# Decision Support System

# In

# Crop Selection

**A PROJECT REPORT**

***Submitted by***


**SAGAR YADAV (400170)**
**MAYUR WARPE (400168)**
**VARAD LAD (400136)**
**CHAITANYA TUPSAMUDRE (400167)**
**OMKAR PATIL (CSE)**


***In partial fulfillment for the award of the degree***

***Of***

**BACHELOR OF TECHNOLOGY**

**IN**

**INFORMATION TECHNOLOGY**

Guided by

**Dr. Parminder Kaur**


Information Technology


**MGM's Jawaharlal Nehru Engineering College, Aurangabad**


YEAR 2020- 2021

# CERTIFICATE

This is to certify that the project report

**"DECISION SUPPORT SYSTEM IN CROP SELECTION"**

Submitted by

| | |
|---|---|
| **SAGAR YADAV** | **400170** |
| **MAYUR WARPE** | **400168** |
| **VARAD LAD** | **400136** |
| **CHAITANYA TUPSAMUDRE** | **400167** |
| **OMKAR PATIL** | **CSE** |

is a bonafide work carried out by them under the supervision of Dr. Parminder Kaur and it is approved for the partial fulfillment for the award of Bachelor of Technology (Information Technology), from Dr. Babasaheb Ambedkar Technological University, Lonere.

Date:

Dr. Parminder Kaur
Guide
**Dept. of Information Technology**

Dr. S. C. Tamane
Head of Department
**Dept. of Information Technology**

**CONTENTS**

## List of Abbreviations

- ML : Machine Learning
- UI : User Interface
- MSE : Mean Square Error
- AI : Artificial Intelligence
- BSCI: Bare Soil Composite Image
- API : Application Programming Interface

# LIST OF FIGURES

# CHAPTER 1
# SYNOPSIS

## 1.1 Project Title

Decision Support System in Crop Selection.

## 1.2 Project Option

Internal Project

## 1.3 Internal Guide

Dr. Parminder Kaur

## 1.4 Technical Keywords

1. Remote Sensing Data
2. Satellite Image
3. Machine Learning
4. Artificial Intelligence
5. (MSE)Mean Square Error

## 1.5 Problem Statement

To design and develop a system which can suggest which crop farmers should sow, so that they will get maximum profit according to their soil properties and various inputs.

## 1.6 Abstract

Nowadays the usage of Remote Sensing techniques are too vast and many Earth observations had been done by Remote Sensing. Remote sensing has been widely used for mapping of soil, its classification and also its texture description. The innovation of this work is to use the LANDSAT image data to assess land information and soil classification in the most coming up area of Aurangabad district was used for soil classification and propose the possible fertilization for this study area. Landsat image is classified by minimum distance classification algorithm and according to the reflectance characteristics of the surface material. From the classified data we can find out the best crops for the best soil using color of the soil. It is proved that, within limitations the classification algorithms and have an important influence on the classification result.

In consideration with the ML algorithms, Image classification is a supervised learning problem: define a set of target classes (objects to identify in images), and train a model to recognize them using labeled example photos. With the help of this classification algorithms, we are able to classify the soil type by training our model with thousands of different types of soil images. We created the model which takes input as image of soil and gives the output as soil type.

Climate data is also one of the most important factor for our project, when we talk about climate data, the type of data can vary significantly. We have very little actual observations at the scale of climate and usually not covering a large area. In consideration to our final model of predicting crop, the climate data like rainfall, temperature, etc. are very most important attributes of dataset which we are using. Our model analysis the data from the dataset and gives the suitable crop which farmer can be sow into his field.

## 1.7 Goals and Objectives

- Goals:
    - To deliver a system which takes the input as the location of the farmer and according to that location it will give the suggestion to farmer which crop should they sow.
    - To build a system which extract soil features from satellite image.
    - Also, a model which can give the soil type using image classification.
    - Gathering the data for making datasets, like climate data, crop data, etc.
    - Analysis of the data and filtering it.

- Extract soil properties from home-based soil testing methods.
- Predict crop from the model based on many fields as input.

- Objectives:

  - To use remote sensing data to extract the features of soil.
  - To use physical data of soil properties rather than depending completely on the Remote sensing data.
  - Gathering a large number of images of different soil types for soil detection model.
  - Deployment of models into android and web application.
  - Making chatbot for user to simplify the methods of soil testing at home.
  - To create a system which can be beneficial in the agriculture.

# CHAPTER 2
# TECHNICAL KEYWORDS

## 2.1 Area of Project

Main objective of our project is to extract the features of soil or extract the soil properties from the remote sensing data (Satellite image). After getting soil type, gathering the different datasets which considers the climate data, data of different crops and their market values. In this project, the model which developed are helpful to the farmers in different kind of areas. It helps to identifying his field soil, the climate around his field location. The most important feature of this project is that it is helpful to farmers regarding which crop he/she can sow into his field so he/she can get more benefits.

## 2.2 Technical Keywords

1. Remote Sensing Data
2. Satellite Image
3. Machine Learning
4. (MSE)Mean Square Error
5. Classification Algorithms
6. Data preprocessing

### 1) Remote Sensing Data:

Remote sensing is the process of detecting and monitoring the physical characteristics of an area by measuring its reflected and emitted radiation at a distance (typically from satellite or aircraft). Special cameras collect remotely sensed images, which help researchers "sense" things about the Earth. Some examples are,

- Cameras on satellites and airplanes take images of large areas on the Earth's surface, allowing us to see much more than we can see when standing on the ground.
- Sonar systems on ships can be used to create images of the ocean floor without needing to travel to the bottom of the ocean.
- Cameras on satellites can be used to make images of temperature changes in the oceans.

Some specific uses of remotely sensed images of the Earth include:

- Large forest fires can be mapped from space, allowing rangers to see a much larger area than from the ground.
- Tracking clouds to help predict the weather or watching erupting volcanoes, and help watching for dust storms.
- Tracking the growth of a city and changes in farmland or forests over several years or decades.
- Discovery and mapping of the rugged topography of the ocean floor (e.g., huge mountain ranges, deep canyons, and the "magnetic striping" on the ocean floor).

## 2) Satellite images:

Satellite images are one of the most powerful and important tools used by the meteorologist. They are essentially the eyes in the sky. These images reassure forecasters to the behavior of the atmosphere as they give a clear, concise, and accurate representation of how events are unfolding. Forecasting the weather and conducting research would be extremely difficult without satellites. Data taken at stations around the country is limited in its representations of atmospheric motion. It is still possible to get a good analysis from the data, but because the stations are separated by hundreds of miles significant features can be missed. Satellite images aid in showing what cannot be measured or seen. In addition, the satellite images are viewed as truth. There is no chance for error. Satellite images provide data that can be interpreted "first-hand".

Satellites images give a good representation of what is happening at every point in the world, especially over oceans where large gaps in data occur. Data can only be taken at certain points around the world, though, without this data, forecasting would be just as difficult as not having satellites. It is essential to have both. Having the two together gives a much better understanding as to how the atmosphere is behaving and greatly improves forecasting accuracy.

## 3) Machine Learning:

Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it to learn for themselves.

Today, examples of machine learning are all around us. Digital assistants search the web and play music in response to our voice commands. Websites recommend products and movies and songs based on what we bought, watched, or listened to before. Robots vacuum our floors while we do . . . something better with our time. Spam detectors stop

unwanted emails from reaching our inboxes. Medical image analysis systems help doctors spot tumors they might have missed. And the first self-driving cars are hitting the road.

## 4) Mean square error:

In statistics, the mean squared error (MSE) of an estimator (of a procedure for estimating an unobserved quantity) measures the average of the squares of the errors — that is, the average squared difference between the estimated values and what is estimated. MSE is a risk function, corresponding to the expected value of the squared error loss. The fact that MSE is almost always strictly positive (and not zero) is because of randomness or because the estimator does not account for information that could produce a more accurate estimate.

Let's define a mathematical equation that will give us the mean squared error for all our points.

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \tilde{y}_i)^2$$

General formula for mean squared error.

Let's analyze what this equation actually means.

- In mathematics, the character that looks like weird E is called summation (Greek sigma). It is the sum of a sequence of numbers, from i=1 to n. Let's imagine this like an array of points, where we go through all the points, from the first (i=1) to the last (i=n).
- For each point, we take the y-coordinate of the point, and the y'-coordinate. The y-coordinate is our purple dot. The y' point sits on the line we created. We subtract the y-coordinate value from the y'-coordinate value, and calculate the square of the result.
- third part is to take the sum of all the $(y\text{-}y')^2$ values, and divide it by n, which will give the mean.

