A PROJECT REPORT ON

Smart Cultivation and Prediction System for Agriculture

SUBMITTED TO THE SAVITRIBAI PHULE PUNE UNIVERSITY, PUNE

IN THE PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF THE DEGREE

OF

BACHELOR OF ENGINEERING (COMPUTER ENGINEERING)

BY

Vedant Parnaik (B150454294)
Kedar Terkhedkar (B150454331)
Apoorva Parnate (B150454295)
Tanay Sapre (B150454312)

Under The Guidance of

Prof. Pradnya Mehta

'येथे बहुतांचे हित'

DEPARTMENT OF COMPUTER ENGINEERING

MARATHWADA MITRA MANDAL'S COLLEGE OF ENGINEERING,

KARVENAGAR, PUNE-411052

SAVITRIBAI PHULE PUNE UNIVERSITY,PUNE

A. Y. 2020-21

CERTIFICATE

This is to certify that the Project Entitled

Smart Cultivation and Prediction System for Agriculture

Submitted by

Vedant Parnaik (B150454294)
Kedar Terkhedkar (B150454331)
Apoorva Parnate (B150454295)
Tanay Sapre (B150454312)

is a bonafide work carried out by students under the supervision of **Prof. Pradnya Mehta** and it is approved for the partial fulfillment of the requirement of Savtribai Phule Pune university, Pune for the award of the degree of Bachelor of Engineering (Computer Engineering)

Prof. Pradnya Mehta
Internal Guide,
Dept. of Computer Engg.

Prof. H. K. Khanuja
HOD,
Dept. of Computer Engg.

Name and Sign of
External Examiner

Dr.S.M. Deshpande
Principal

PROJECT APPROVAL SHEET

A Project Report Titled as

Smart Cultivation and Prediction System for Agriculture

is verified for its originality in documentation, problem statement, proposed work and implementation successfully completed by

Vedant Parnaik (B150454294)
Kedar Terkhedkar (B150454331)
Apoorva Parnate (B150454295)
Tanay Sapre (B150454312)

at

DEPARTMENT OF COMPUTER ENGINEERING
MARATHWADA MITRA MANDAL'S COLLEGE OF ENGINEERING,
KARVENAGAR, PUNE-411052
SAVITRIBAI PHULE PUNE UNIVERSITY,PUNE

ACADEMIC YEAR 2020-21

Prof. Pradnya Mehta
Internal Guide,
Dept. of Computer Engg.

Prof. H. K. Khanuja
HOD,
Dept. of Computer Engg.

ACKNOWLEDGEMENT

We take this to express our deep sense of gratitude towards our esteemed guide Prof. **Pradnya Mehta** for giving us this splendid opportunity to select and present this project and also providing facilities for successful completion.

We thank **Dr.S.M.Deshpande** ,Principal,Marathwada Mitra Mandal College of Enginnering Karvenagar ,Pune 52 and **Prof. Harmeet Khanuja** , Head, Department of Computer Engineering, for opening the doors of the department towards the realization of the project, all the staff members, for their indispensable support, priceless suggestion and for most valuable time lent as and when required. With respect and gratitude, We would like to thank all the people, who have helped us directly or indirectly.

At last we want to express our sincere gratitude to all the staff members of Computer Engineering Department who helped us directly or indirectly during this course of work.

Vedant Parnaik
Kedar Terkhedkar
Apoorva Parnate
Tanay Sapre

]

ABSTRACT

A vast fraction of the population of India considers agriculture as their primary occupation. Production of crops plays an important role in our country. Bad quality crop production is often due to either excessive use of fertilizer or using not enough fertilizer. For efficient crop growth, it is essential to measure the level of nutrients present in the soil. The proposed system of IoT and ML is enabled for soil testing using the sensors, is based on measuring and observing soil parameters. This system lowers the probability of soil degradation and helps maintain crop health. Different sensors such as soil temperature, soil moisture, NPK, pH, are used in this system for monitoring temperature, humidity, soil moisture and soil pH along with NPK nutrients of the soil respectively. The data sensed by these sensors is stored on the microcontroller and analyzed using machine learning algorithms based on which suggestions for the growth of the suitable crop is made. This project also has a methodology that focuses on using a convolutional neural network as a primary way of identifying if the plant is at risk of a disease or not. Using image processing and machine learning along with wireless sensor networks, the crop is monitored and sensor values of various climatic conditions are generated. These machine learning algorithms take these sensor values as input and predict if the crop is healthy or not and its lifetime, which will help the farmer make effective decisions to increase the yield of the crop.

Contents

List of Figures

Technical Keywords

0.1 Domain Name

Internet of Things and Machine Learning

0.2 Technical Keywords

Internet of Things, Machine Learning, Firebase Cloud Database, Raspberry Pi, Sensor network, Random forest, Convolutional Neural Network (CNN), Streamlit.

Chapter 1

Introduction

1.1 Overview

Achieving a maximum yield rate of crop using limited land resources is a goal of agricultural planning in an agro-based country. Prior determination of the problems associated with crop yield can increase yield rate of crops. To get a good crop it is important that the land should have adequate fertilizer. Hence, Soil testing is used to do chemical analysis of soil and to find the level of fertility of soil. Based on which suggestion of appropriate crop and prediction of required nutrients of the crops can be done. Soil testing includes testing of soils for properties like pH, moisture, Nitrogen (N), Phosphorus (P), Potassium (K).Soil testing also includes monitoring temperature and humidity of soil as well as atmosphere. All the data of nutrients of the soil is sensed by various sensors is stored on the microcontroller and analyzed using Random Forest algorithm based on which suggestions are made for the suitable crop for the soil and additional nutrients to be used for good growth of the crop.

Pests and diseases affecting the crops/plants also cause a tremendous decrease in production of the crops. In the majority of cases pests or diseases are seen on the leaves or stems of the crop/plant. Therefore, identification of plants, leaves, stems and finding out the diseases, percentage of pest or disease, symptoms of the pest or disease attack, plays a key role in successful cultivation of crops.

Hence, in order to increase crop productivity, farmers need to approach the experts to seek their advice regarding the treatment of incidence of pests and diseases to their crops and suggestions for control of the pests and diseases. To make it easier the crop images can be scanned and uploaded which will be analyzed using image processing and Convolution Neural Network. CNN will identify and predict if the plant is at a risk of a disease or not. If a plant is at a risk of some disease the model can suggest ways to the farmer to curb the spread of the disease. It will help the farmer to make effective decisions to increase crop yield and helps maintain crop health.

1.2 Motivation

Every day we find that the environment is changing continuously which is harmful to the crops and leading farmers towards debt and suicide. In many cases like this and with a growing population to maximize yield, farmers are using more pesticides and fertilizers which are leading to the soil infertility as well as decreasing the holding capacity of the soil and increasing toxicity of the soil. Farming land is used by growing industrialization, so again increasing rate of soil pollution which affects the quality of plants.

1.3 Problem Definition and Objectives

Our country is facing various problems related to agriculture such as lack of knowledge about the suitable crop, the concentration of the nutrients, plant pathology. Smart cultivation and Prediction System for Agriculture provides the best solution for the suitable crop, nutrients issues and plant pathology.

1.4 Project Scope and Limitations

Project scope:- This system is used to recommend crop to be sown based on the nutrients present in the soil. Also, it predicts the diseases the crop can have based on the crop images and suggest suitable remedies for the same.

Limitations:-

1. The sensor module can be easily affected by water. Being an agricultural project, special care has to be taken for damage done by rains hence, all the hardware needs to be waterproofed.

2. Proper care needs to be taken of voltage supplied by the battery. Due to this issue, IC and sensors can get heavily damaged beyond repair.

3. The system cannot detect diseases for all the plants.

4. The lighting conditions as well as extra head gears can affect the system performance.

Chapter 2

Literature Survey

Sr.No	Title	Methodology	Pros	Cons
1	Seasonal Crops Disease Prediction and Classification Using Deep Convolutional Encoder Network	Convolutional encoder network	High Accuracy (95 percent)	Requires high processing power
2	Cucumber disease detection using artificial neural network.	ANN, GLCM(Gray level co-occurrence method)	Uses deep learning	Classification accuracy low
3	Svm Classifier Based Grape Leaf Disease Detection.	K-means clustering algorithm with SVM, Color co-occurrence method.	Low processing power	Low accuracy
4	Detection of leaf disease and classification using digital image processing.	GLCM, SVM, K-means	High accuracy	High processing power

Table 2.1: Literature review

Sr.No	Title	Methodology	Pros	Cons
5	Smart Agriculture Using Internet of Things with Raspberry Pi	All the sensors are mounted in a box with connections done through breadboards	All the sensors are located at a single place which will make debugging easy.	In case of short circuit of the breadboard, the complete apparatus will be damaged.
6	Integrated optical sensor for NPK Nutrient of Soil detection.	The mechanism uses colorimetry. The wavelengths of suitable colors are measured and thus we find respective nutrients.	Fibre Optic sensors give precised readings in an enclosed space.	Interference of external light affect the readings.
7	Real-time and Low-cost IoT based farming using raspberry Pi.	The mechanism focuses on real-time observation with efficient use of cheapest security system.	The sensor data monitored is sent to the cloud for processing.	In case of network failure, temporary data logging in not taken into consideration.
8	Precision agriculture monitoring system using wireless sensor network and Raspberry Pi local server.	The mechanism focuses on deploying a low-cost sensor system, gathering field data, and displaying the data through a graphical user interface (GUI).	The whole system was tested and proven to work by the application of fertilizer to the soil and seeing its response in the GUI.	Readings collected are used for data visualization and not for any further analysis.

Table 2.2: Literature review

2.1 Description

Description:

Paper 1-
"Seasonal Crops Disease Prediction and Classification Using Deep Convolutional Encoder Network"

The paper titled "Seasonal Crops Disease Prediction and Classification Using Deep Convolutional Encoder Network" focuses on modified custom CNNs to achieve higher prediction accuracy. The proposed system has a hybridized deep learning neural network and named it a convolutional encoder network. It is a combination of both CNN and autoencoders but uses only the encoding part of the autoencoders to obtain useful features. This system uses immense processing power but compensates it with higher accuracy.

Paper-2
" Recent developments of the Internet of Things in Agriculture: A Survey"

The research paper titled 'Recent developments of the Internet of Things in Agriculture: A Survey' outlines recent insights in the development and advancement of the internet of things in agriculture. The paper also briefs about layered architecture of Agro-IoT, Multimedia Internet of Things (MIoT), Industrial Internet of Things (IIoT). It also discusses hardware platforms along with the sensors based on their applications. It focuses on IoT based farm management systems, irrigation systems, crop monitoring and disease prediction.

Paper-3
"Machine Learning-based Grape Leaf Disease Detection"

The paper titled "Machine Learning-based Grape Leaf Disease Detection", is imagined to aid the identifying and arranging leaf illnesses utilizing a Multiclass Support Vector Machine (SVM) grouping system. In the first place, the influenced area is found by K-means clustering, and after that highlights like shading and surface are extricated. The framework can effectively characterize the analyzed infection with an accuracy of 88.89.

Paper-4
"Detection of Leaf Diseases and Classification using Digital Image Processing"

In the Research paper "Detection of Leaf Diseases and Classification using Digital Image Processing", the segmentation of leaves is done using the K- Means algorithm. Texture features are extracted using GLCM(gray-level co-occurrence matrix)and then classification is done using SVM. The Framework is fairly simple as opposed to deep neural networks, yet it achieves an accuracy of more than 90

Paper-5
"Smart Agriculture, Using Internet of Things with Raspberry Pi"

The paper titled 'Smart Agriculture, Using Internet of Things with Raspberry Pi' has all the sensors mounted in a box with connections done through breadboards. All the sensors are located in a single place which will make debugging easy

Paper-6
"Integrated optical sensor for NPK Nutrient of Soil detection"

The research paper titled 'Integrated optical sensor for NPK Nutrient of Soil detection' uses The mechanism uses colorimetry. The wavelengths of suitable colors are measured and thus we find respective nutrients. Fibre Optic sensors give precise readings in an enclosed space

Paper-7
"Real-time and Low-cost IoT based farming using raspberry Pi"

The research paper titled 'Real-time and Low-cost IoT based farming using raspberry Pi' focuses on real-time observation with efficient use of the cheapest security system. The sensor data monitored is sent to the cloud for processing

Paper-8
"Precision agriculture monitoring system using wireless sensor network and Raspberry Pi local server"

The paper titled 'Precision agriculture monitoring system using wireless sensor network and Raspberry Pi local server ' focuses on the mechanism focusing on deploying a low-cost sensor system, gathering field data, and displaying the data through a graphical user interface (GUI). The whole system was tested and proven to work by the application of fertilizer to the soil and seeing its response in the GUI.

Paper-9

"Security and Privacy for Green IoT-Based Agriculture: Review, Blockchain Solutions, and Challenges"

The research paper titled 'Security and Privacy for Green IoT-Based Agriculture: Review, Blockchain Solutions, and Challenges' presents research challenges on security and privacy issues in the field of green IoT-based agriculture. It starts by describing a four-tier green IoT-based agriculture architecture and summarizing the existing surveys that deal with smart agriculture. Then, it provides a classification of threat models against green IoT-based agriculture into five categories, including, attacks against privacy, authentication, confidentiality, availability, and integrity properties. Moreover, it provides a taxonomy and a side-by-side comparison of the state-of-the-art methods toward secure and privacy-preserving technologies for IoT applications and how they will be adapted for green IoT-based agriculture. In addition, it analyzes the privacy-oriented blockchain-based solutions as well as consensus algorithms for IoT applications and how they will be adapted for green IoT-based agriculture.paper link

Paper-10

"Automated disease classification in (Selected) agricultural crops using transfer learning"

The paper titled "Automated disease classification in (Selected) agricultural crops using transfer learning" is based on different CNN based architectures namely AlexNet,VGG16, VGG19, GoogLeNet, ResNet, DenseNet, etc.,Which have been developed and adopted for solving the problem of disease classification in various crops to achieve higher prediction accuracy. The augmented image dataset was used for training and validation of the six deep learning models. These trained and validated models were deployed to classify the given image. This system uses immense processing power but compensates it with higher accuracy .

Paper-11
"Smart Sensing System for Precision Agriculture"

The paper titled “Smart Sensing System for Precision Agriculture” is based on a sensing system using proximity sensors and low-cost smartphones to utilise all its sensors like accelerometers, gravity sensors, GPS,etc. This paper concludes that new technologies need to be adapted as they arise and overlooked technologies resurrected. Large sections of the electromagnetic spectrum are being used, as well as ultrasonic, electrical resistivity, and physical measurements but others, such as magnetic susceptibility, seem to be underexploited, which might be useful for future investigation.

Paper-12
"A Survey on the Role of IoT in Agriculture for the Implementation of Smart Farming."

