%']
print (yield_list[p])
```

```
80-100%
```

Fig 4.9 Crop yield predicted

CHAPTER 5

CONCLUSION

The system developed recommends an accurate crop that can be cultivated using a machine learning algorithm called the Rain forest algorithm. The yield that can be produced by the field is also predicted using the Random Forest algorithm. The Rain Forest algorithm achieves the best accuracy when compared to other modules in recommending the best crop for the field. The Rain Forest algorithm achieves the best accuracy with the lowest number of decision trees built. It is suitable for massive crop yield prediction in agricultural planning. The soil parameters are analyzed using respective sensors and the data is fed into the Random forest model which recommends the crop and predicts the yield strength of the soil. The outcomes of the recommendation system make the farmers take the right decision for the right crop such that the agricultural sector will be developed with innovative ideas. The outcomes of the crop yield prediction system give an idea of the soil strength and the excepted harvest with the present soil conditions. As a whole, the outcomes of the developed system ensure the farmer chooses the right crop to cultivate and gives him a clear picture of the profit margins.

CHAPTER 6

FUTURE ENHANCEMENTS

The system developed recommends a crop and predicts yield of the field based on soil parameters. In future, the system can be extended with development of a website or app which the farmer can use to view the soil parameters and crop recommended, estimated yield of his field. The machine learning models of crop recommendation system and crop yield predicted system can be trained using more real time sensor data. This ensures even more accuracy of the system.

REFERENCES

[1] Sonika R,"Automatic Soil Nutrients and crop Detection Management using IOT", IJIRST,2021.

[2] Rahul BV,"Monitoring the soil Parameters using Iot and Android Based Application for Smart Agriculture", IEEE, 2021.

[3] Ramya S,"IOT Framework for Smart Farming using Machine Learning Technique",IJRMST, 2020.

[4] R Madhumathi,"Soil NPK and Moisture analysis using Wireless Sensor Networks", IEEE, 2020.

[5] Niizar Ali Alhaj Abdellah,"Real Time Application of Iot for the Agriculture in the Field along with Machine Learning Algorithm", IEEE, 2020.

[6] Smita Pawar,"Design of an IoT enabled Automated Hydroponics System using NodeMCV and Blynk",IEEE, 2019.

[7] Dr. R Kezia Rani,"A Novel IOT Based Solution for Agriculture Field Monitoring and Crop Prediction Using Machine Learning", IJIRST, 2019.

[8] Setia Gumilar,"Real-Time Monitoring System for Measurement of soil Fertility Parameters, Susanto Nugarha", IEEE, 2019.

[9] Suhas Athani,"Soil Moisture monitoring using IoT enabled arduino sensors with neural networks for improving soil management for farmers and predict seasonal rainfall for planning future harvest in North Karnataka- India",IEEE, 2017.

[10] M.K.Gayathri, J.Jayasakthi, Dr.G.S.Anandhamala, "Providing Smart Agriculture Solutions to Farmers for Better Yield using IOT",IEEE, 2015.

APPENDIX

1. Soil Analysis

```
#include <SoftwareSerial.h> #include "DHT.h"
#include <Wire.h>
#include <LiquidCrystal_I2C.h> LiquidCrystal_I2C lcd(0x27, 16, 2); SoftwareSerial ph(5,6);
DHT dht; String pH;
long lastMsg = 0; char msg[50];
int value = 0; String message=""; char val[50];
String top; char a[20]; char c[20]; char b[20];
String getStringPartByNr(String data, char separator, int index)
{
int stringData = 0; String dataPart = "";
for (int i = 0; i < data.length(); i++) { if (data[i] == separator) { stringData++;}
 else if (stringData == index) { dataPart.concat(data[i]);
} else if (stringData > index) { return dataPart;
break;
}
} return dataPart;
}
void setup() { Serial.begin(9600); ph.begin(9600); dht.setup(8);
lcd.begin(); lcd.backlight();
}
void loop() { String t;
```

```
if (ph.available())  {
String rcv = ph.readString();
String p = getStringPartByNr(rcv, ',', 0); pH = getStringPartByNr(p, ':', 1); Serial.print("pH:
    ");
Serial.println(pH);
}
float humidity  = dht.getHumidity();
float temperature  = dht.getTemperature(); if  (isnan(humidity)  || isnan(temperature))
    {Serial.print("  ");
}
else { Serial.print("Humidity");  Serial.println(humidity,  1);
t="Temperature:  " + String(temperature);
}
int moisture = analogRead(A1);
String m = "moisture: " + String(moisture); String p = "PH: " + String(pH); Serial.println(m);
Serial.println(p);   Serial.println(t);   delay(1000);   lcd.clear();
lcd.print(m);  delay(500);  lcd.clear();
lcd.print(p); delay(500);  lcd.clear();
lcd.print(t);  delay(500);
}
```

2. Crop Recommendation System

```
import pandas as pd
import numpy as np
```

```
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.metrics import classification_report
from sklearn import metrics
from sklearn.model_selection import cross_val_score
import pickle
df = pd.read_csv('../Data-processed/crop_recommendation.csv')
features = df[['N', 'P','K','temperature', 'humidity', 'ph', 'Soil Moisture']] target = df['label']
acc = []
model = []
from sklearn.model_selection import train_test_split

Xtrain, Xtest, Ytrain, Ytest = train_test_split(features,target,test_size = 0.2,random_state =2)

from sklearn.ensemble import RandomForestClassifier

acc_score = []
n_estimators = [1, 2, 4, 8, 16, 32, 64, 100, 200]
for i in n_estimators:
    RF = RandomForestClassifier(n_estimators=i, random_state=2)