## 5) Classification algorithms:

Classification is the process of recognizing, understanding, and grouping ideas and objects into preset categories or "sub-populations." Using pre-categorized training datasets, machine learning programs use a variety of algorithms to classify future datasets into categories.

Classification algorithms in machine learning use input training data to predict the likelihood that subsequent data will fall into one of the predetermined categories.

One of the most common uses of classification is filtering emails into “spam” or “non-spam.”

In short, classification is a form of “pattern recognition,” with classification algorithms applied to the training data to find the same pattern (similar words or sentiments, number sequences, etc.) in future sets of data.

### Types of classification algorithms:

- Logistic Regression
- Naive Bayes Classifier
- K-Nearest Neighbors
- Decision tree
- Support vector Machine

## 6) Data preprocessing:

Data preprocessing is a process of preparing the raw data and making it suitable for a machine learning model. It is the first and crucial step while creating a machine learning model.

When creating a machine learning project, it is not always a case that we come across the clean and formatted data. And while doing any operation with data, it is mandatory to clean it and put in a formatted way. So, for this, we use data preprocessing task.

### Why do we need data processing?

A real-world data generally contains noises, missing values, and maybe in an unusable format which cannot be directly used for machine learning models. Data preprocessing is required tasks for cleaning the data and making it suitable for a machine learning model which also increases the accuracy and efficiency of a machine learning model.

It involves below steps:

- Getting the datasets
- Importing libraries
- Importing datasets
- Finding Missing Data
- Encoding Categorial Data
- Splitting dataset into training and test set
- Feature scaling

# CHAPTER 3
# INTRODUCTION

## 3.1 Project Idea

In our project we are making the system for farmers so they can yield their production of crops in very digitization manner. Farmer is the most important person for country, who gives the food for us but due to uncertainty of climate, soil properties and good fertilization, he could face less production in the field. In our system we are going to make a system which can predict the suitable crop for farmers which should be sown into the farm, so that they can achieve maximum production and gain maximum profit.

So, our main objective of our project is to extract the features of soil or extract the soil properties from the remote sensing data (Satellite image). After getting soil type, gathering the different datasets which considers the climate data, data of different crops and their market values. In this project, the model which developed are helpful to the farmers in different kind of areas. It helps to identifying his field soil, the climate around his field location. The most important feature of this project is that it is helpful to farmers regarding which crop he/she can sow into his field so he/she can get more benefits.

## 3.2 Motivation of Project

Agriculture in India is undergoing a structural change leading to a crisis situation. The rate of growth of agricultural output is gradually declining in the recent years. The relative contribution of agriculture to the GDP has been declining over time steadily. The performance of agriculture by crop categories also clearly indicates the slowing down process of agriculture in India.

The Majority of farmers are not able to do soil testing and due to this they are not getting what actually they deserve. With the increase in technology the quality of satellite image is getting clearer and lots of people are working on it. We can extract the features of the soil using the reflection of the soil and its texture. These soil type getting model can help to us to enhance their accuracy of crop prediction model. Hence, we propose a system which will take location and other parameters from the farmer and suggest the different crops to the farmer according to his soil properties by using remote sensing data.

## 3.3 Literature Survey

- Remote sensing applications, P. S. Roy, R. S. Dwivedi, D. Vijayen, "national Remote Sensing Center, ISRO".

    - **ABSTRACT**: By using Remote sensing technique, we can identify the different kinds of soils with the help of satellite data and then get confirmation of this data from laboratory to verify satellite data is correct or not. For that we can test the soil sample of same area and cross check it with satellite data. The different soils observed within an area are put into a limited number of groups. Each group of soils is represented by a Pedon (The smallest unit or volume of soil that contains all the soil horizons of a particular soil type). Such groups of soils may occur on the landscape in a repeated manner and are delineated on a map as "map unit". The delineated units on a map will be given a specific symbol, colour and name. Soil maps contain several mapping units along with a legend that give description of each soil unit. The soil maps are prepared on different scales varying from 1:1 million to 1:4,000 to meet the requirements of planning at various levels.

- Multi-Temporal Satellite Images on Topsoil Attribute Quantification and the Relationship with Soil Classes and Geology Bruna C. Gallo 1,2 , José A. M. Demattê 1,2,*, Rodnei Rizzo 1 , José L. Safanelli 1 , Wanderson de S. Mendes 1 , Igo F. Lepsch 1 , Marcus V. Sato 1 , Danilo J. Romero 1 and Marilusa P. C. Lacerda 3, Department of Soil Science, College of Agriculture Luiz de Queiroz, University of São Paulo, Rua Pádua Dias, 11, Piracicaba, Cx Postal 09, São Paulo, CEP 13416900, Brazil;

    - **ABSTRACT**: The FAO (Food and Agriculture Organization) estimates that the world population will increase from 7.3 billion to approximately 9 billion by 2050. . However, arable soil areas are decreasing due to the anthropogenic impact of degradation and urbanization. Poor soil management practices will threaten its capacity to provide soil resources, and possibly not enhance climate change policy. The proposed method is to assess bare soil, by masking green vegetation and crop residue targets from Landsat images. The compilation of time series images creates a BSCI, providing soil surface reflectance data for modeling topsoil properties.

- Soil mapping approach in GIS using Landsat satellite imagery and DEM data Erturul Aksoy, Gökhan Özsoy* and M. Sabri Dirim Uludag University, Faculty of Agriculture, Department of Soil Science, 16059 Görükle-Bursa, Turkey.

    - **ABSTRACT**: The objective of this study was to create base soil survey maps of the studied lands using Landsat satellite imagery and Digital Elevation Model (DEM) data in a GIS framework. Specific goals were to generate soil maps and to test the usage probability of slope class map overlies colour composite images as a preliminary map for soil survey in a hilly terrain.

Surrogate soil-landscape data layers were derived from Landsat satellite imagery and a 10 m DEM. The data were also used to produce 3D-view with slope class boundaries superimposed Landsat image and relief shaded map as a colour map in order to select possible site of soil profile pits and to define physiographic units. Six soil series formed on two different physiographic units were determined, described and sampled. Soil profiles have been classified according to Soil Taxonomy and FAO-Unesco soil map of the World legend classification systems. The methodology was adequate for soil survey and mapping of some types of soils.

- Soil Type Classification and Mapping using Hyperspectral Remote Sensing Data, Amol D. Vibhute1 *, K. V. Kale1 Senior Member IEEE 1. Dept. of Computer Science & IT, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad-431004 (MS), India *amolvibhute2011@gmail.com, kvkale91@gmail.com Rajesh K. Dhumal1, 2, S. C. Mehrotra1, 2 2. Srinivasa Ramanujan Geospatial Chair, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad-431004 (MS), India dhumal19@gmail.com, mehrotra.suresh15j@gmail.com

    - **ABSTRACT**: Hyperspectral remote sensing has been widely used for mapping of soil, its classification and also its texture description. It is beneficial in urban and rural management. The present work reports the study regarding classification soil analysis using Support Vector Machine (SVM). Hyperion Hyperspectral satellite data with 10nm fine spectral resolution of Phulambri region of Aurangabad district of Maharashtra (India) which lies between 20° 06' N latitude and 75° 25' E longitude was used for soil classification. Gaussian Radial Basis Function (RBF) kernel of SVM was used to extract five various soils types and achieved overall accuracy of 71.18% and with Kappa Value of 0.57 having sufficient training samples. It was found that the soil of the region may be classified in five categories. The maximum area (51 %) was covered by the brown sandy soil, whereas the minimum (.02%) by gray clay soil. The result is of great significance for soil analysis of very complex region without reduction of dimensionality in satellite data.

- Application of Satellite Remote Sensing to find Soil Fertilization by using Soil Colour Study Area Vellore District, Tamil Nadu, India. http://dx.doi.org/10.3991/ijoe.v9i2.2530 N. Suresh Kumar, S. Margret Anouncia, M. Prabu VIT University, Vellore, Tamil Nadu

    - **ABSTRACT**: Nowadays the usage of Remote Sensing and GIS techniques are too vast and many Earth observations had been done by Remote Sensing

and GIS techniques. The innovation of this work is to use the LANDSAT image data and GIS Techniques to assess land information and soil classification in the most coming up area of Vellore district and propose the possible fertilization for this study area. Landsat image is classified by minimum distance classification algorithm and according to the reflectance characteristics of the surface material. From the classified data we can find out the best fertilization for the best soil using colour of the soil. It is proved that, within limitations the classification algorithms and threshold parameters have an important influence on the classification result.