The research paper titled 'A Survey on the Role of IoT in Agriculture for the Implementation of Smart Farming' aims to develop a complete precision agriculture system based on IoT and cloud computing. A lot of sensors are used in the system to monitor real time data of crops/ plants, animals, gases and climate. It can be also used in the greenhouses. The complete system is of great use for collection and organising the data.

Paper-13

"Internet of Things in Smart Agriculture: Enabling Technologies"

The research paper titled 'Internet of Things in Smart Agriculture: Enabling Technologies' proposed an IoT technology research and innovation roadmap for the field of precision agriculture (PA) is presented. Many recent practical trends and the challenges have been highlighted. Some important objectives for integrated technology research and education in precision agriculture are described. Effective IoT based communications and sensing approaches to mitigate challenges in the area of precision agriculture are presented.

Paper-14

"IoT Applications in Agriculture: A Systematic Literature Review"

The research paper titled 'IoT Applications in Agriculture: A Systematic Literature Review' provides farmers and researchers with a clear perspective of IoT applications in agriculture. It presents a systematic literature review of IoT-based tools and applications for agriculture. The objective of this paper is to offer an overview of the IoT applications in agriculture through topics such IoT-based software applications for agriculture available in the market, IoT-based devices used in agriculture, as well as the benefits provided by this kind of technologies.

Paper-15

"How Can Heterogeneous Internet of Things Build our Future: A Survey"

The research paper titled 'How Can Heterogeneous Internet of Things Build our Future: A Survey' proposes a four-layer HetIoT architecture consisting of sensing, networking, cloud computing, and applications. Then, the state of the art in HetIoT research and applications have been discussed. It also suggests several potential solutions to address the challenges facing future HetIoT, including self-organizing, big data transmission, privacy protection, data integration and processing in large-scale HetIoT

Paper-16
"Tea Leaf Diseases Recognition using Neural Network Ensemble"

The paper titled "Tea Leaf Diseases Recognition using Neural Network Ensemble" proposes a tea leaf disease recognizer (TLDR)an initiative to recognize diseases of the tea leaf. In TLDR, at first the image of the tea leaf is cropped, resized and converted to its threshold value in the image processing.Then feature extraction method is applied. Neural NetworkEnsemble (NNE) was used for pattern recognition. The extracted features are passed to the ANN along with the disease type and the ANN is trained.After going through the testing process 91 % of accuracy was found

Chapter 3

Software Requirement Specification

3.1 Assumptions and Dependencies

Assumptions:-

1. The battery voltage should be in the required range.
2. There should not be any loose connections in case of mishandling.
3. The apparatus should be clean before usage. Murky apparatus might lead to vague readings.
4. The user should be able to use the system properly.

Dependencies:-

1. Low battery voltage may not initialize the sensor, hence no parameters will be measured.
2. Loose connections will lead to loss in the data.
3. If the apparatus is murky before its usage then the readings will the imprecise.

3.2 Functional Requirements

3.2.1 Measure necessary soil parameters

Sensing and measuring NPK, pH, soil moisture, soil temperature, surrounding humidity and temperature is an important aspect.

3.2.2 Analyze the data measured

The data will be well organized and sent to raspberry pi to process it by the machine learning algorithms.

3.2.3 Suggest a suitable crop/plant

Analyzing the input information by using machine learning algorithms to provide accurate crop suggestions for maximum yield.

3.2.4 Plant pathology

Predict if any diseases are affecting the crop/plant. Also, suggest remedies to avoid further spread of diseases.

3.3 External Interface Requirements

3.3.1 User Interface

- Sign up - It allows to user to create an account.
- Login - It allows user to enter his username and password.
- Crop suggestion screen - Displays all the values sensed by sensors and gives crop suggestion based on the same. It also displays ways to increase crop production by suggesting fertilizers if needed for the crops.
- Disease identification screen - Displays the disease prediction and the suggestions for its cure.

3.3.2 Hardware Interface

- This project will need an Arduino Mega 2560 and Raspberry Pi 4.
- It will need sensors to check soil parameters such as pH, NPK, soil moisture, soil temperature, surrounding temperature and humidity sensor.
- It will need a camera to click pictures of the crops.
- It will also need a tablet to display results.

3.3.3 Software Interface

- Python Version - 2.7.1 or above.
- Operating System - Linux 16.04 above.
- Arduino IDE - 1.7 or above.

3.3.4 Communication Interface

Serial Communication between arduino and raspberry pi

3.4 Non-functional requirements

3.4.1 Performance Requirement

Response time : The system should perform with minimum delay. Loose connections result in skipping the data while transferring hence, there should not be any loose connections.

3.4.2 Security Requirements

System is secure as it is automated.

3.4.3 Safety Requirements

- The emergency kill switch will keep the complete setup safe in case of any malfunctions.
- Reverse voltage protection circuit will be used to keep the battery safe.

3.4.4 Software Quality Attributes

- Correctness : The values sensed should be correct.
- Reliability : This system is reliable as it has sensors mounted on it for accurate result. Machine learning models give accurate results consistently with the desired data input. It also uses high pixel cameras on it for accurate result and image processing to detect the disease.
- Learnability : The user interface should present information as close to reality as possible and permit efficient utilization of the software failures.The prediction and recommendation model should be as precise as possible.
- Robustness : This system performs wired communication and hence there is no data loss.
- Maintainability : System should be properly maintained and battery life of the system should be properly managed.Deep learning models should be routinely re-trained with newer data from customer for maintaining accurate results and/or an increase in accuracy.
- Extensibility : The product should be able to work with extended features.
- Efficiency : The product should work efficiently under heavy work-load. machine learning prediction and recommendation models should be able to output results in lowest possible time with highest accuracy.
- Availability : The data is processed on the local host and machine learning models are pre-trained hence does not require internet. Thus, making it available in rural regions as well.

- Usability: The customer should be able to specify nutrients and put up picture of plants/crop.

3.5 System Requirements

3.5.1 Database Requirements

Labelled dataset required for training the model.

3.5.2 Software Requirements

1. OpenCV version- 3.0 and above.
2. Python version- 2.7.16 and above.
3. Operating Systems- Linux 16.04 and above.
4. Cloud - Firebase Realtime Database

3.5.3 Hardware Requirements

1. Raspberry Pi 4
2. Arduino Mega 2560
3. Soil pH sensor
4. Soil moisture sensor
5. Soil temperature sensor
6. Surrounding temperature and humidity sensor
7. Soil NPK sensor
8. Soil pH sensor
9. Camera

10. Touchscreen display
11. Processor - Intel 6th Gen. onwards.
12. Hard Disk - 50GB.
13. Memory - 4GB RAM

3.6 Analysis Models

SDLC expands to Software Development Life Cycle. It is a process used to design, develop and test, maintain high quality software. The SDLC aims to produce a high quality software that meets or exceeds customer expectations, reaches completion within times and cost estimates. We used incremental SDLC model because in this system every time when there a match occurs we need to work on data, add the data to system so that it will be helpful for further work.
SDLC is a process, which is followed for a software project. It consists of a detailed plan describing how to develop, maintain, replace and alter or enhance specific modules within a software. The life cycle defines a methodology for improving the quality of software and the overall development process.

Diagram of Incremental model:

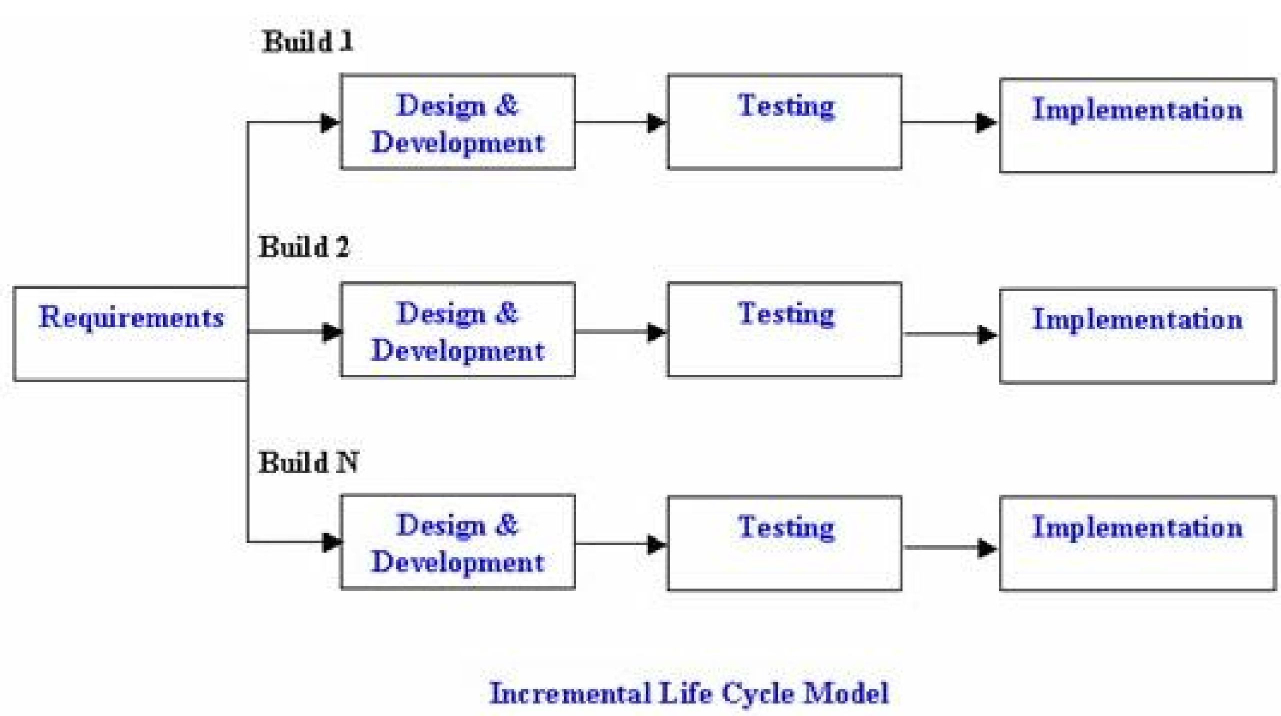

Figure 3.1: SDLC model

So we have used incremental model for building of our project. We have built the systems in iterations.We keep on building the part of the system and testing it with the expected results,so finally combining all the iterations we got the final system's output.

21

Chapter 4

System Design

4.1 System Architecture

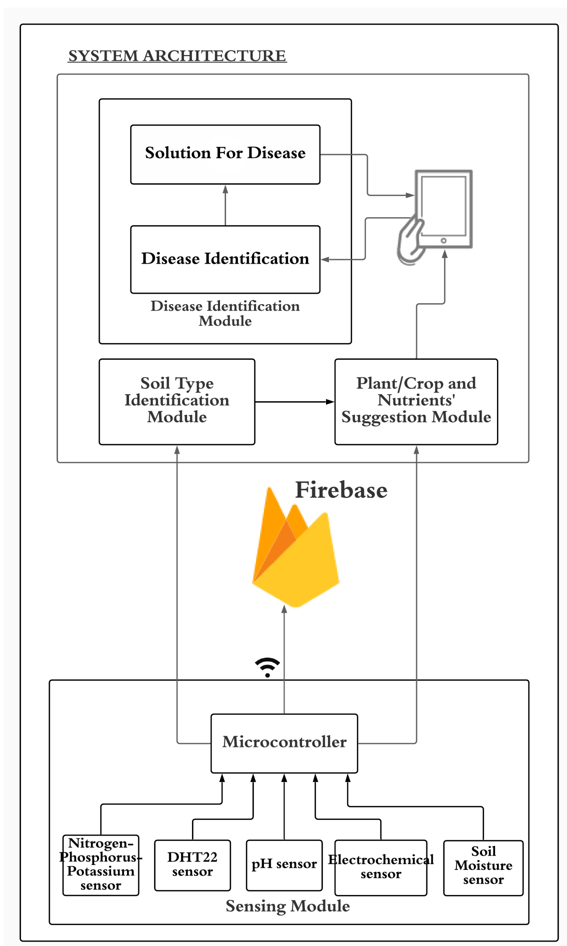

Figure 4.1: System Architecture

4.2 Data Flow Diagrams

A data-flow diagram (DFD) is a graphical representation of the flow of data through an information system. A DFD shows what kind of information will be input to and output from the system, where the data will come from and go to, and where the data will be stored. It does not show information about the timing of processes, or information about whether processes will operate in sequence or in parallel.