    RF.fit(Xtrain,Ytrain) predicted_values = RF.predict(Xtest)


    x = metrics.accuracy_score(Ytest, predicted_values) acc_score.append(x)
RF = RandomForestClassifier(n_estimators=16, random_state=2) RF.fit(Xtrain,Ytrain)
predicted_values = RF.predict(Xtest)


x = metrics.accuracy_score(Ytest, predicted_values) acc.append(x)
```

```
model.append('RF') print("RF's Accuracy is: ", x)
print(classification_report(Ytest,predicted_values))
temp=float(input("Enter  the temperature:"))
hum=float(input("Enter  the humidity:"))
ph=float(input("Enter  the ph value:"))
mois=float(input("Enter   the Soil Moisture"))
data = np.array([[57,49,42,temp,hum,ph,mois]])
prediction = RF.predict(data)
print(prediction)
```

3. Crop yield prediction system

```
import numpy as np import warnings
warnings.filterwarnings('ignore')   import pandas as pd
data = pd.read_csv("recom.csv")
clay1=np.repeat('clay',1200)
clay1=clay1.reshape(1200,1)
sandy1=np.repeat('sandy',1300)
sandy1=sandy1.reshape(1300,1)
good1=np.repeat('good_bio_cond',1300)
good1=good1.reshape(1300,1)
l1=np.repeat('alkaline',1300)
al1=al1.reshape(1300,1)
excellent=np.repeat('excellent_bio_cond',1100)
excellent=excellent.reshape(1100,1)
chalky1=np.repeat('chalky  ',177)
chalky1=chalky1.reshape(177,1)
poor=np.repeat('poor_bio_cond',1577)
```

```
poor=poor.reshape(1577,1)
soil1=np.concatenate((clay1,sandy1, al1, chalky1),axis=0)
bio_cond = np.concatenate((good1,excellent, poor),axis=0)
from random import shuffle
print (soil1)
np.random.shuffle(soil1)
np.random.shuffle(bio_cond)
data['soil']=soil1
df = pd.DataFrame(data)

one = pd.get_dummies(df['soil'])

one['alkaline']

df['alkaline']=one['alkaline']

df['sandy']=one['sandy']
df['chalky'] = one['chalky ']

df['clay']=one['clay']

df.drop(['soil'],axis=1)

data['bio']=bio_cond
df = pd.DataFrame(data)

one = pd.get_dummies(df['bio'])

df['excellent_bio_cond']=one['excellent_bio_cond']

df['poor_bio_cond']=one['poor_bio_cond']
```

```
df['good_bio_cond'] = one['good_bio_cond']

data['excellent']=df['excellent_bio_cond']

data['poor']=df['poor_bio_cond']

data['good']=df['good_bio_cond']

data.drop(['bio','soil'],axis=1)
df2 = data.values
data['yield']  = 0

a= 0

b = 0

c= 0

d = 0

arr = np.zeros((3977,1))

for i in  range(3977):

if(((df2[i,  0] < 12.8) or (df2[i,  1] > 0.005)) and (df2[i,  2] <60) and (df2[i,  3] >
55) and ((df2[i,  10] == 1.0) or (df2[i,  8] == 1.0))):

#print i b = b + 1
arr[i,0]  = 2



for i  in  range(3977):
```

```
if(((df2[i, 0] < 12.8) or (df2[i, 1] > 0.5)) and (df2[i, 2] <50) and (df2[i, 3] > 45)
and ((df2[i, 10] == 1.0) or (df2[i, 9] == 1.0) or (df2[i, 8] == 1.0))):

#print i c= c + 1
arr[i,0] = 3



for i in range(3977):
if(((df2[i, 0] < 12.8) or (df2[i, 1] == 0.0)) and (df2[i, 2] <90) and (df2[i, 3] < 90)
and ((df2[i, 10] == 1.0) or (df2[i, 9] == 1.0) or (df2[i, 7] == 1.0) or (df2[i, 8] ==
1.0))):

#print i d= d + 1
arr[i,0] = 4


for i in range(3977):

        if(((df2[i, 0] > 12.6 and df2[i, 0] < 13.5) or (df2[i, 1] > 0.000001)) and (df2[i, 2]
      > 40 and df2[i, 2] < 100) and (df2[i, 3] > 52 and df2[i, 3] < 100) and (df2[i, 10] ==
1.0) ):

#print i a = a + 1
arr[i,0] = 1 data['final yield']=arr
data = data.drop(['excellent_bio_cond','poor_bio_cond','good_bio_cond'],axis=1)
data = data.drop(['soil','bio','excellent','poor','good','yield'],axis = 1)
a = np.unique(data['final yield'])
data.to_csv('finalyield.csv')
from sklearn.preprocessing import OneHotEncoder
```

```
dt = pd.read_csv('finalyield.csv')
X = dt.drop(['final yield'],axis=1)
Y = dt['final  yield']
Y= Y.to_frame()
Y = Y.values

onehot_encoder = OneHotEncoder(sparse=False) Y = Y.reshape(len(Y), 1)
onehot_encoded = onehot_encoder.fit_transform(Y) print (onehot_encoded)
from sklearn.model_selection  import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, Y,test_size=0.3) from sklearn.tree import
    DecisionTreeClassifier
from sklearn.ensemble import RandomForestClassifier rf =
    RandomForestClassifier(n_estimators=140) rf.fit(X_train,y_train)
yeild  = rf.predict(X_test)
from matplotlib import pyplot as plt test_vector =
np.reshape(np.asarray([12.737998,0.026821,61,56,70,42,1.0,0.0,0.0,0.0,0.0]),(1,11
))

p = int(rf.predict(test_vector)[0])

yield_list  = ['20-30%','30-40%','50%','60-70%','80-100%']

print (yield_list[p])
```