# CHAPTER 4
# PROBLEM DEFINITION AND SCOPE

## 4.1 Problem Statement

To design and develop a system which suggest which crop farmers should plant so that they will get maximum profit according to their soil properties and various inputs.

### 4.1.1 Goals and Objectives

- Goals:

1. To deliver a system which takes the input as the location of the farmer and according to that location it will give the suggestion to farmer which crop should they plant.
2. To build a system which extract soil features from satellite image.
3. Also, a model which can give the soil type using image classification.
4. Gathering the data for making datasets, like climate data, crop data, etc.
5. Analysis of the data and filtering it.
6. Extract soil properties from home-based soil testing methods.
7. Predict crop from the model based on many fields as input.

- Objectives:

1. To use remote sensing data to extract the features of soil.
2. To use physical data of soil properties rather than depending completely on the Remote sensing data.
3. To create a system which can be beneficial in the agriculture.
4. Gathering a large number of images of different soil types for soil detection model.
5. Deployment of models into android and web application.
6. Making chatbot for user to simplify the methods of soil testing at home.
7. To create a system which can be beneficial in the agriculture.

### 4.1.2 Statement of scope

1. Scope of our project is not limited up to soil properties but also the climate conditions. It can be used by anyone who needs to sow a crop which is according to the soil properties and many other constraints like climate data respectively.

2. A simple application having wide scope of crops according to the soil properties and gives suggestion according to the input of different constraints.

## 4.2 Major Constraints

1. Input should be a location or a longitude and latitude of the farm or the place.
2. For the suggestion of crops, it should be done in a single application.
3. Input should be soil type for getting better result
4. Weather data is also one of the most important input constraint in this project.

## 4.3 Methodologies of problem solving and efficiency issues.

## Methodologies of Problem Solving:

**Reading the Literature survey:**

The information about the soil and its type is mandatory to do this type of project. Also, the information about the remote sensing data, satellite images and how to get the features of soil from the satellite image. Reading research papers gives more clear idea about how to do project in efficient way so that it gives maximum output to the farmers.

**Gathering of an Images:**

We decided to collect the time series data from year 2010-2020 which gives more accurate information about the land. We are gathering the images from https://earthexplorer.usgs.gov/ of 10 years specifically of the month march, April, may as in this month the most of the farms are empty.

We collected more than 1000 images of each soil type for training our soil identification model.

**Preprocessing of the Image:**

After gathering all the images, the images are classified and the particular are is going to be selected. As the image is covering a large area, we are going to select the small area on which we can do our next step of extraction

**Extraction of features from soil data:**

After selecting of the data, the features of the soil is going to extracted using different software like ArcGIS. This gives us the data which is required to extract the features of the soil by using the reflective image.

**Physical data:**

After the extraction of the features from the satellite images, we need to verify that the data gathered from the satellite image is matching to the actual data. For this the physical testing of the soil is required. So, we are planning to the soil testing of a farm.

In the soil identification model we just have to give input as image of soil and it gives us the type of the soil.

**Compilation of data:**

After the verification of the data, we are going to compile data into a specific manner so that it matches to the dataset of the different crops.

**Build UI:**

Android based UI is developed where user can locate their location using the map or using the different input criteria. He is able to identify the soil type using soil identification model provided in application. In UI, there is our final model which takes input as soil type obtained from soil identification model, and the climate data from the datasets which we have provided and according to the live weather. On the basis of these inputs it gives the output as crop name which farmer can sow.

## 4.4 Outcomes

The famer is able to see what is the properties of his soil without doing the physical testing and also, he is able to see what are the different types of crops he can plant in his farm which is suitable to the properties of his soil and different constraints required for good agriculture.

## 4.5 Advantages and Applications

### 4.5.1 Advantages

1. The main advantage of this system is that farmer will get to know about his farms soil properties.
2. Farmer will also be able to see what are the different types of crops he can plant according to his soil condition.
3. This system is based on the Satellite images so it doesn't required time to get the details of the soil.

4. This can also help with costs, and also increase the productivity by knowing the properties.

### 4.5.2 Applications

1. In industries which are trying to get the soil properties and its productivity.
2. For the farmers to get different suggestion of crops for his farm.
3. In agriculture sector it can be a big enhancement as decision support system for crop suggestion.

# Chapter 5

# Project Plan

## 5.1 Project Estimates

### 5.1.1 Reconciled Estimates

- Cost Estimates – NA
- Time Estimates – 7 Months

### 5.1.2 Project Resources

1. Windows/Linux operating system.
2. Datasets of climate of last 10 years.
3. Datasets of different crops and their requirement.
4. Various research papers
5. Jupiter notebook
6. Android studio
7. Weather API's
8. Different software's and frameworks, for ex. XAMPP, FLASK, etc.

## 5.2 Project Schedule

### 5.2.1 Project Task Set

- Task set:

Task 1: Induction and Research of Decision Support System

Task 2: Identifying the factors influencing the crops and soil

Task 3: Technologies we can use and select one for our project

Task 4: Finding Datasets and Cleaning

Task 5: Developing the ML prediction model

Task 6: Creation of Frontend.

Task 7: Soil testing methods that can be implemented at home

Task 8: Creation and implementation of chat bot to assist for the methods (JavaScript)

Task 9: Creating a ML model that consist of the parameters for the prediction of crops

Task 10: Deployment of the ML model in android and Website

Task 11: Testing

## 5.3 Team Organization

### 5.3.1 Team Structure

In our team there are 5 team members and one guide assigned to us, the structure is given as follows-

1. Dr. Parminder Kaur (Guide)
2. Sagar Govind Yadav (IT) (RN- 400170)
3. Mayur Sahebrao Warpe (IT) (RN-400168)
4. Varad Sudhir Lad (IT) (RN-400136)
5. Chaitanya Mahadeo Tupsamudre (IT) (RN-400167)
6. Omkar Pandurang Patil (CSE) (RN-406229)

### 5.3.2 Management, reporting and communication

Progress of project is submitted on every meet scheduled with guide. A log book is maintained showing all work and their respective dates which we have mentioned above in the task set.

# Chapter 6

# Soil identification model in ML and home-based soil testing methods implementation

## 6.1 Soil identification model

The soil identification model which is able to give us output of soil type. We have created this model using google teachable machine and some libraries of python. first model predicts the Soil type for which we have collected around 1200 images of different soil types like alluvial, black, clay, saline, etc. For the second model, we have taken input from the user and matched the input with the given data set.

Given below, the provide images shows that how our model of soil identification is and how it is implemented in the android as well as in the web application