DFD Level 0 Diagram

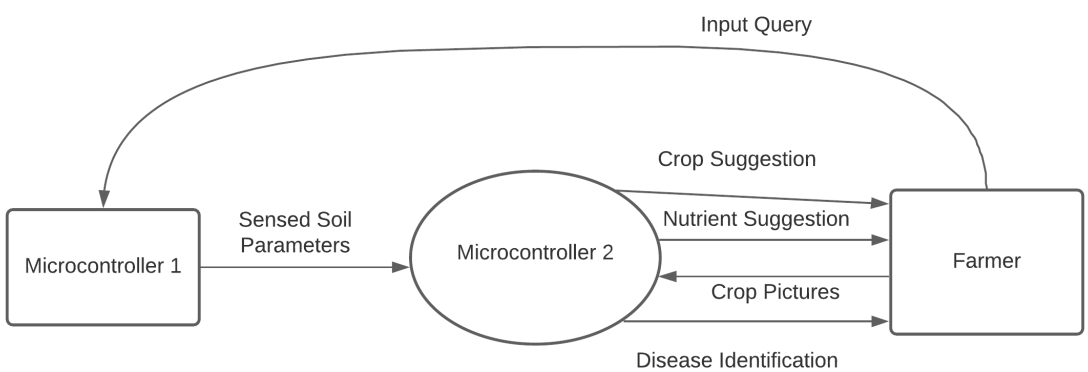

Figure 4.2: DFD Level 0

DFD Level 1 Diagram

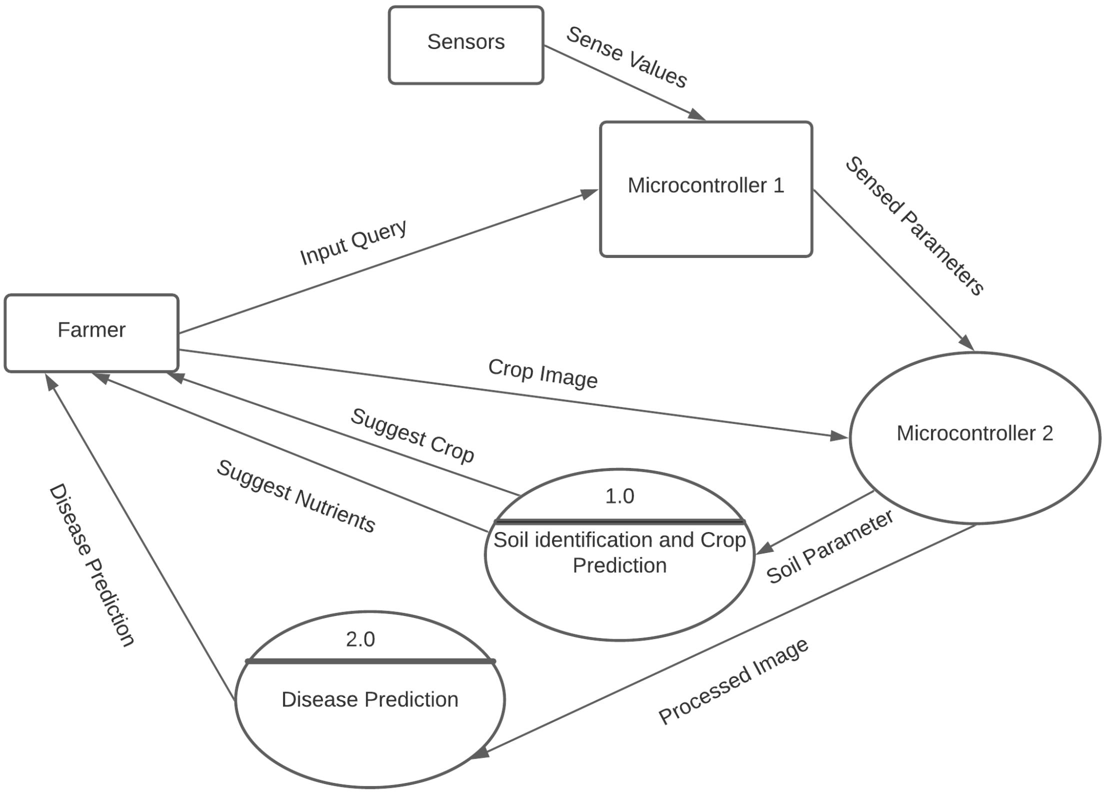

Figure 4.3: DFD Level 1

DFD Level 2 Diagram

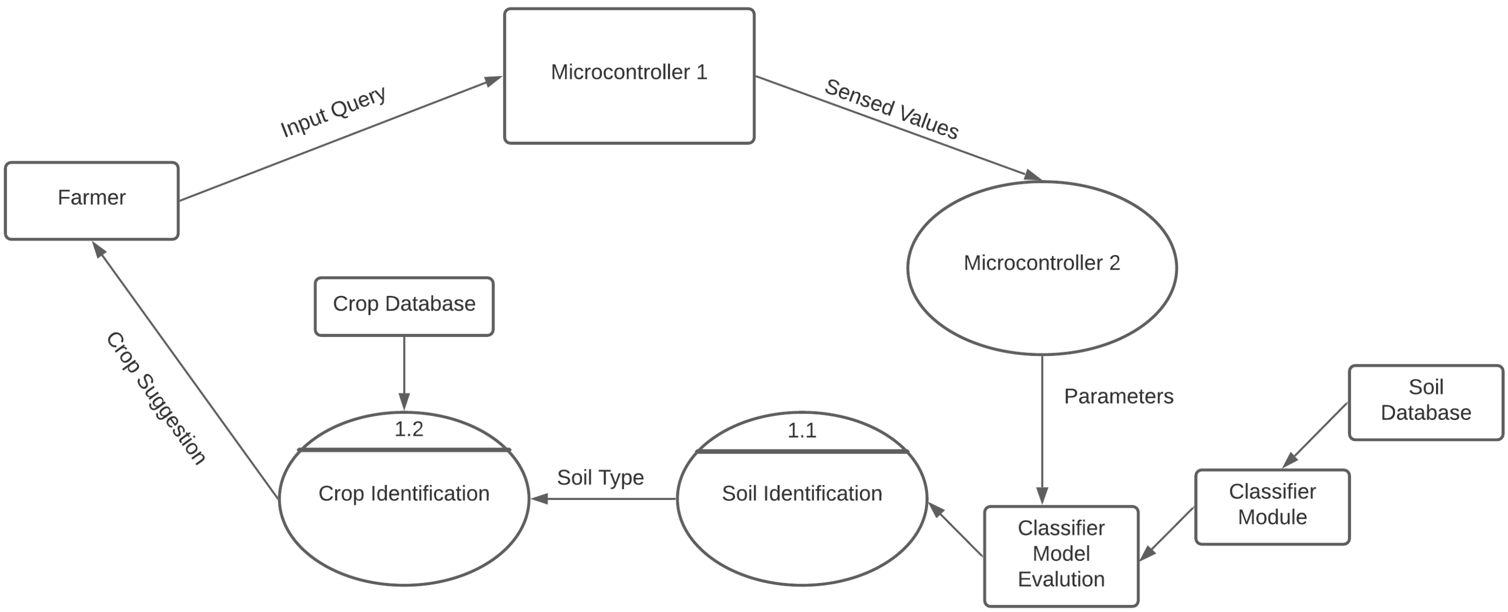

Figure 4.4: DFD Level 2

DFD Level 3 Diagram

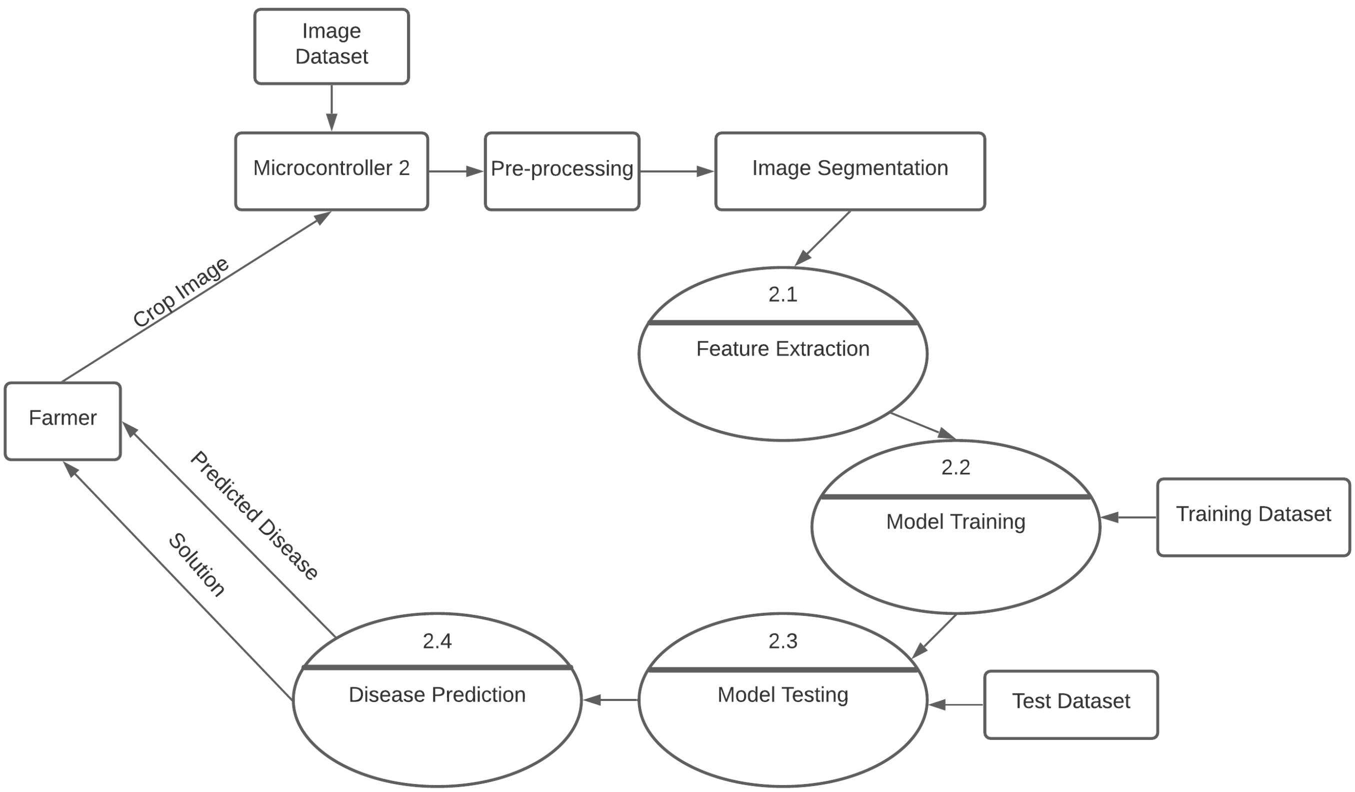

Figure 4.5: DFD Level 3

4.3 Entity Relationship Diagram

It shows us the relation between the entities.

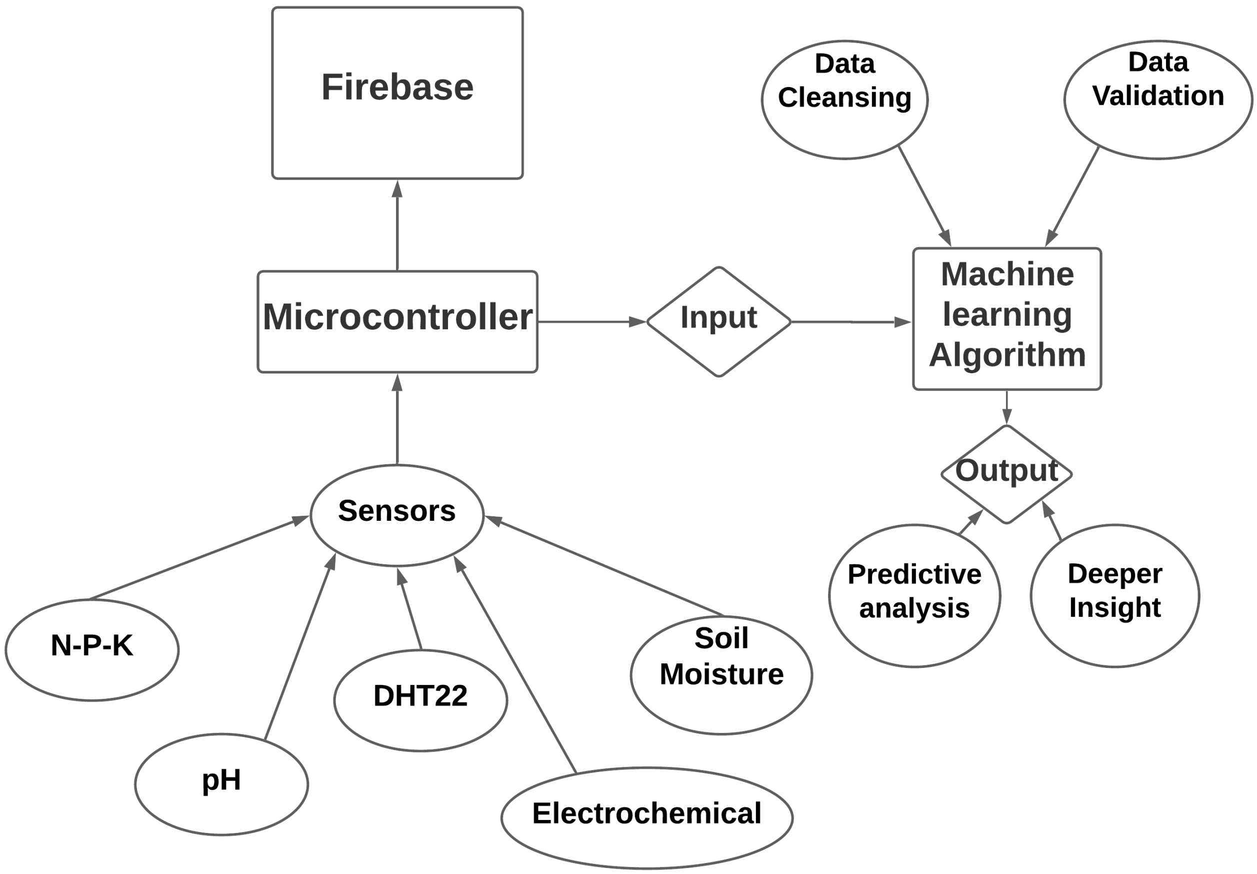

Figure 4.6: Entity Relationship Diagram

4.4 UML Diagrams

4.4.1 Use case Diagram

The use case diagram shows the interaction between the actor and the system. It consists of:

1. User(GCC)
2. Surrounding environment
3. Automated system

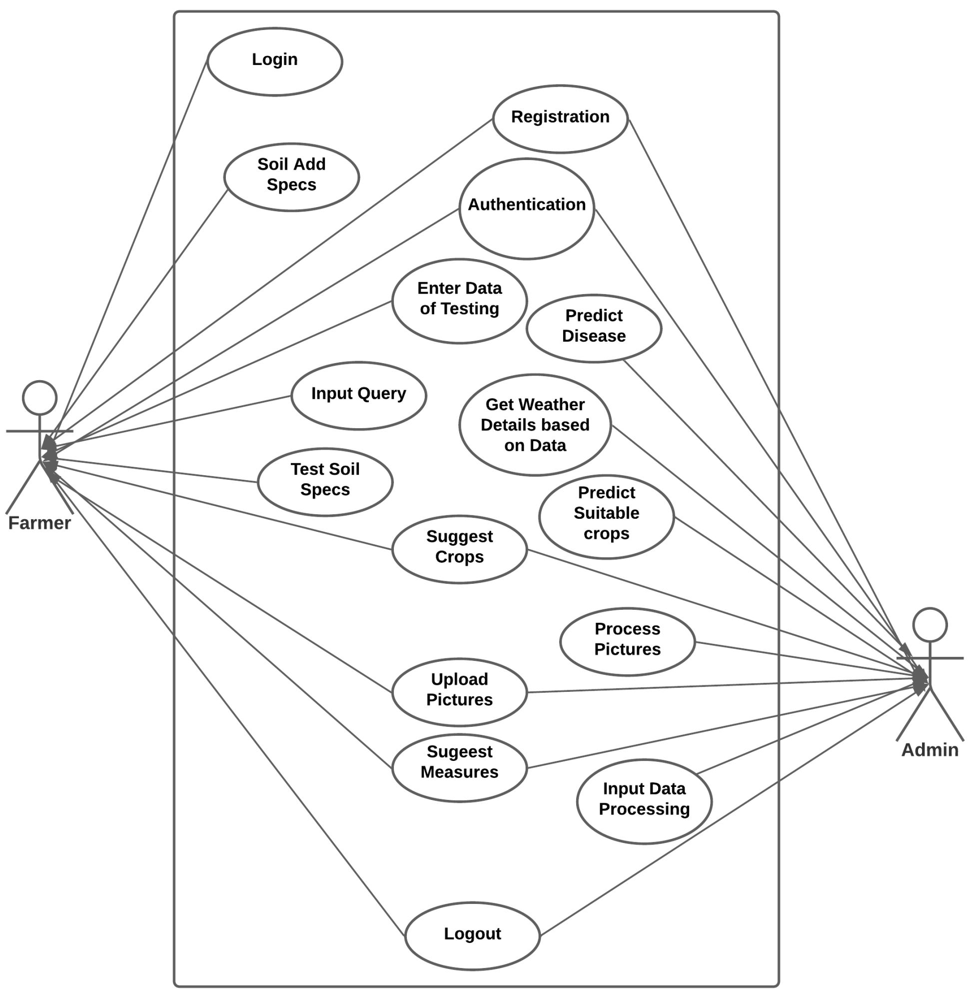

Figure 4.7: Use Case Diagram

4.4.2 Class Diagram

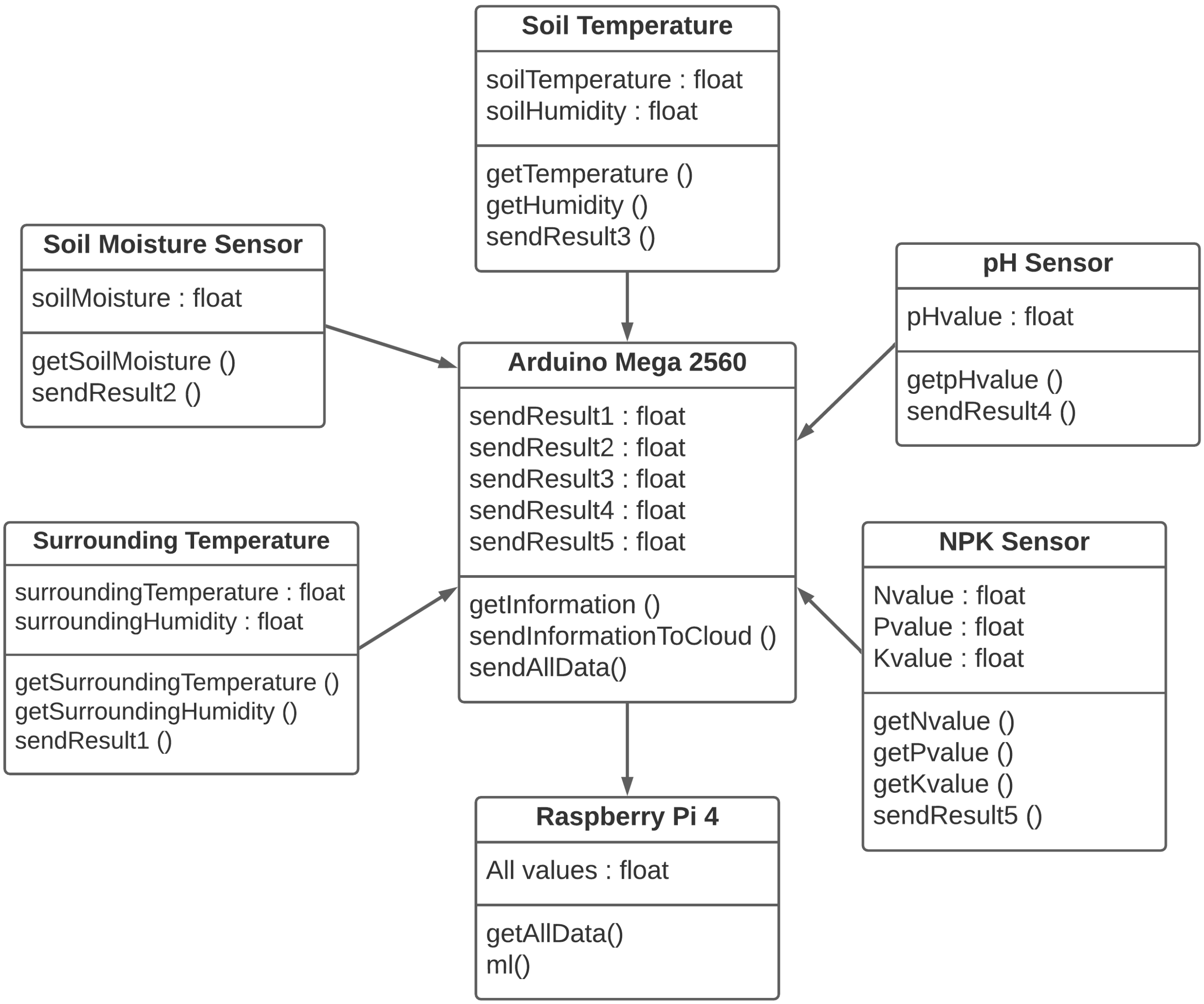

Figure 4.8: Class Diagram

4.4.3 Activity Diagram

In general Activity diagrams is used to show different activities which are performed by the system.

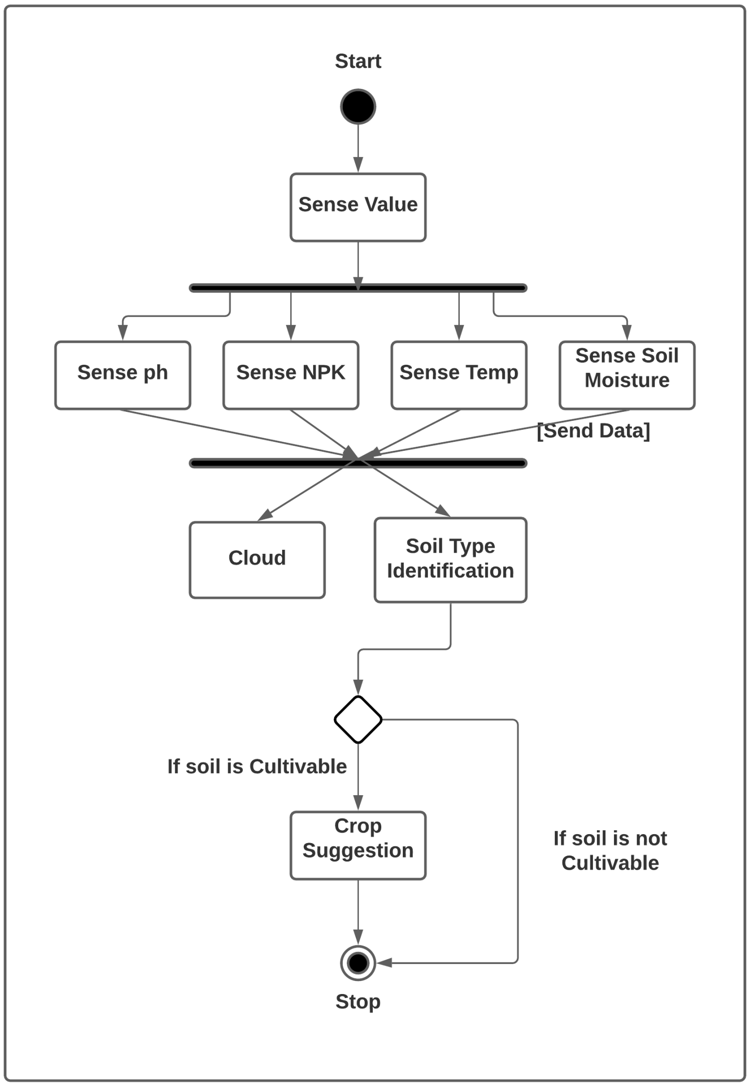

Figure 4.9: Activity Diagram

Chapter 5

Project Plan

5.1 Project Estimate

The COnstructive COst MOdel (COCOMO) is most widely used software estimation model in the world. The COCOMO model predicts the efforts and duration of a project based on inputs relating to the size of the resulting systems and number of "cost drive" that affect productivity.
Three Development Models:

- Organic Mode: Relatively small, simple software projects in which small teams with a good application experience work to a set of less than rigid requirements.

- Semi-detached Mode: An intermediate, (in size and complexity), software project in which teams with missed experience levels must meet a mix of rigid and less than rigid requirements.

- Embedded Mode: A software project that must be developed within a set of tight hardware, software and operation constraints.

Effort Computation: The basic COCOMO model computes effort as a function of a program size.
The basic COCOMO equation is :
Effort = a*KLOCb , in person/months with
KLOC=Lines of code, (in the thousands) inpersons/months.

For duration and staffing,
Duration = C*Efforts in months.

5.1.1 Project Resources

People	4 team members, 1 internal guide, 1 external guide
Hardware	Major components-RPI4B and Arduino mega2560
Software	Technology used: Python3.7. Libraries: Firebase-Admin, Tensorflow Keras, OpenCV, Streamlit, OS, Numpy, Pandas

5.2 Risk Management

Risk is inevitable in a business organization when undertaking projects. However, the project manager needs to ensure that risks are kept to a minimal. Risks can be mainly divided between two types, negative impact risk and positive impact risk.

5.2.1 Risk Identification

Risks, such as operational or business risks will be handled by the relevant teams. The risks that often impact a project are supplier risk, resource risk and budget risk. Supplier risk would refer to risks that can occur in case the supplier is not meeting the timeline to supply the resources required.
Resource risk occurs when the human resource used in the project is not enough or not skilled enough. Budget risk would refer to risks that can occur if the costs are more than what was budgeted.

5.2.2 Risk Analysis

Risk analysis is the systematic study of uncertainties and risks we encounter in business, engineering, public policy, and many other areas. Risk analysts seek to identify the risks faced by an institution or business unit, understand how and when they arise, and estimate the impact (financial or otherwise) of adverse

outcomes. Risk managers start with risk analysis, then seek to take actions that will mitigate or hedge these risks.
Various risk identified for our project are :

1. Communication failure between system and cloud
2. Loose connections in the system hardware

5.2.3 Overview of Risk Mitigation, Monitoring, Management

Risk Mitigation

Risk mitigation planning is the process of developing options and actions to enhance opportunities and reduce threats to project objectives. Risk mitigation implementation is the process of executing risk mitigation actions. Risk mitigation progress monitoring includes tracking identified risks, identifying new risks, and evaluating risk process effectiveness throughout the project. Mitigation for identified Risks are:

1. Checking for a Wi-Fi connection and establishing a link between the system and hardware before sending the data from the system to the cloud database.
2. Soldering all the wired connections so that there won't be any loose connections.

Risk Monitoring

Risks can be monitored on a continuous basis to check if any change is made. New risks can be identified through the constant monitoring and assessing mechanisms.

Risk ID	1
Risk Description	Failing to converge to optimum solution.
Category	Software.
Source	Identified during early development and testing.
Impact	High.
Response	Mitigate.
Strategy	Cross validation can be used for training.
Risk Status	Identified.

Risk ID	2
Risk Description	Wrong data might be fed.
Category	Software.
Source	Identified during early development and testing.
Impact	High.
Response	Mitigate.
Strategy	Validation can be put upon the data that is delivered to the service.
Risk Status	Identified.

Risk ID	3
Risk Description	Failing to converge to optimum solution.
Category	Software.
Source	Identified during early development and testing.
Impact	High.
Response	Mitigate.
Strategy	Cross validation can be used for training.
Risk Status	Identified.

5.3 Project Schedule

Project scheduling is a mechanism to communicate what tasks need to get done and which organizational resources will be allocated to complete those tasks in what time frame. A project schedule is a document collecting all the work needed to deliver the project on time.

5.3.1 Project Task Set

Work Task	Description	Duration.
Basic Study	Related work done for Compendium	3 weeks.
Review of papers	Analysis of different techniques used	3 weeks.
Problem Formulation	Critical analysis and results achieved in research	2 weeks.
Literature Survey	Comparison of technology Studied and result analysis	3 weeks.
Objective of Topic	find some objectives related to system	2 weeks.
SRS	installation and understanding of software required	2 weeks.
Design documentation	UML diagrams and System Structure	2 weeks.
Implementation of base Paper	coding and first module	4 weeks.
Selection of hardware	hardware implentation	2 weeks.
Testing of system	Test system quality, fix errors if any and improve if needed. Test system for different inputs	2 weeks.
Report writing	Prepare initial report	2 weeks
Final report with modification	Prepare and modify initial report	1 week

5.3.2 Task Network

A Task Network (Activity Network) is a graphic representation of the task flow for a project. A task network depicts each software engineering task, its dependency on other tasks, and its projected duration. The task network is used to compute the critical path, a timeline chart and a variety of project information.The task network is a useful mechanism for depicting inters task dependencies and determining the critical path.
Project task and their dependencies are noted in this diagrammatic form:

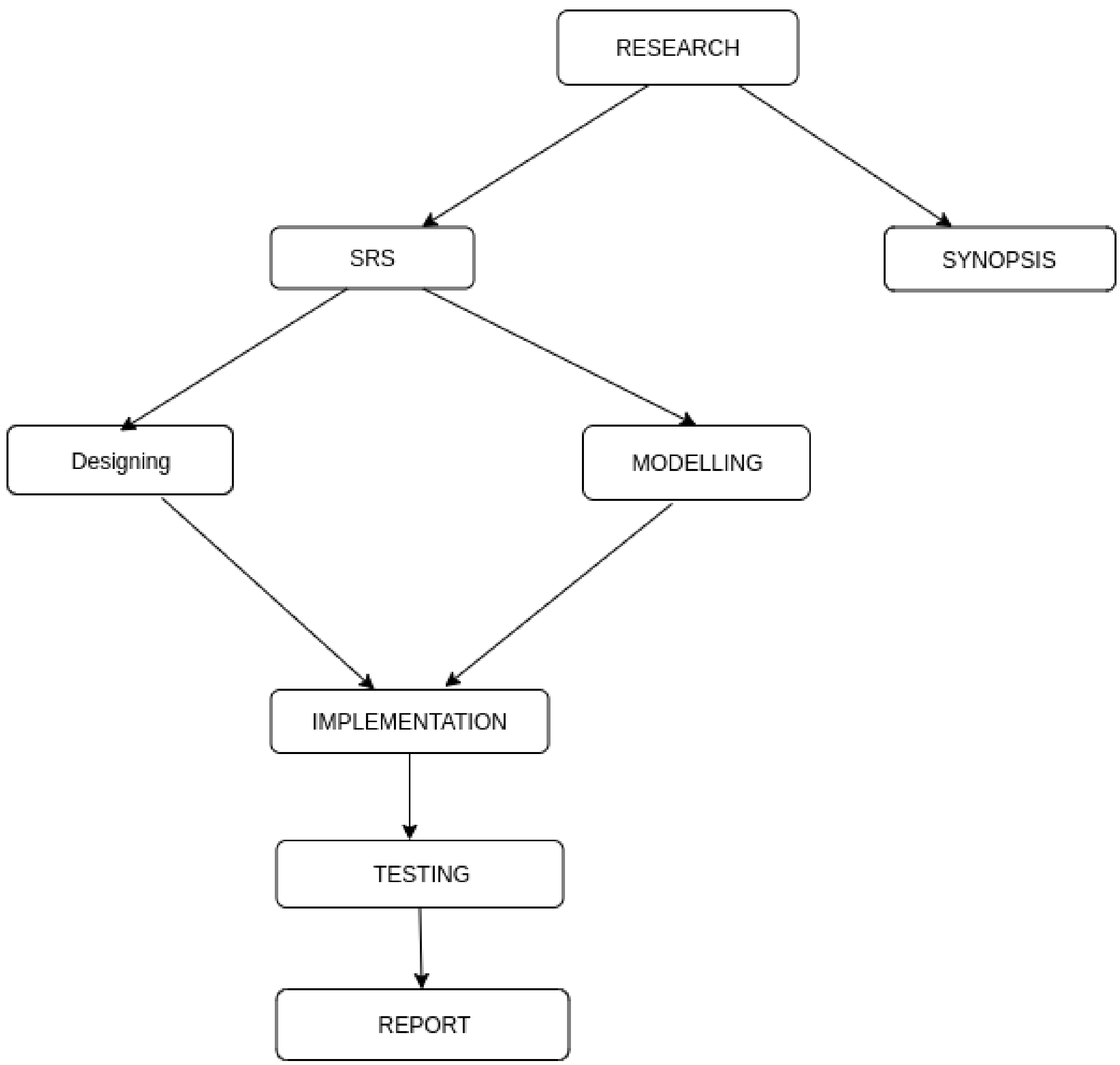

Figure 5.1: Task Network

5.3.3 Timeline Chart

When creating a s/w project schedule, the planner begins with a set of tasks (The work breakdown structure). If automated tools are used, work breakdown is input as a task network. As a consequence, a timeline chart is generated. A timeline chart (also called Gantt chart) can be developed for the entire project. Also, separate charts can be developed for each project function. A timeline chart enables you to determine what works will be conducted at a given point in time.
Gantt chart:
Gantt charts are excellent models for scheduling and for budgeting, and for reporting and presenting and communicating project plans and progress easily and quickly. We have used Gantt chart as a project planning tool. It measures most basic project management functions like Gantt chart for project scheduling task, and doing resource management using resource load charts.

5.4 Team Organization

Proper project team organization is one of the key constraints to project success. If the project has no productive and well-organized team, there's an increased probability that this project will be failed at the very beginning because initially the team is unable to do the project in the right manner. Without right organization of teamwork, people who form the team will fail with performing a number of specific roles and carrying out a variety of group/individual responsibilities. Hence, when you plan for a new project, first you must take care of the best project team organization through team building activities.

5.4.1 Team structure

Name	Skill Set
Vedant Parnaik	Python, Embedded-C
Apoorva Parnate	Python
Tanay Sapre	Python
Kedar Terkhedkar	JavaScript

5.4.2 Management reporting and communication

- Online and offline communication.
- Daily meeting with team members.
- Weekly meeting with Internal Guide.
- Monthly meeting with external Guide.

Chapter 6

Project Implementation

6.1 Overview of Project Modules

There are following major modules as:

A. Sensing module :
The module consists of a interconnected sensor network and sending the data to the cloud database. Various sensors connected to arduino board like N-P-K sensor, pH sensor, soil temperature sensor, soil moisture sensor, surrounding temperature sensor, surrounding humidity sensor collect the data from the soil and environment. This data is organized and later sent to raspberry pi to send to the firebase and other machine learning algorithms.

B. Plant/ Crop suggestion module :
A classification model using Random Forest Algorithm is trained with the dataset containing various features in relation with the needs of crop such as N-P-K values, pH value, soil temperature, soil moisture, surrounding temperature, surrounding humidity and the location area of the field. The input values are taken from the firebase and passed to the model. After the processing the result containing recommendation for the suitable crop to be cultivated are displayed. The system also displays suggestions regarding what fertilizer treatment is required for the crop to have good production on the basis of existing N-P-K availability in the soil.

C. Disease Prediction module :
The pre-trained models and weights of CNN (Convolutional Neural Network) are loaded. The image of leaf uploaded by the user is passed through the CNN. It detects if the plant is healthy or has any sort of disease and the prediction results are displayed. If the crop has some disease; name of the disease, its cause and methodologies to cure that disease are also displayed.

6.2 Tools and Technologies Used

1. Python:
Python is an interpreter, high-level, general-purpose programming language. Created by Guido van Rossum and first released in 1991, Python has a design philosophy that emphasizes code readability, notably using significant whitespace. It provides constructs that enable clear programming on both small and large scales.

2. Firebase Cloud Database:
Firebase is a platform developed by Google for creating mobile and web applications. It was originally an independent company founded in 2011.[1] In 2014, Google acquired the platform[2] and it is now their flagship offering for app development.

3. Internet of Things:
The Internet of things (IoT) describes the network of physical objects—a.k.a. ”things”—that are embedded with sensors, software, and other technologies for the purpose of connecting and exchanging data with other devices and systems over the Internet. In the consumer market, IoT technology is most synonymous with products pertaining to the concept of the ”smart home”, including devices and appliances (such as lighting fixtures, thermostats, home security systems and cameras, and other home appliances) that support one or more common ecosystems, and can be controlled via devices associated with that ecosystem, such as smartphones and smart speakers. The IoT can also be used in healthcare systems.

4. Machine Learning:
Machine learning (ML) is the study of computer algorithms that improve automatically through experience and by the use of data. It is seen as a part of artificial intelligence. Machine learning algorithms build a model based on sample data, known as ”training data”, in order to make predictions or decisions without being explicitly programmed to do so. Machine learning algorithms are used in a wide variety of applications, such as in medicine, email filtering, and computer vision, where it is difficult or unfeasible to develop conventional algorithms to perform the needed tasks.

5. Streamlit:
Streamlit is an open-source Python library that makes it easy to create and share beautiful, custom web apps for machine learning and data science. ****

6.3 Algorithm Details

Algorithm for Sensing module :-
Step 1: Start.
Step 2: Initiate the microcontrollers.
Step 3: Put the sensors in the soil for measurement and collect the data in arduino.
Step 4: Send the collected data from arduino to raspberry pi.
Step 5: Send all the data to the machine learning modules and the significant data to the firebase.
Step 6: Stop

Algorithm for Plant/ Crop Suggestion module :-
Step 1: Start.
Step 2: Train a classification model(Random forest) with a dataset, which contains various features in relation with the needs of a crop.
Step 3: Take input values from the microcontrollers as well as the user.
Step 4: Display the result containing recommendations for the suitable crop.
Step 5: Stop.

Algorithm for Plant leaf image Detection :-
Step 1: start
Step 2: Capture the image of the leaf of the plant.
Step 3: Convert the image into the suitable format by using image processing techniques.
Step 4: if (image is clear and usable):
Else:
error message('Try again')
Step 5: Stop

Algorithm for Disease Prediction :-
Step 1: Start
Step 2: Load the pre-trained models and weights of CNN (Convolutional Neural Network)
Step 3: Pass the image through the CNN and capture the prediction results.
Step 4: Display the results.
Step 5: Stop

Chapter 7

Software Testing

Testing is an investigation conducted to provide stakeholders with information about the quality of the product or service under test. Software testing also provides an objective, independent view of the software to allow the business to appreciate and understand the risks of software implementation. Test techniques include, but are not limited to, the process of executing a program or application with the intent of finding software bugs. It describes software test environment for testing, identifies the tests to be performed, and provides schedules for test activates.

7.1 Type of Testing

Unit testing technique was used in this project as modules. Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. It is the testing of individual software units of the application .It is done after the completion of an individual unit before integration. This is a structural testing, that relies on knowledge of its construction. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications.

7.2 Test cases and Test Results

Test Case ID	TC001
Test Case Description	Establishing communication with the hardware system and cloud database.
Steps and Input	Establish Link.
Expected Output	It should establish a proper link.
Actual Output	Link is established.
Status	Pass.

Test Case ID	TC002
Test Case Description	Reading values from different sensors.
Steps and Input	It should read correct values from the environment.
Expected Output	Correct analysis of surrounding.
Actual Output	Surrounding is analyzed correctly.
Status	Pass.

Test Case ID	TC003
Test Case Description	Establish a communication between firebase and the ML module.
Steps and Input	Build a connection.
Expected Output	Connection should be established.
Actual Output	Connection is established.
Status	Pass.

Test Case ID	TC004
Test Case Description	Login with correct credentials.
Steps and Input	Authenticate with database.
Expected Output	User should be logged in.
Actual Output	Login Successful.
Status	Pass.

Test Case ID	TC005
Test Case Description	Login with incorrect credentials.
Steps and Input	Authenticate with database.
Expected Output	User should not be logged in.
Actual Output	Login unsuccessful.
Status	Pass.

Test Case ID	TC006
Test Case Description	Write data to cloud and crosscheck the output of Case 3.
Steps and Input	Perform write operation to the cloud.
Expected Output	Correct information should be written to cloud.
Actual Output	Write operation successful.
Status	Pass.

Test Case ID	TC007
Test Case Description	Read data to cloud and crosscheck the output of Case 3.
Steps and Input	Perform read operation to the cloud.
Expected Output	Correct information should be read.
Actual Output	Read operation successful.
Status	Pass.

Chapter 8

Results

8.1 Outcomes

The proposed system successfully senses accurate values from the soil and surrounding. It successfully suggest a suitable plant/crop for the given soil parameters and also gives suggestions for fertilizers if required. It predicts if any diseases have affected the plant/crop also gives remedial measures.