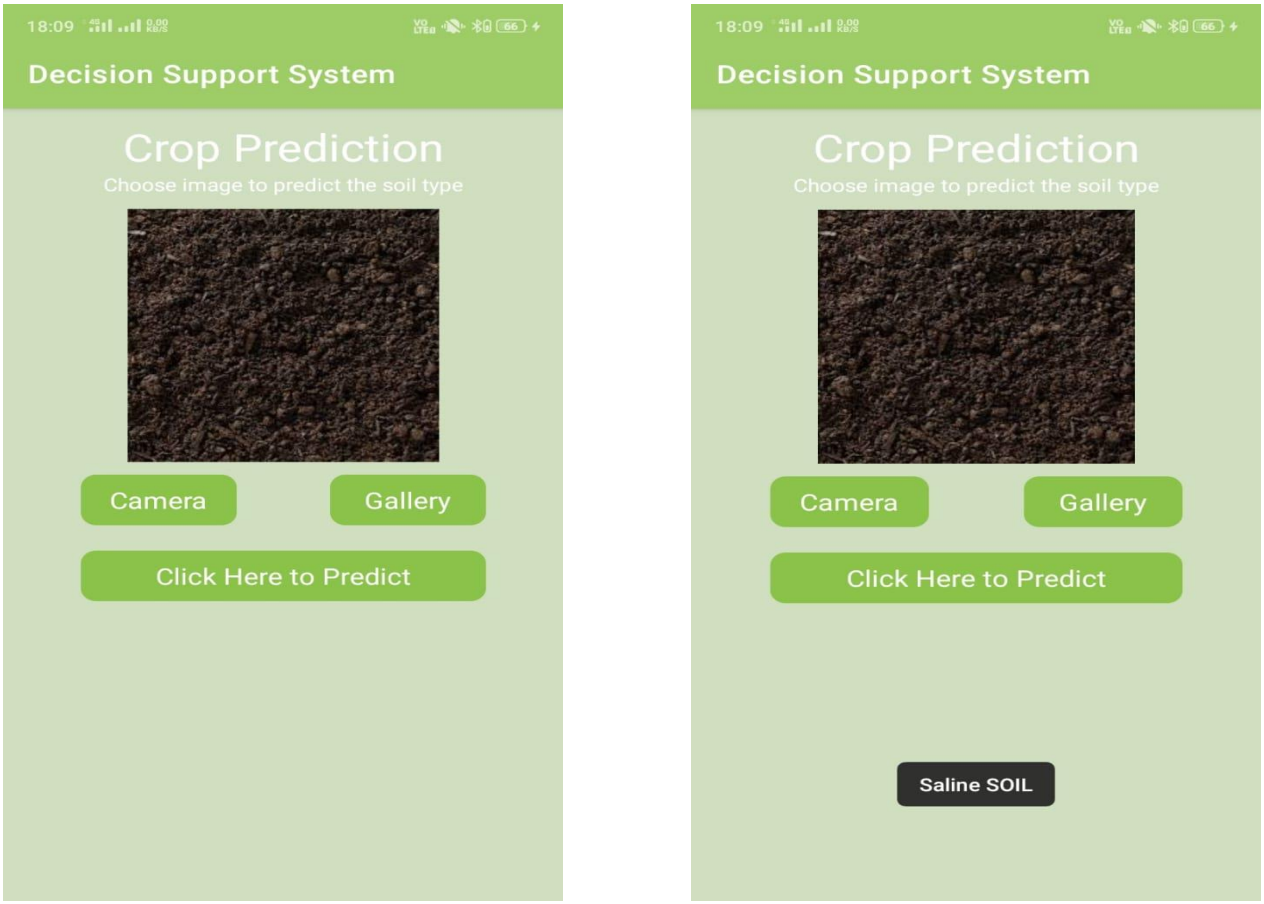


Image 6.1: Screen shots of the soil identification model

## 6.2 Home-based soil testing methods

A soil test is important for several reasons: to optimize crop production, to protect the environment from contamination by runoff and leaching of excess fertilizers, to aid in the diagnosis of plant culture problems, to improve the nutritional balance of the growing media and to save money and conserve energy by applying only the amount of fertilizer needed. Pre- plant media analyses provide an indication of potential nutrient deficiencies, pH imbalance or excess soluble salts. This is particularly important for growers who mix their own media.

Soil testing in laboratories is quite expensive and it is not affordable for the farmers. Instead of doing soil testing at laboratories farmers can test their field soil at home using home-made soil testing methods. Given below the method of soil testing at home which we have mentioned in the application.

1. The jar soil test
2. The pH tests
3. Earthworm test
4. Soil aggregate stability test
5. Soil compaction test

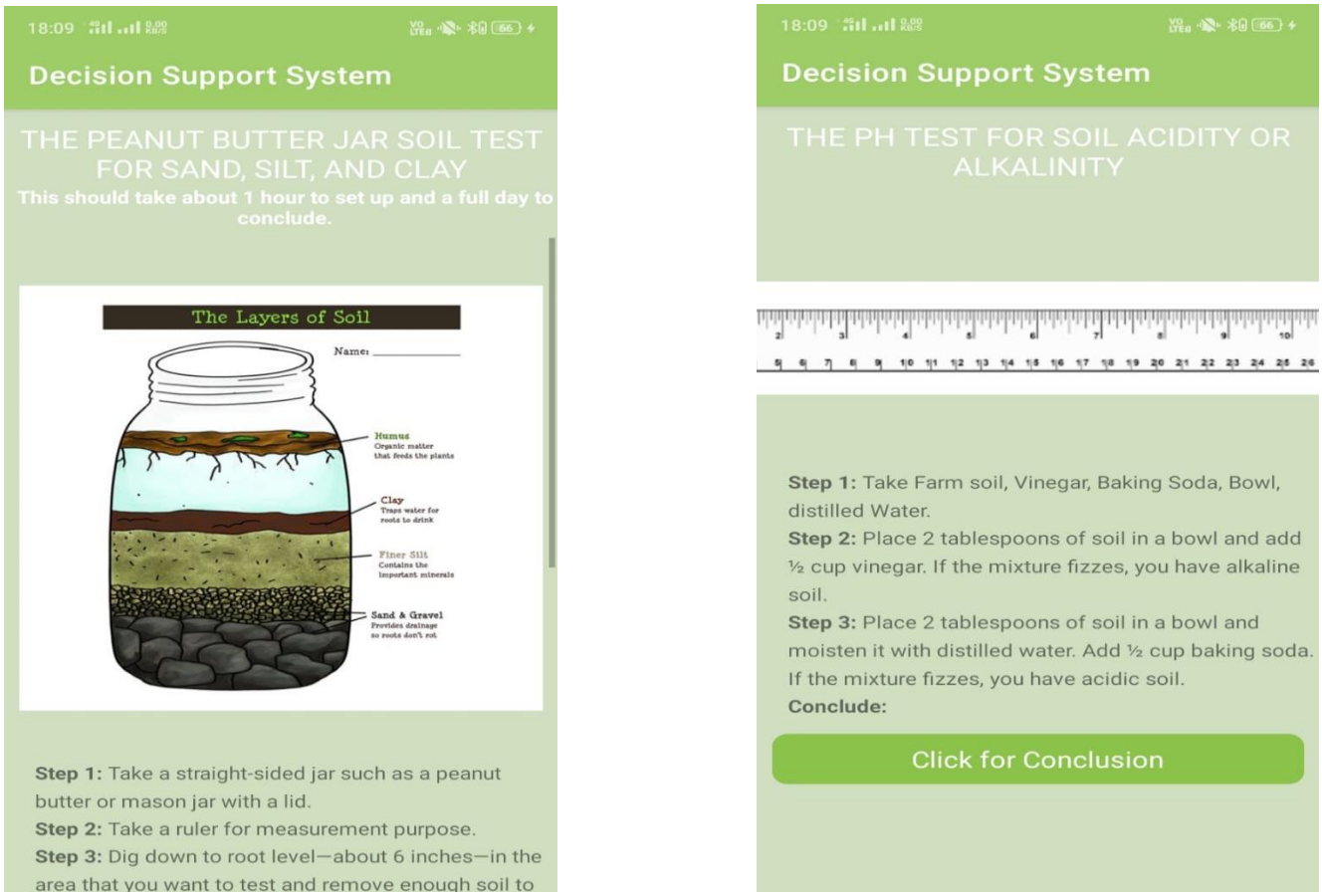


Image 6.2: Screen shots of the soil Properties methods

## 6.3 Chatbot for guidance of soil testing

The chatbot which we have created using JavaScript is helpful for farmers by direct guidance. In this chatbot farmer has to choose the soil testing method and accordingly the chatbot gives the step-by-step instructions to farmer what steps should he has to do.

Below some screenshots are there which shows that how the chatbot model is and how it gives the instruction to the user.