8.2 Screenshots

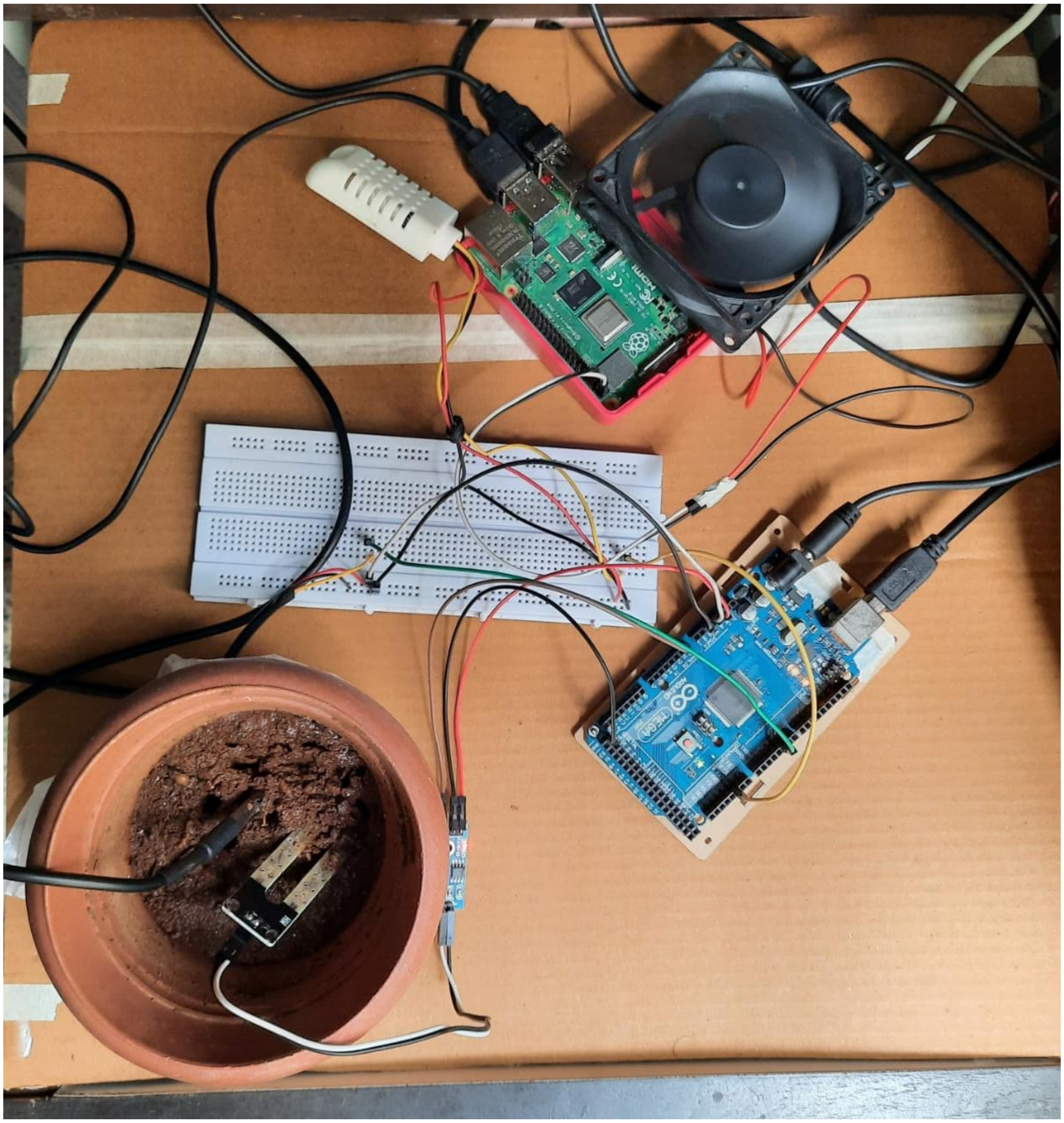

Figure 8.1: Hardware Setup

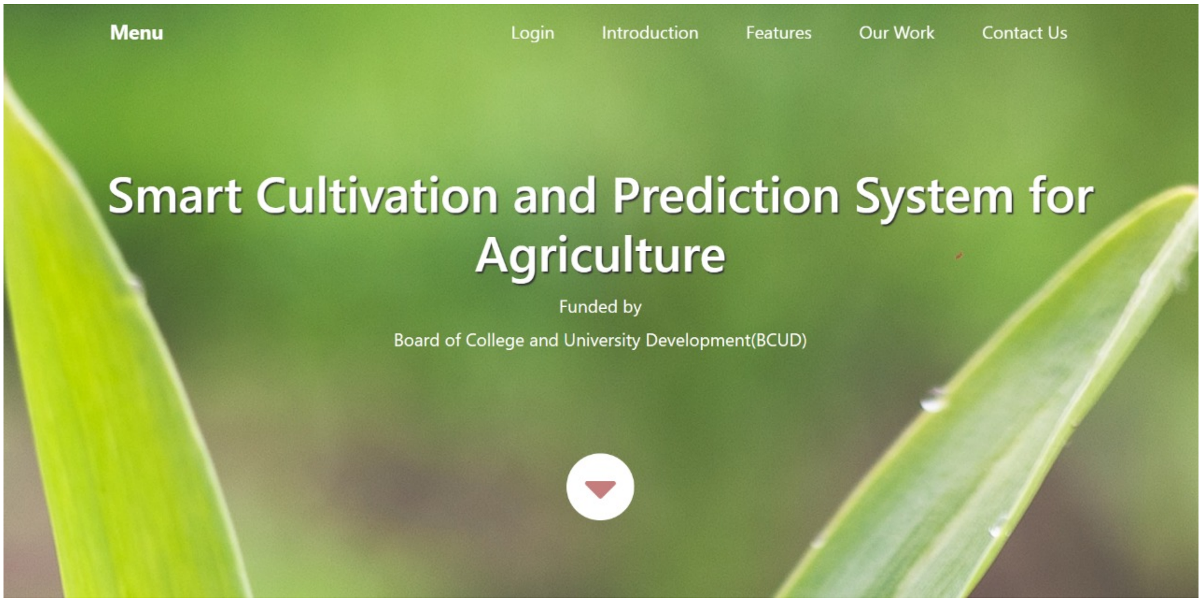

Figure 8.2: Front Page

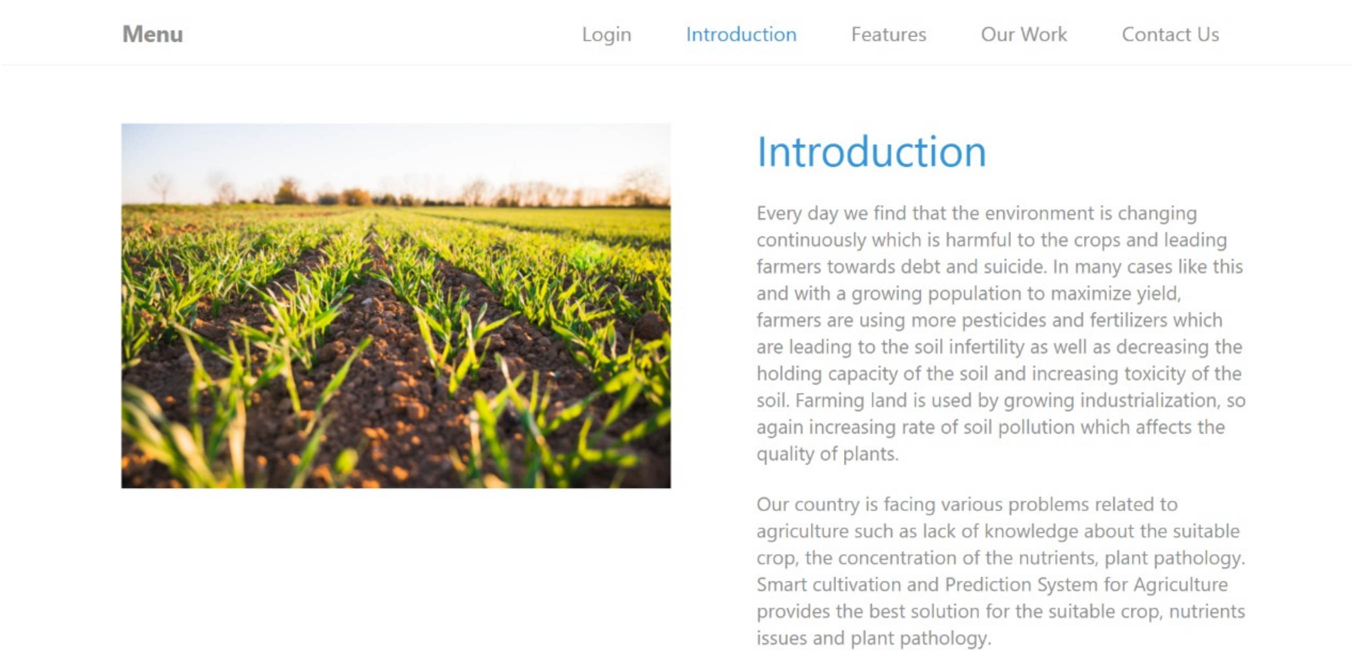

Figure 8.3: Introduction

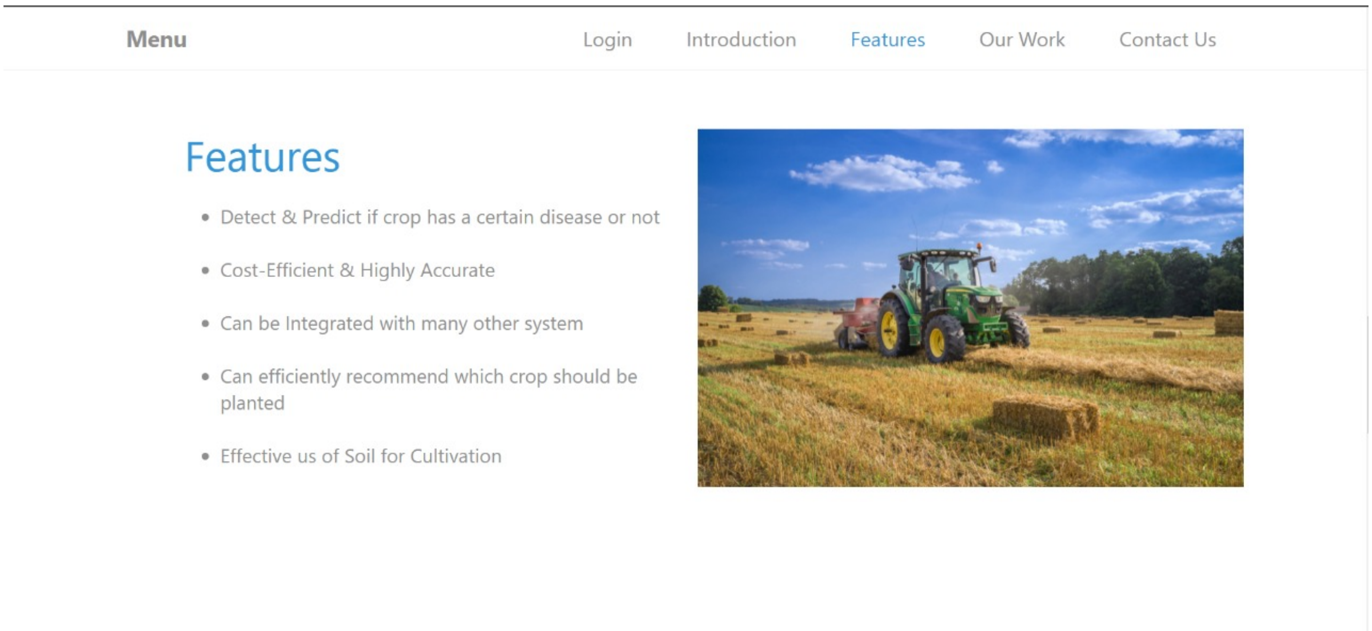

Figure 8.4: Features

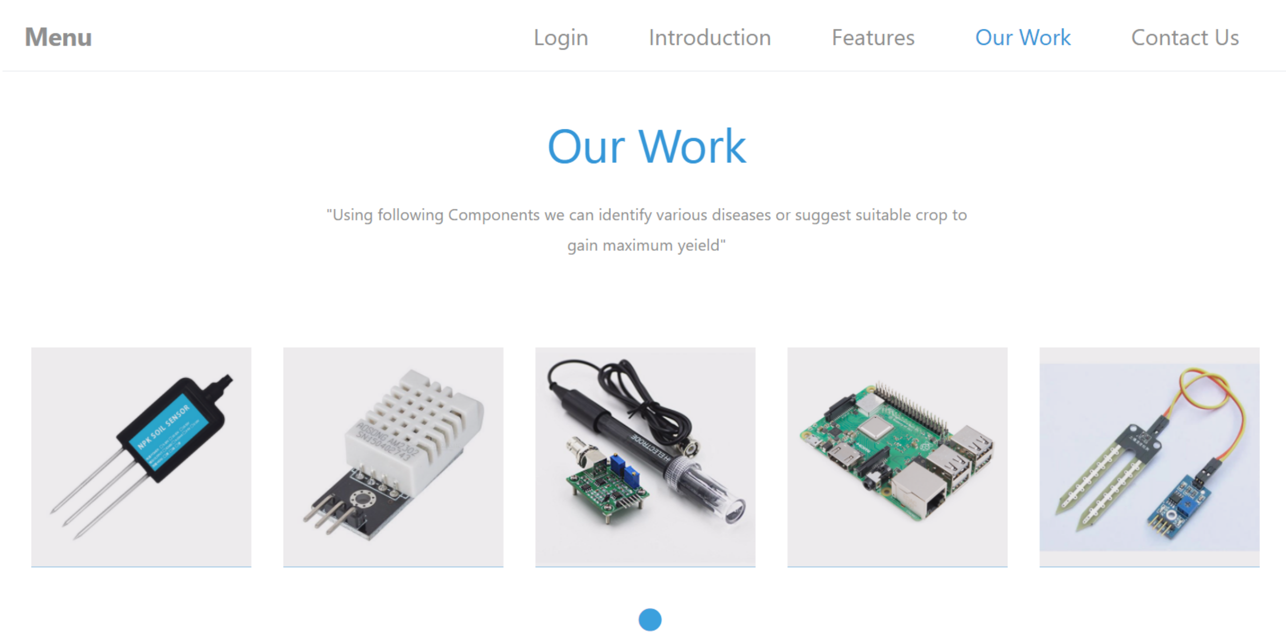

Figure 8.5: Our Work

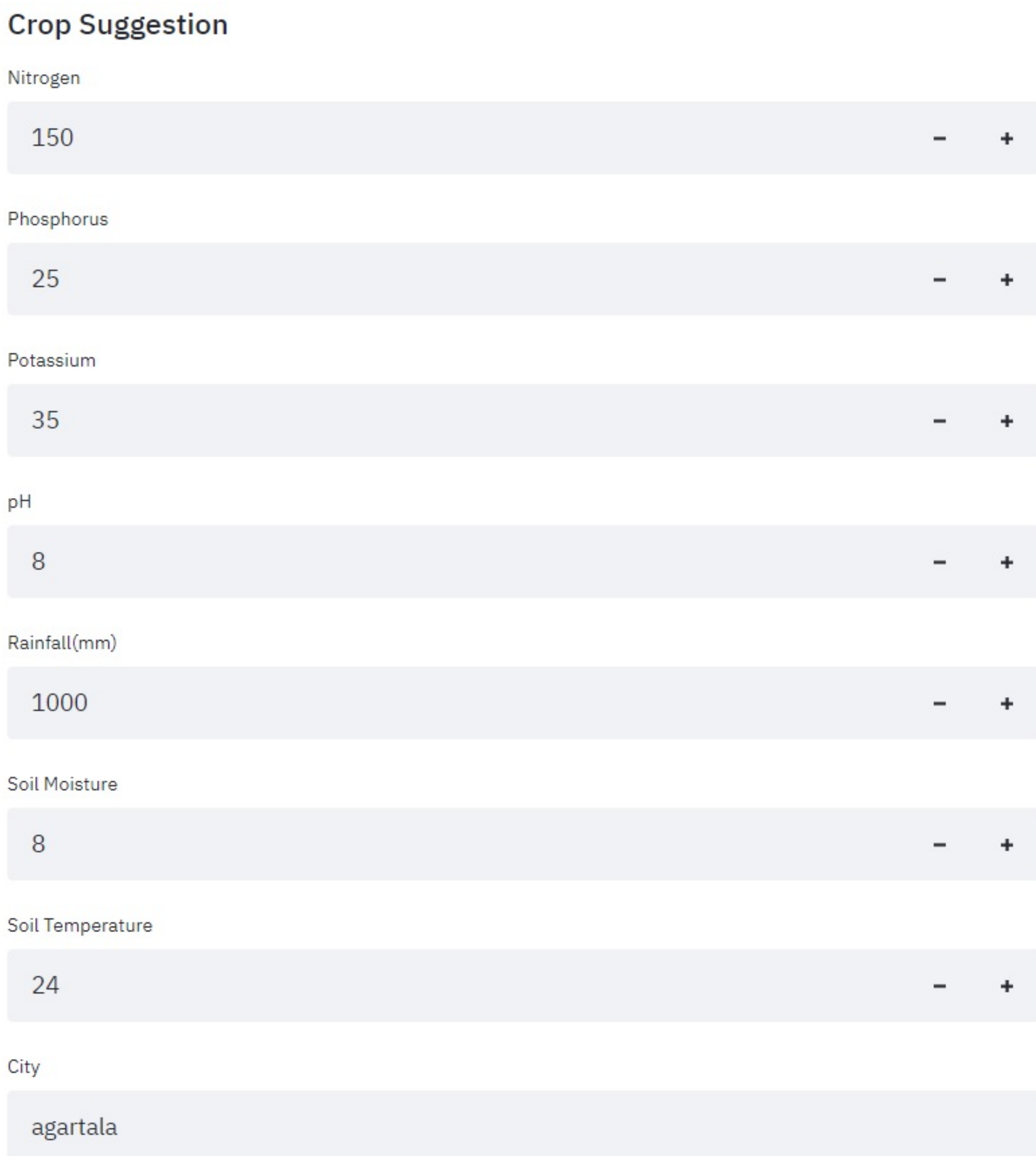

Figure 8.6: Crop Suggestion

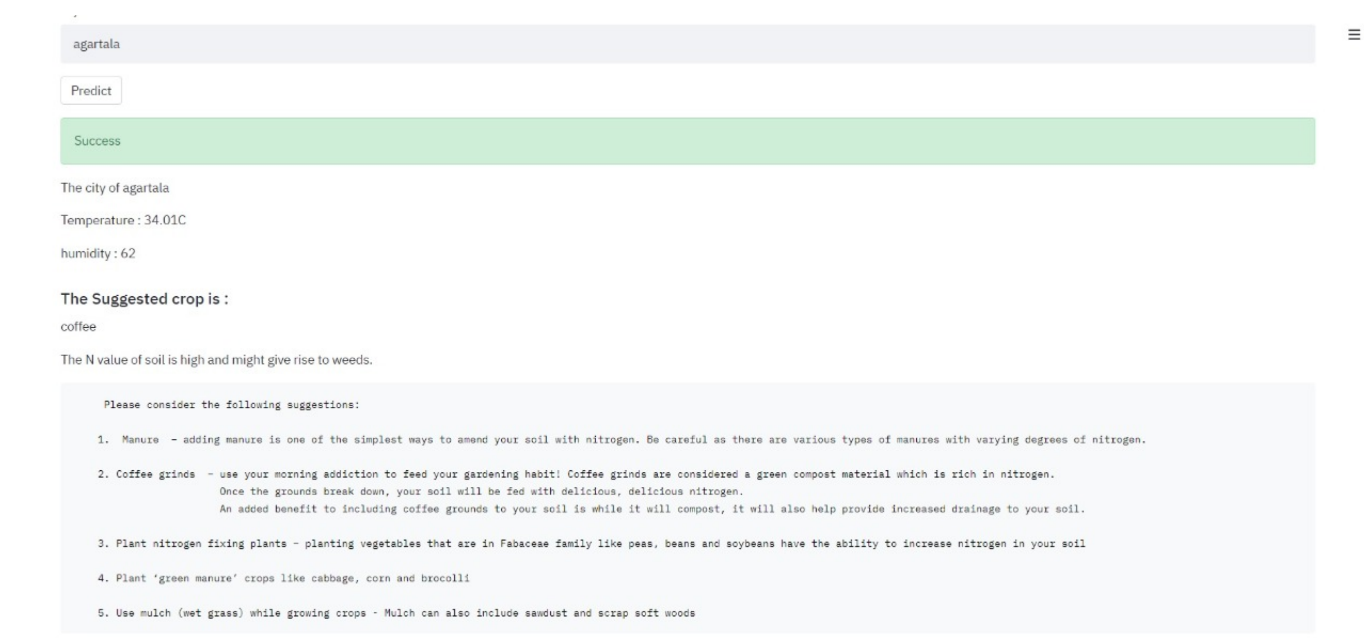

Figure 8.7: Fertilizer Suggestion based on NPK values

Figure 8.8: Disease Identification

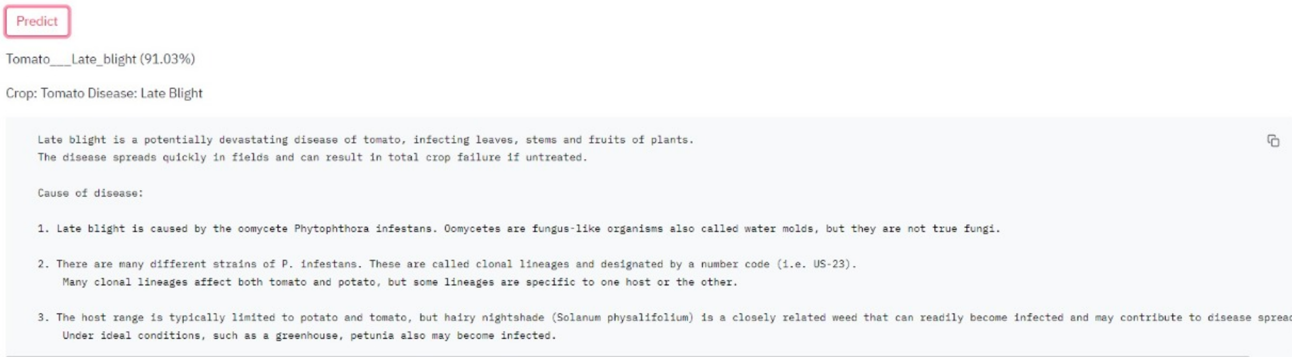

Figure 8.9: Details about cause of disease and ways to cure it

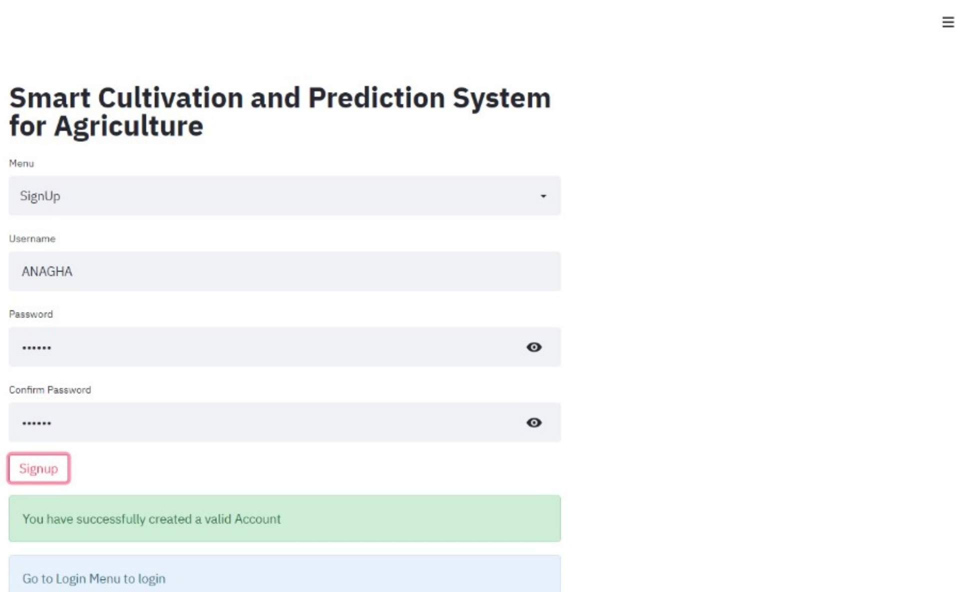

Figure 8.10: Sign Up

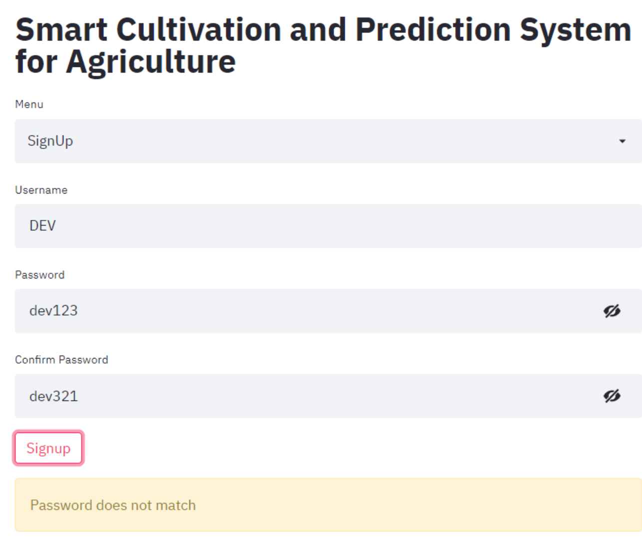