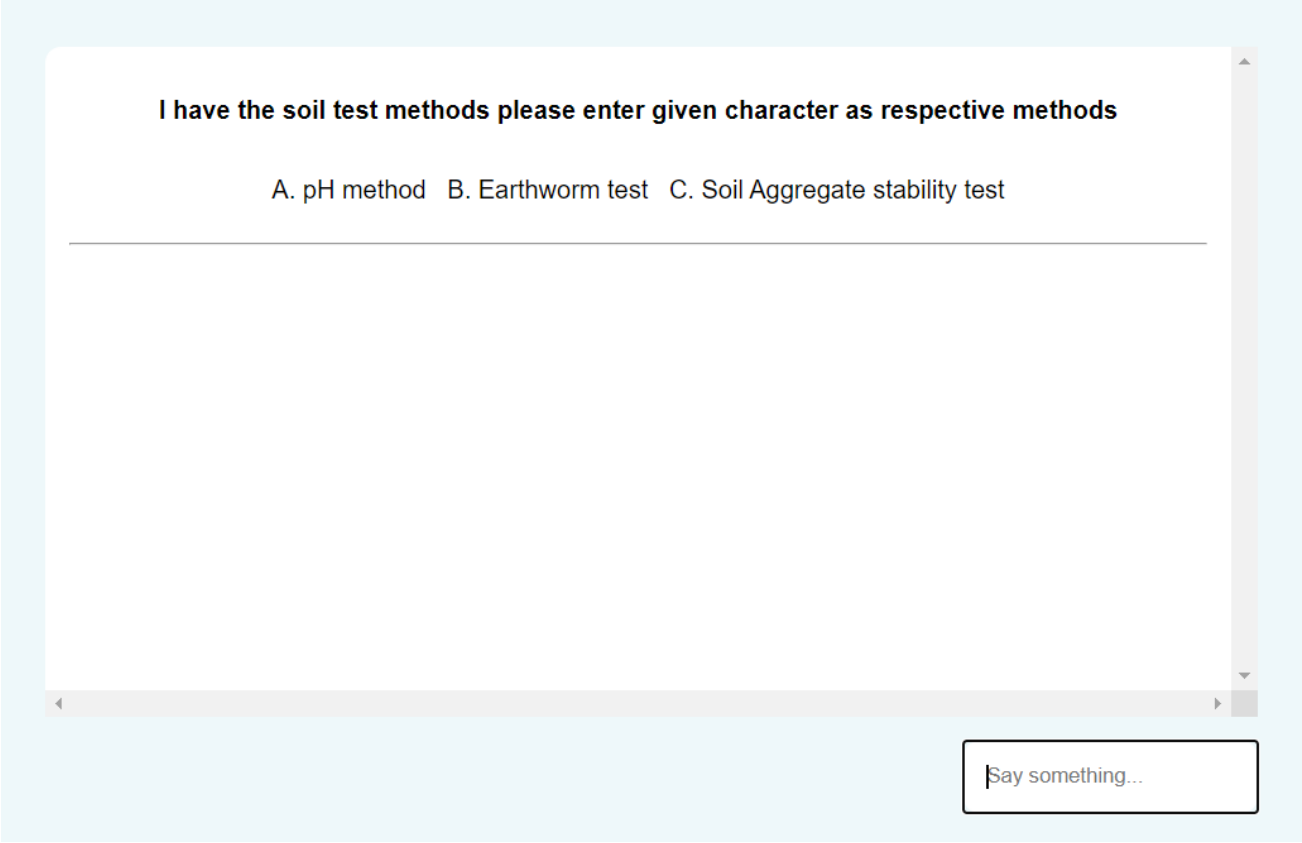


Image 6.3: UI of Chabot for soil testing methods

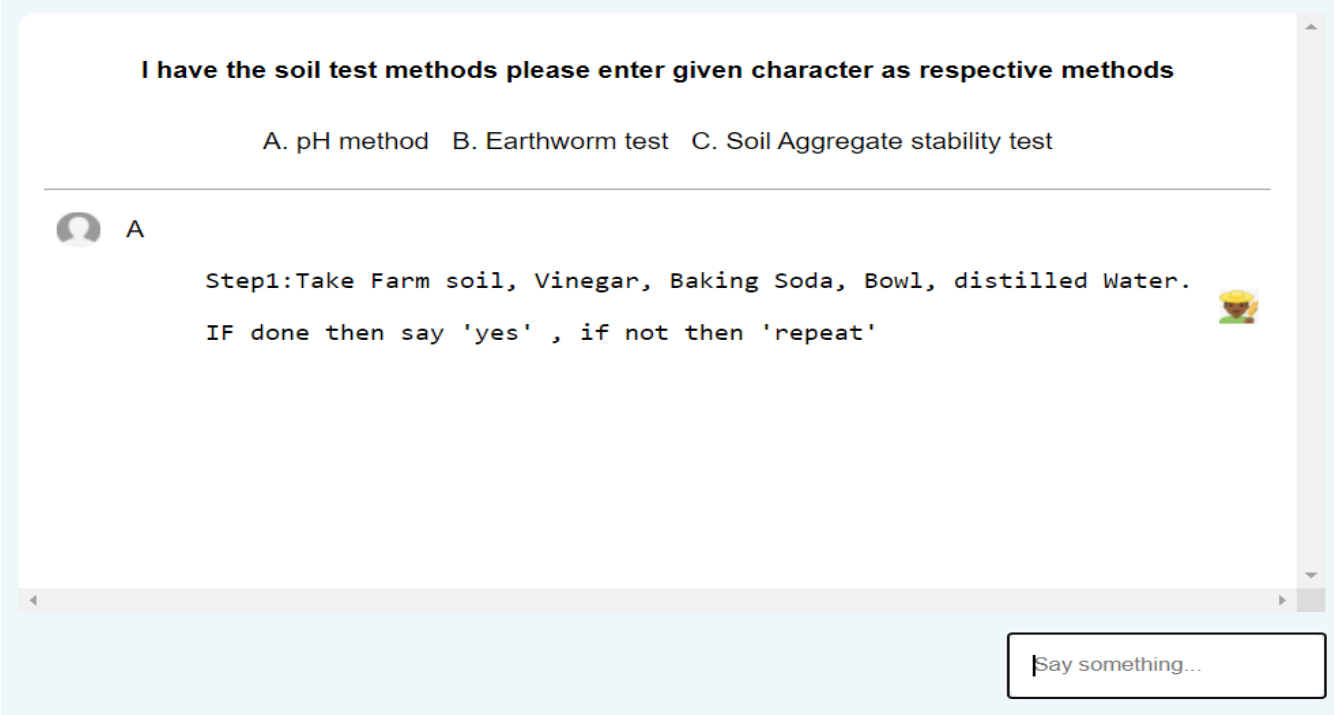


Image 6.4: Step-by-step instruction level 1

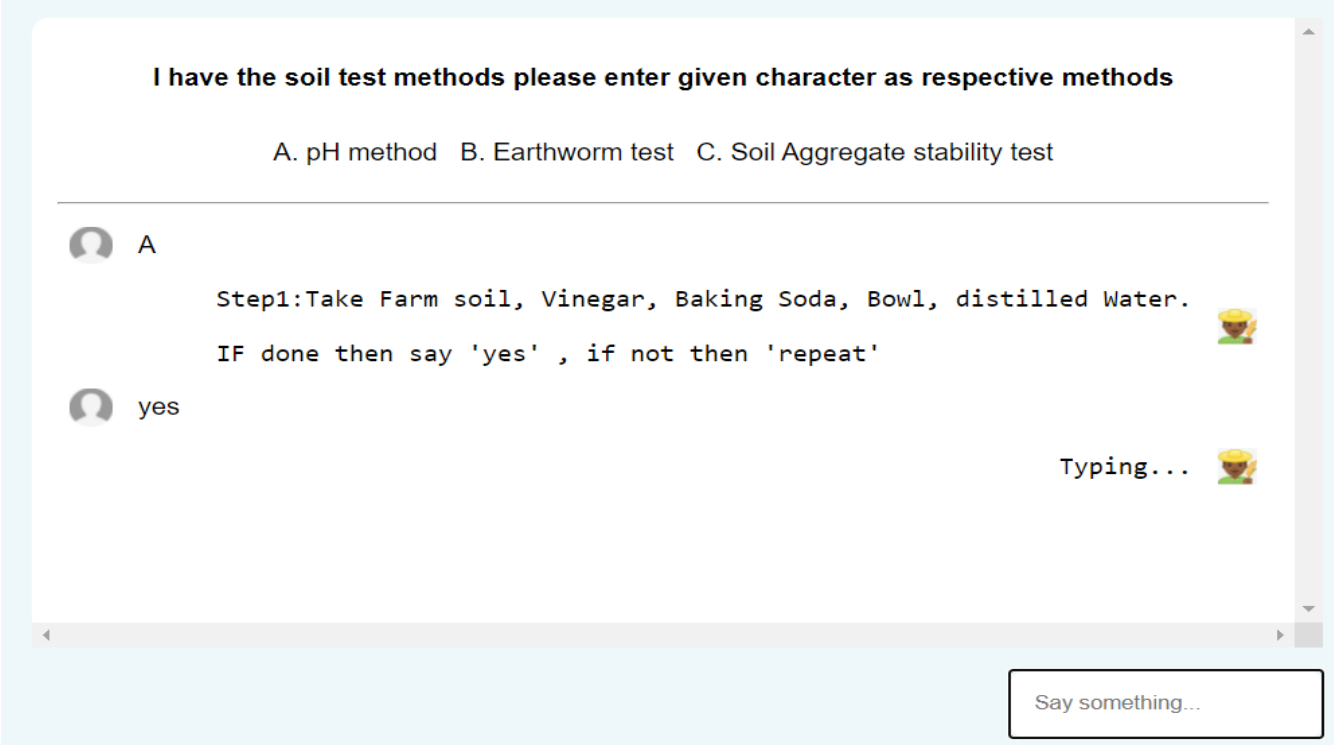


Image 6.5: step-by-step instruction level 2

# CHAPTER 7
# DEPLOYMENT OF ML MODEL INTO ANDROID APPLICATION

## 7.1 ML model

We have created two machine learning models out of which one is for the prediction of the soil type and another is for the crop suggestion. The Machine learning model is created with the help of google teachable machine and different libraries. The first model predicts the Soil type for which we have collected around 1200 images of different soil types like alluvial, black, clay, saline, etc. For the second model, we have taken input from the user and matched the input with the given data set.