Figure 8.11: Unsuccessful sign up for wrong credentials

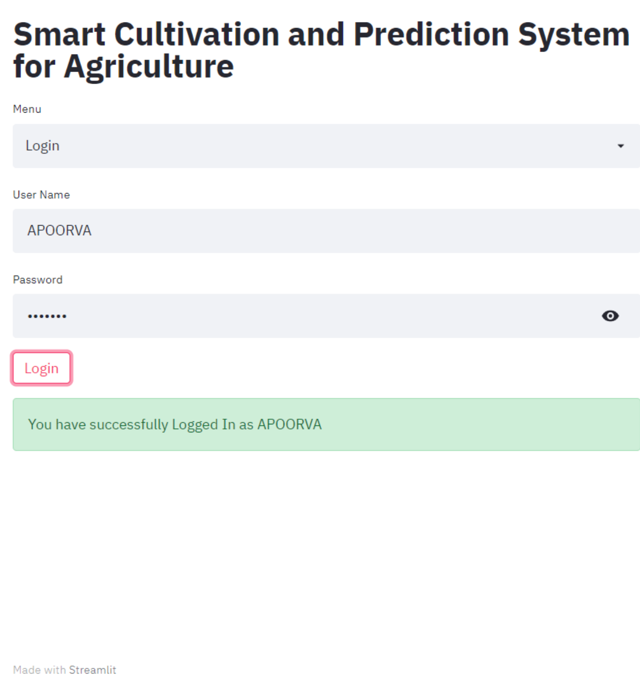

Figure 8.12: Log in

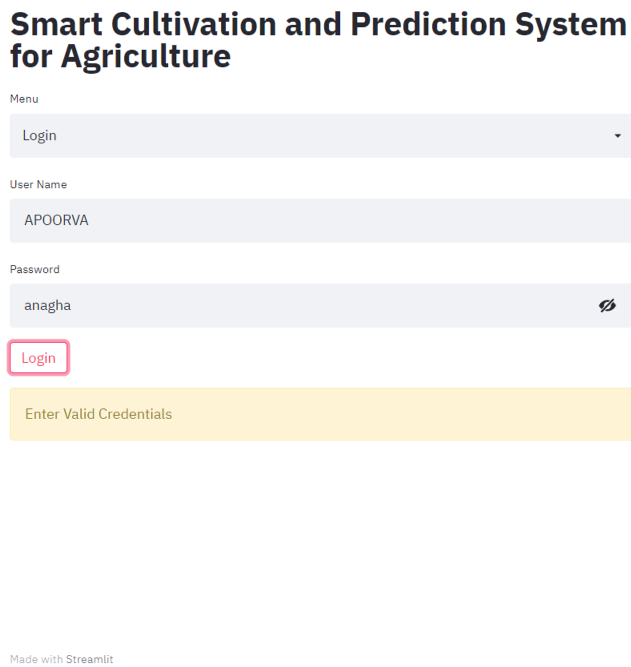

Figure 8.13: Unsuccessful log in for invalid credentials

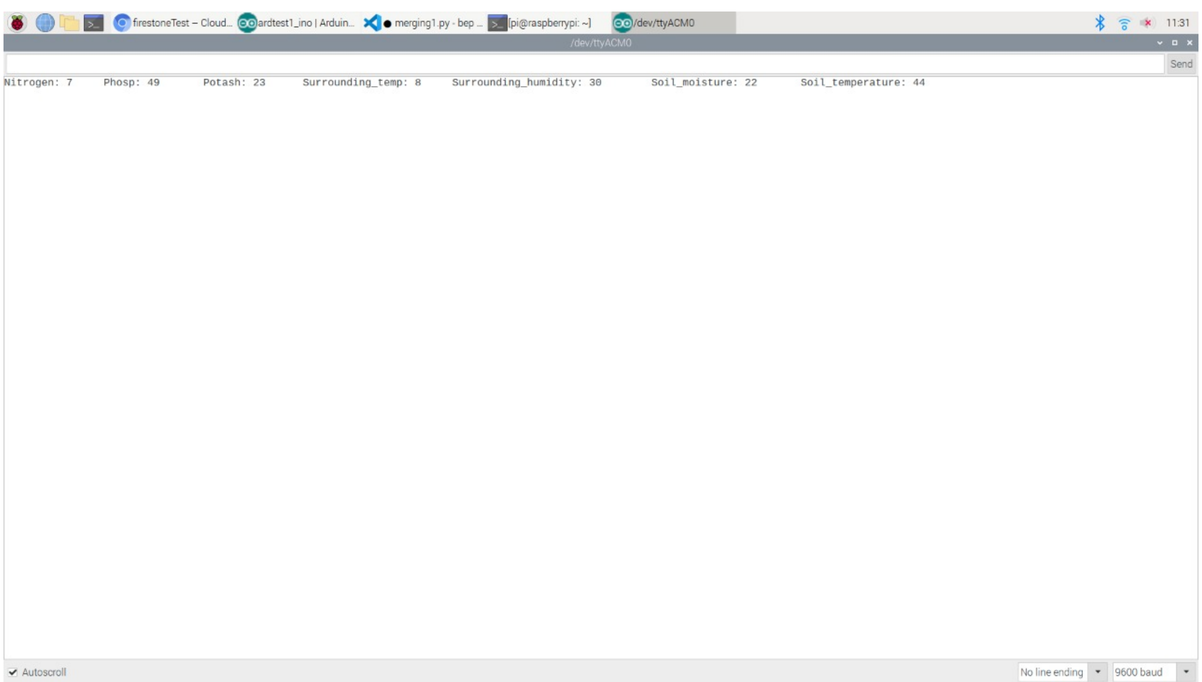

Figure 8.14: Arduino Comboard

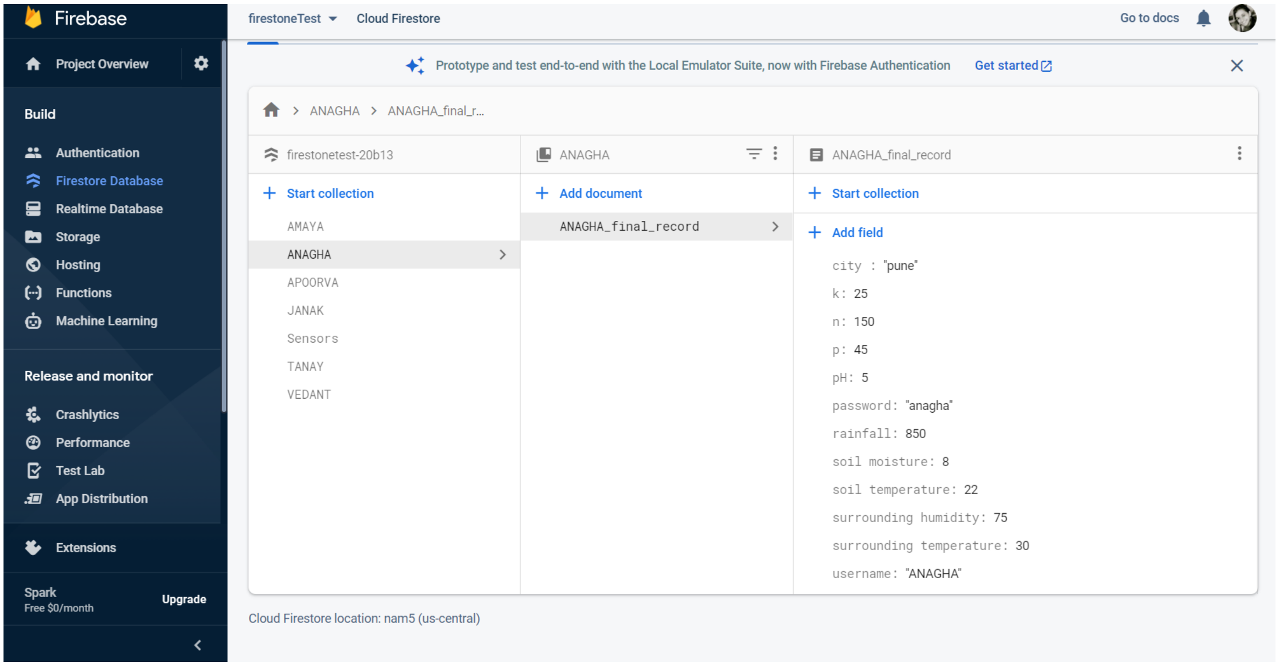

Figure 8.15: Firestore Record example 1

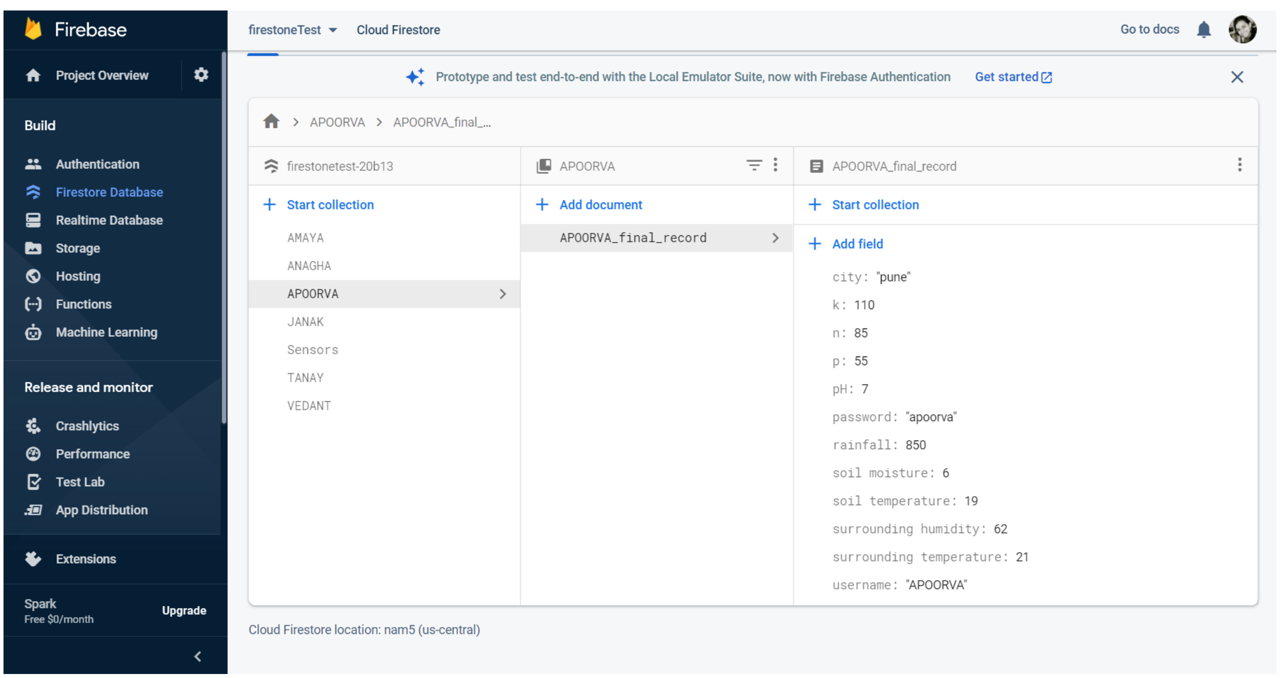

Figure 8.16: Firestore Record example 2

Chapter 9

Conclusions

9.1 Conclusions

We will be able to detect macro nutrient levels in the soil. We can predict suitable crops/plants according to the nutrient levels present in the soil. We will be able to predict if any diseases may or may not affect the crop as well as show suggestions to prevent the diseases.