After the Creation of the model, we have a major task to deploy that model into the android application and web application. In this chapter, we will look into the android application.

## 7.2 Android Application

Android application is made using java and has features used like intents, shared preferences, and networking. In the android application there are different pages as follows:

- Login Page
- Registration Page
- Home Page for the weather details, different methods
- Separate pages for each method
- Current weather detection page
- ML model page (soil type recognition)
- Crop prediction page.

## 7.3 Creation of Splash Screen

As the user opens an application user will able to see the splash screen for 10 seconds, we have created a splash screen which the contents name and logo of the application. We have attached the screenshot of the splash screen of the application.


Image 7.1: Splash Screen

After opening the application, the above screen opens for 5 seconds and it redirects to the home page for the details.

## 7.4 Home Page

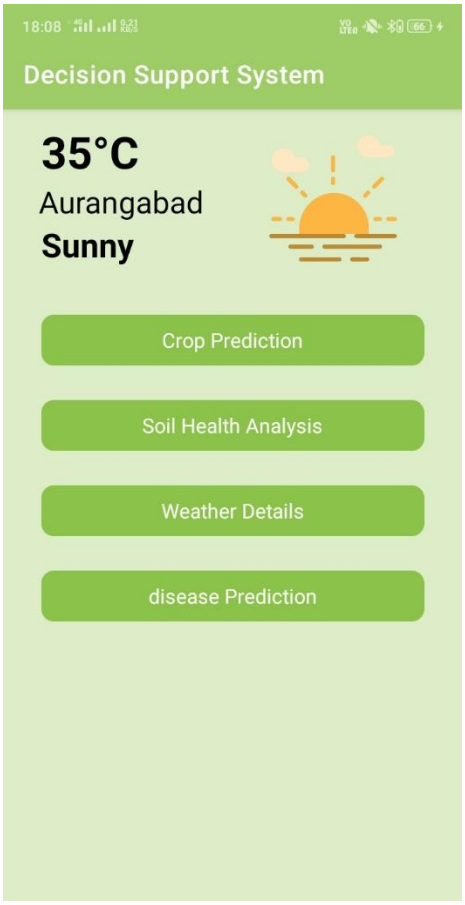


Image 7.2: Home Page

We want the home page should net and clean and it looks more attractive, user can use the application without any difficulty, and he is able to find anything easily, to find the details easily we have created a page/intent as simple as a single static page which consists of current weather details, the different buttons for the implementation of methods like ML model. On the home page, we have the following things.

- **Current Weather Details: -** We used the open Weather API for the current weather conditions and temperature, on the home page from accessing the location of the user it shows the details about the weather of his current city.

- **Crop Prediction: -** This button is basically to predict which type of soil the user has, once we click on this button it redirects to the crop prediction page where we can select the image/click image from the camera and it predicts the soil type as output. We have shown the image of the process in the application and how it works.

- On clicking the button, it redirects to the page of crop prediction as shown in the Image Screenshot of crop prediction On this page, we have to get the image or click the image as I have clicked the image from the gallery and it is displayed the area given in image Screenshot of crop prediction, once we click on the button Click here to predict button it shows you the type of soil, in this case, it shows the saline soil.

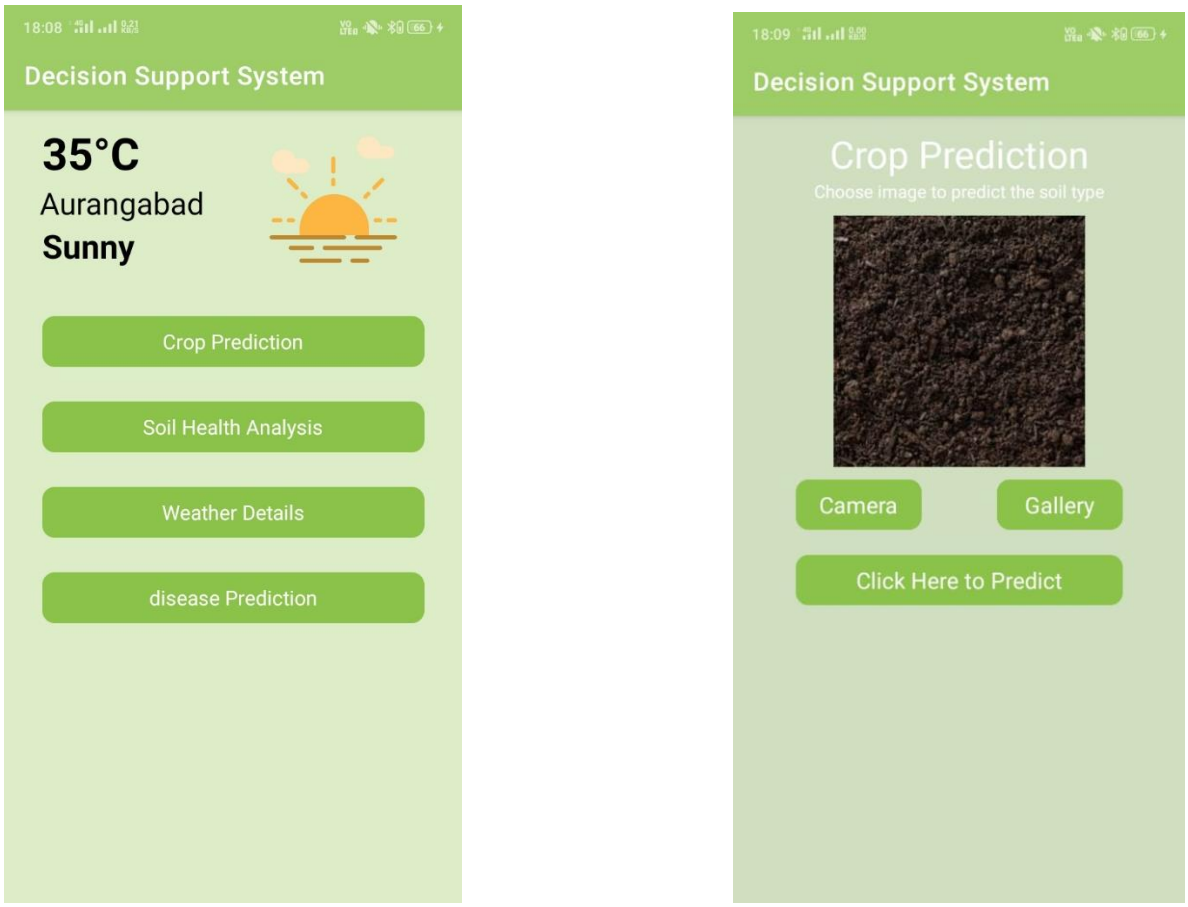


Image 7.3: Home page and Crop Prediction Model

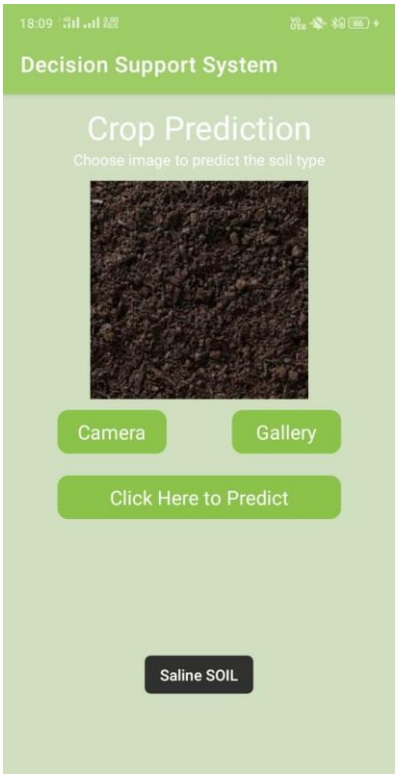


Image 7.4: Screenshot of Soil Prediction Model

- **Soil Health Analysis: -** Soil Health Analysis is the most important part of this project. Extraction of soil properties is the main task for the farmers because on the soil properties the type of crop that can be grown is depended also on how much soil I healthy and what fertilizers they need to add to the soil so that soil can become healthier. We have added some homemade methods from which the farmer can do the methods by sitting at home, it methods take 1-24 hours' time. We have implemented assistance or user can also access it by clicking on the button in the android application. We have added the image of soil health analysis below to get more about it. Currently, we have added the 5 methods as you can see in the image Screenshot of Soil health Analysis, the methods currently working are the peanut butter jar soil test for sand, silt, and clay, the pH test for soil acidity or alkalinity, the earthworm test to gauge organic matter, soil aggregate stability test and soil compaction test as shown in the below images.

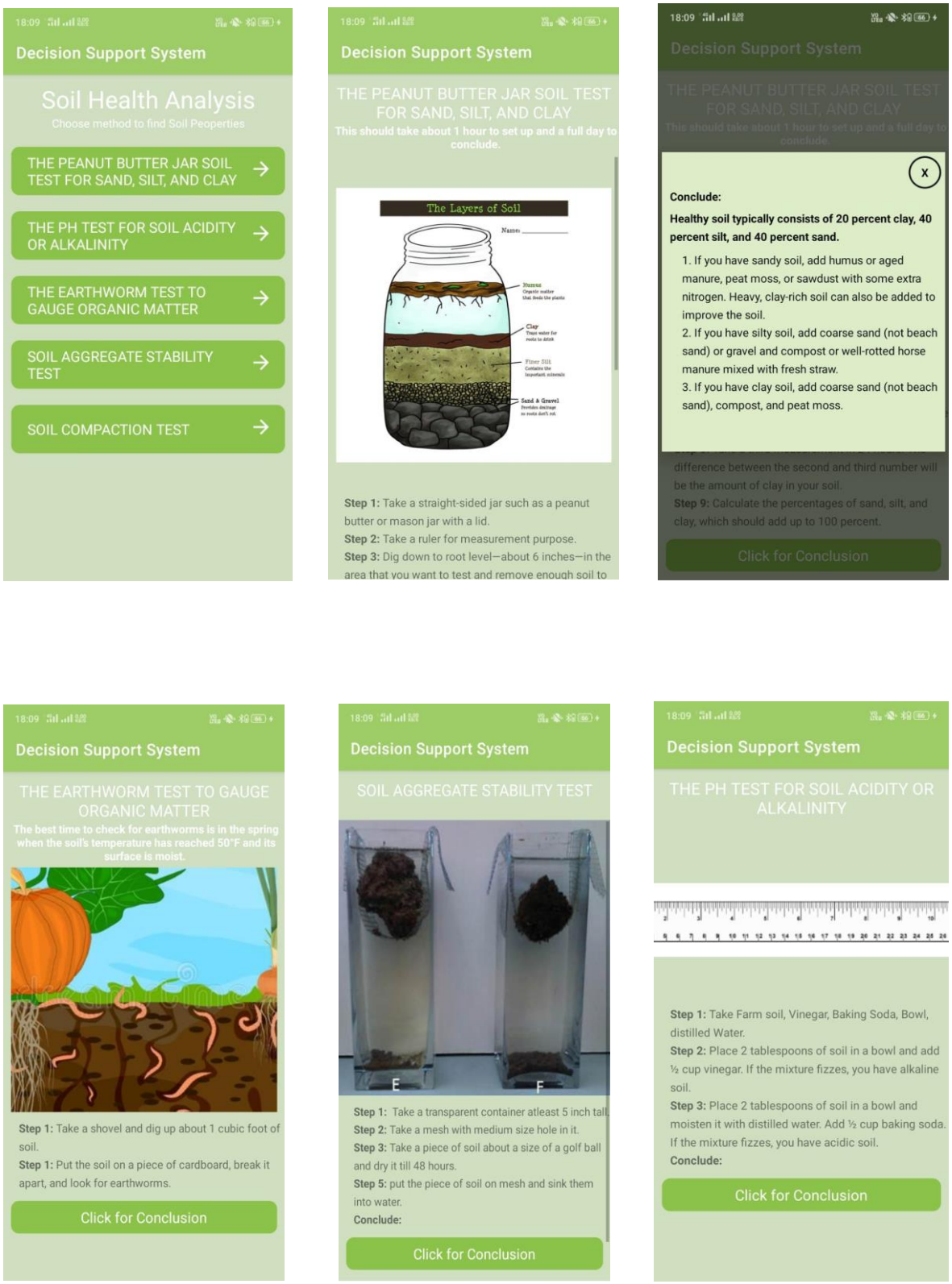


Image 7.5: Screenshots of Different Soil Health Analysis Methods

- **Weather Details: -** Weather Details are the important factor to get the current weather situation and future weather situation because a lot more things depend on the weather. We have implemented the Open Weather API for the details of weather it shows the details of the current weather as well as future data. We have attached an image of the application which has implemented the method weather.

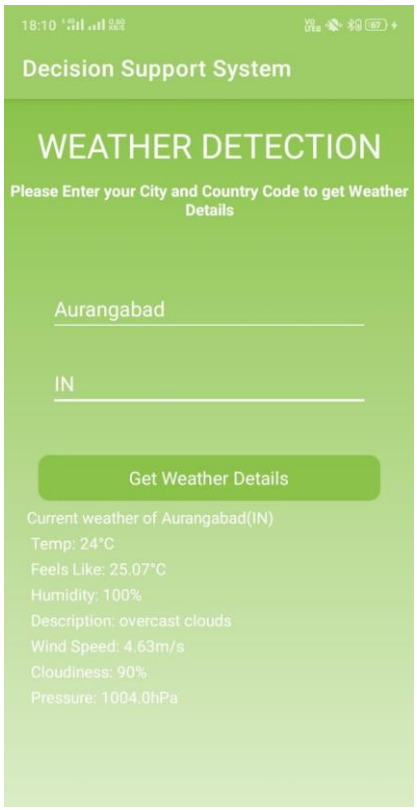


Image 7.6: Screenshots of Weather Detection Page

CHAPTER 8

# DEPLOYMENT OF ML MODEL INTO WEB APPLICATION

## 8.1 ML model

We have created two machine learning models out of which one is for the prediction of the soil type and another is for the crop suggestion. The Machine learning model is created with the help of google teachable machine and different libraries. The first model predicts the Soil type for which we have collected around 1200 images of different soil types like alluvial, black, clay, saline, etc. For the second model, we have taken input from the user and matched the input with the given data set.

After the Creation of the model, we have a major task to deploy that model into the android application and web application. In this chapter, we will look into the web application.

## 8.2 Web Application

Web Application is created using several web technologies. We used HTML, CSS, Bootstrap and Python for frontend web designing and we developed backend using MySQL, flask, Python and Machine Learning models.

In the web application there are different sections included as follows:

- Login Page
- Registration Page
- Home Page
- Current Whether API
- ML model section (soil type recognition)
- Crop prediction page
- Soil Health Analysis

## 8.3 Creation of Login Page

As the user opens an application user will have to login or register himself/herself. In this section, we will look into the login module. We created a login module using HTML, CSS, MySQL and PHP. We have made this module as a responsive module that will adjust itself according to screen size of users. Here are some images of login page.


Image 8.1: Login page desktop view

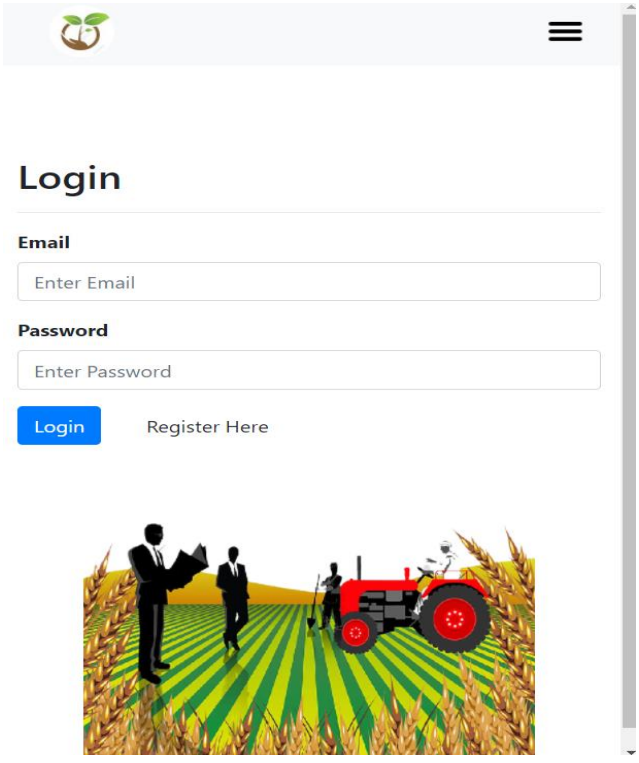


Image 8.2: Login page Mobile view

## 8.4 Creation of Register Page

If user nor registered then he must be registered himself/herself at this page.
Here are some sample images of Register page