9.2 Future Work

Creating a widely usable and highly portable hardware module to detect various components of soil, capable of doing all the work that can be done using multiple and multi-level systems currently. Mapping the centralized network to a GPS network connected through LoRa to cover larger agricultural fields. The whole unit could be mounted on an autonomous mini 4WD off-road landrover capable of doing all the manual testing work all by itself. An all inclusive, interactive software which would be available in various regional languages, which will track the progress of the crop, predict diseases, and give highly efficient recommendations to boost crop production.

9.3 Applications

1. Farming.
2. Precision Agriculture.
3. Effective use of soil for cultivation.

APPENDIX A

NP HARD PROBLEM:- NP-hard,is a class of problems that are informally,least as hard as the hardest problem in NP”.More precisely,a problem H is NP-hard when every problem L in NP can be reduced in polynomial time.As a consequence finding a polynomial algorithm to solve any NP-hard problem would give polynomial algorithms for all the problem is NP,which is unlikely as many of them are considered as hard.

NP COMPLETE:- NP-complete in computational complexity theory, a decision problem is NP-complete when it is both in NP and NP-hard.The set of NP-complete problems is often denoted by NP-C or NPC. The abbreviation NP refers to non deterministic polynomial time. In computational complexity theory, a problem is NP-complete when it can be solved by a restricted class of brute force search algorithms and it can be used to simulate any other problem with a similar algorithm. More precisely, each input to the problem should be associated with a set of solutions of polynomial length, whose validity can be tested quickly), such that the output for any input is ”yes” if the solution set is non-empty and ”no” if it is empty. A problem is said to be NP-hard if everything in NP can be transformed in polynomial time into it, and a problem is NP-complete if it is both in NP and NP-hard. The proposed project is feasible.We can implement it within time limit and use actuators,sensors and hardware components to do so.So it is guaranteed to work according to our expectations once analyzed surrounding conditions and implemented.

APPENDIX B

Publication Details:

Cited as: Vedant Parnaik, Kedar Terkhedkar, Apoorva Parnate, Tanay Sapre, Pradnya Mehta, "Survey Paper on Smart Cultivation and Prediction System for Agriculture", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 7, Issue 3, pp.554-557, May-June-2021.

- Published in : International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT) Volume 7 - Issue 3 - May-June 2021
- UGC Journal No : 64718
- Impact Factor of the Publication : 6.135
- Date of Publication : 2021-06-30
- License : This work is licensed under a Creative Commons Attribution 4.0 International License.
- Manuscript Number : CSEIT21739
- Publisher : Technoscience Academy
- https://ijsrcseit.com/CSEIT21739

Figure 9.1: Paper Publication in IJSRCSEIT

Continuity	70%
Text structure	80%
References	90%
Understanding and Illustrations	91%
Explanatory power	85%
Detailing	75%
Relevance and practical advice	92%
Comments	
Your submitted paper is Accepted **Unique content 87% complete below process.**	

Figure 9.2: Review Report

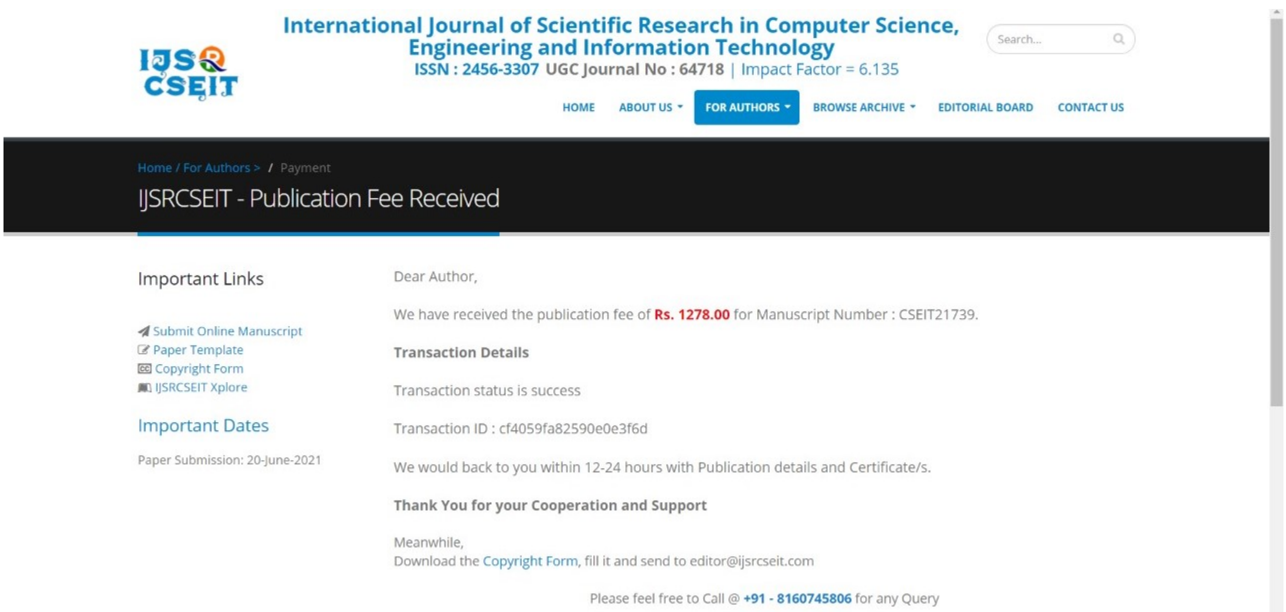

Figure 9.3: Publication Transaction Details

APPENDIX C

Plagarism Report of Project Report:

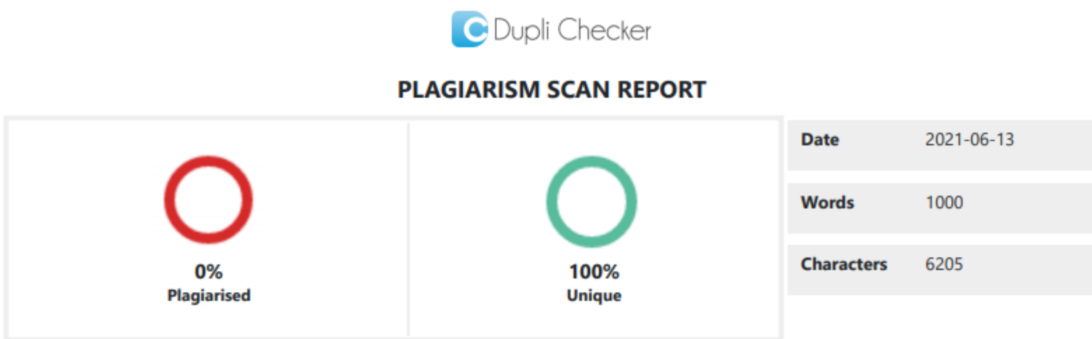

Figure 9.4: Plagiarism Project Report 1

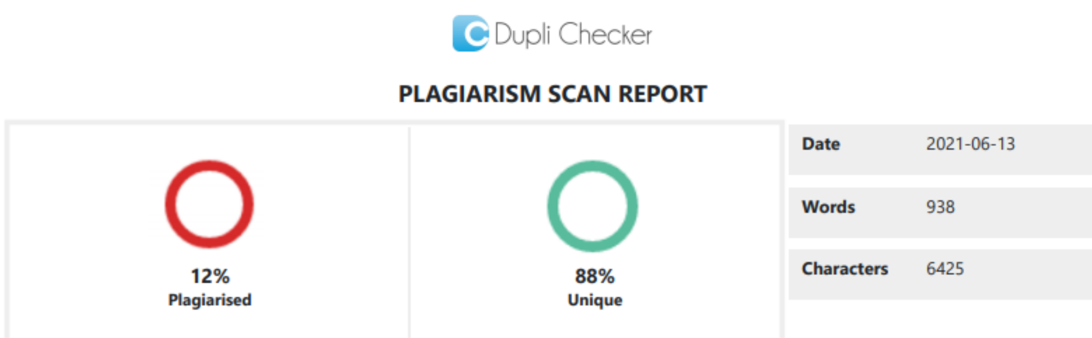

Figure 9.5: Plagiarism Project Report 2

Bibliography

[1] MS Prasad Babu, B Srinivasa Rao, et al. Leaves recognition using back propagation neural network-advice for pest and disease control on crops. *IndiaKisan. Net: Expert Advisory System*, 2007.

[2] Dan Claudiu Ciresan, Ueli Meier, Jonathan Masci, Luca Maria Gambardella, and Jürgen Schmidhuber. Flexible, high performance convolutional neural networks for image classification. In *Twenty-second international joint conference on artificial intelligence*, 2011.

[3] Douglas M Hawkins. The problem of overfitting. *Journal of chemical information and computer sciences*, 44(1):1–12, 2004.

[4] Smita Kapse, Samanta Kale, Sharanya Bhongade, Shivani Sangamnerkar, and Yashada Gotmare. Iot enable soil testing npk nutrient detection. 9. 2020.

[5] Aditya Khamparia, Gurinder Saini, Deepak Gupta, Ashish Khanna, Shrasti Tiwari, and V.H.C. Albuquerque. Seasonal crops disease prediction and classification using deep convolutional encoder network. *Circuits, Systems, and Signal Processing*, 39, 02 2020.

[6] Rakesh Kumar, MP Singh, Prabhat Kumar, and JP Singh. Crop selection method to maximize crop yield rate using machine learning technique. In *2015 international conference on smart technologies and management for computing, communication, controls, energy and materials (ICSTM)*, pages 138–145. IEEE, 2015.

[7] A-K Mahlein, T Rumpf, P Welke, H-W Dehne, L Plümer, U Steiner, and

E-C Oerke. Development of spectral indices for detecting and identifying plant diseases. *Remote Sensing of Environment*, 128:21–30, 2013.

[8] Anne-Katrin Mahlein, Erich-Christian Oerke, Ulrike Steiner, and Heinz-Wilhelm Dehne. Recent advances in sensing plant diseases for precision crop protection. *European Journal of Plant Pathology*, 133(1):197–209, 2012.

[9] M. Masrie, A. Z. M. Rosli, R. Sam, Z. Janin, and M. K. Nordin. Integrated optical sensor for npk nutrient of soil detection. In *2018 IEEE 5th International Conference on Smart Instrumentation, Measurement and Application (ICSIMA)*, pages 1–4, 2018.

[10] Marianah Masrie, Mohamad Syamim Aizuddin Rosman, Rosidah Sam, and Zuriati Janin. Detection of nitrogen, phosphorus, and potassium (npk) nutrients of soil using optical transducer. In *2017 IEEE 4th International Conference on Smart Instrumentation, Measurement and Application (ICSIMA)*, pages 1–4. IEEE, 2017.

[11] Sally A Miller, Fen D Beed, and Carrie Lapaire Harmon. Plant disease diagnostic capabilities and networks. *Annual review of phytopathology*, 47:15–38, 2009.

[12] Subhadra Mishra, Debahuti Mishra, and Gour Hari Santra. Applications of machine learning techniques in agricultural crop production: a review paper. *Indian Journal of Science and Technology*, 9(38):1–14, 2016.

[13] Z. Muhammad, M. A. A. M. Hafez, N. A. M. Leh, Z. M. Yusoff, and S. A. Hamid. Smart agriculture using internet of things with raspberry pi. In *2020 10th IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*, pages 85–90, 2020.

[14] Pranjali Padol and Anjali Yadav. Svm classifier based grape leaf disease detection. pages 175–179, 06 2016.

[15] Pooja Pawar, Varsha Turkar, and Pravin Patil. Cucumber disease detection using artificial neural network. pages 1–5, 08 2016.

[16] Meena R., G.P. Saraswathy, G. Ramalakshmi, K.H. Mangaleswari, and T. Kaviya. Detection of leaf diseases and classification using digital image processing. pages 1–4, 03 2017.

[17] Krishnaswamy Rangarajan Aravind and Purushothaman Raja. Automated disease classification in (selected) agricultural crops using transfer learning. *Automatika*, 61(2):260–272, 2020.

[18] Angie K Reyes, Juan C Caicedo, and Jorge E Camargo. Fine-tuning deep convolutional networks for plant recognition. *CLEF (Working Notes)*, 1391:467–475, 2015.

[19] S. F. Syed, D. Varghese, and A. K. Tripathy. Remote sensor networks for chilli crop disease prediction using thermal image processing techniques. In *2020 3rd International Conference on Communication System, Computing and IT Applications (CSCITA)*, pages 38–43, 2020.

[20] Shafaque Fatma Syed, Ditty Varghese, and Amiya Kumar Tripathy. Remote sensor networks for chilli crop disease prediction using thermal image processing techniques. In *2020 3rd International Conference on Communication System, Computing and IT Applications (CSCITA)*, pages 38–43. IEEE, 2020.