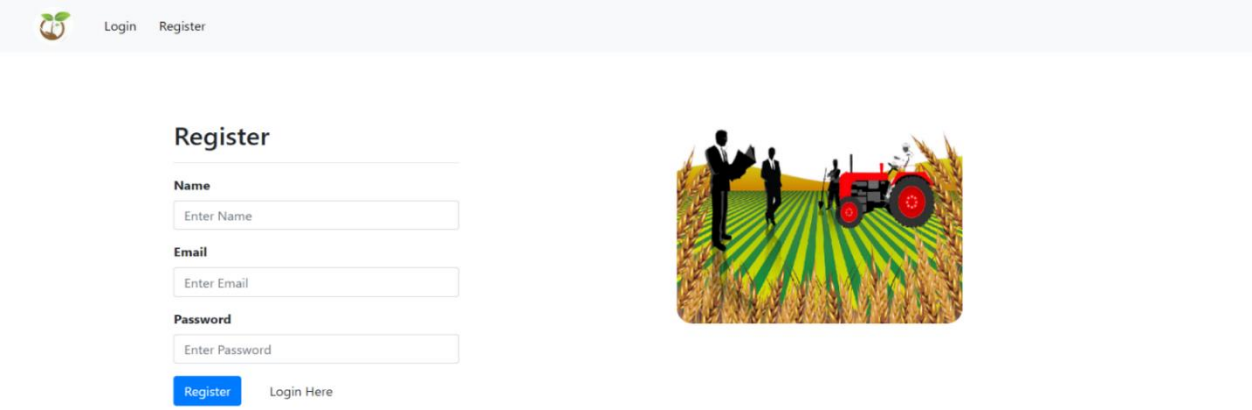


Image 8.3: Register page desktop view

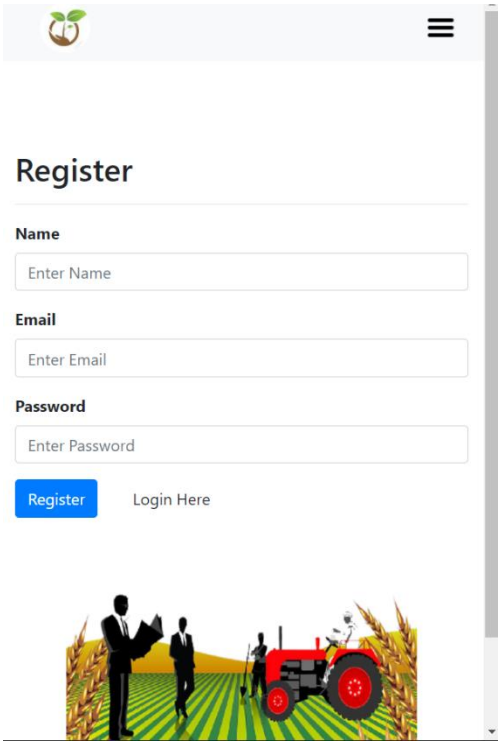


Image 8.4: Register page mobile view

## 8.5 About Home Page

We want the home page should net and clean and it looks more attractive, user can use the application without any difficulty, and he is able to find anything easily, to find the details easily we have created multiple sections on single page which consists of current weather details, The section for crop prediction ML model, Soil Health analysis.

Here are some sample images of Home page


Image 8.5: Home page Desktop view

## 8.6 Deploying current weather API

**Current Weather Details: -** We used the open Weather API for the current weather conditions and temperature, on the home page from accessing the location of the user it shows the details about the weather of his current city. Weather Details are the important factor to get the current weather situation and future weather situation because a lot more things depend on the weather.


Image 8.6: weather API Desktop view


Image 8.7: Home page mobile view

## 8.7 ML model section (Soil Type Recognition)

- **Crop Prediction: -** This section contains a button which basically predict which type of soil the user has, by taking a soil image as input we predict the type of soil. Further this button will redirect user to crop prediction page. We have shown the image of the process in the application and how it works.

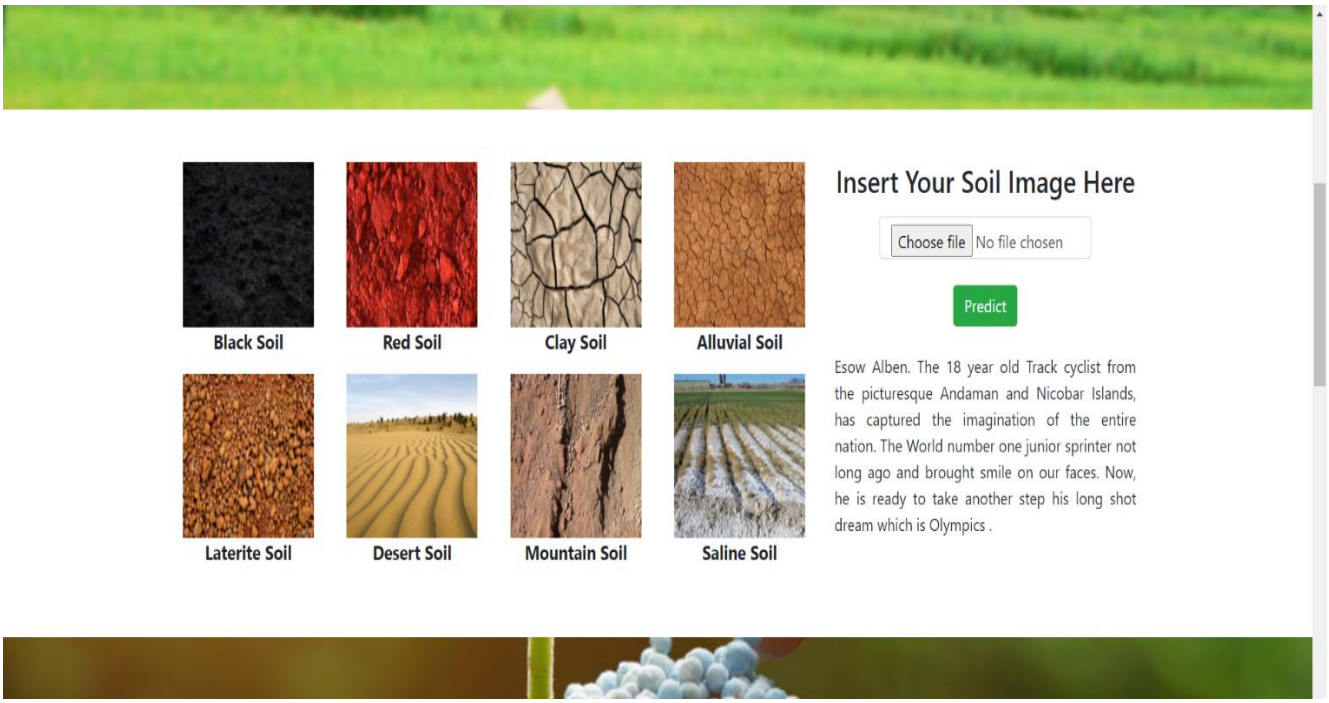


Image 8.8: ML model desktop view

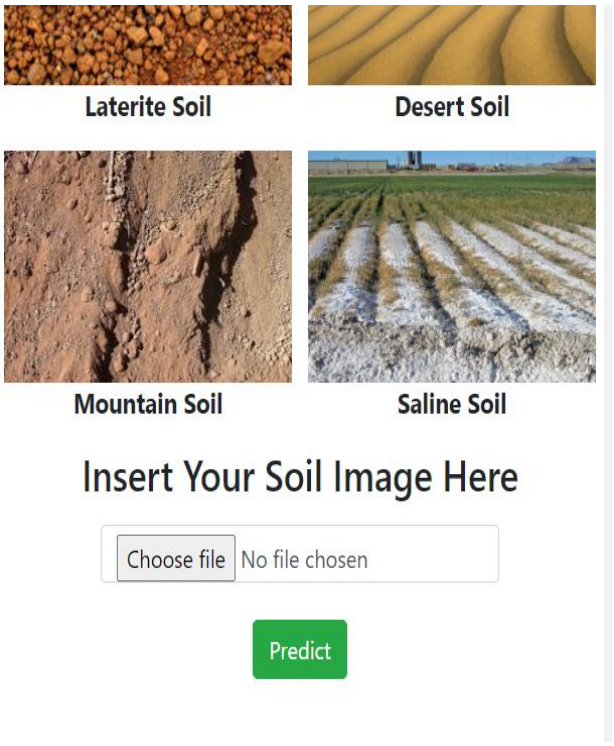


Image 8.9: ML model mobile view

- **Image Prediction page:** On clicking the button it redirects to the page of crop prediction as shown in the Image Screenshot of crop prediction on this page, we have to get the image or click the image as I have clicked the image from the gallery and it is displayed the area given in image Screenshot of crop prediction, once we click on the button Click here to predict button it shows you the type of soil and plants that farmer can sow. In this case, it shows the alluvial soil.

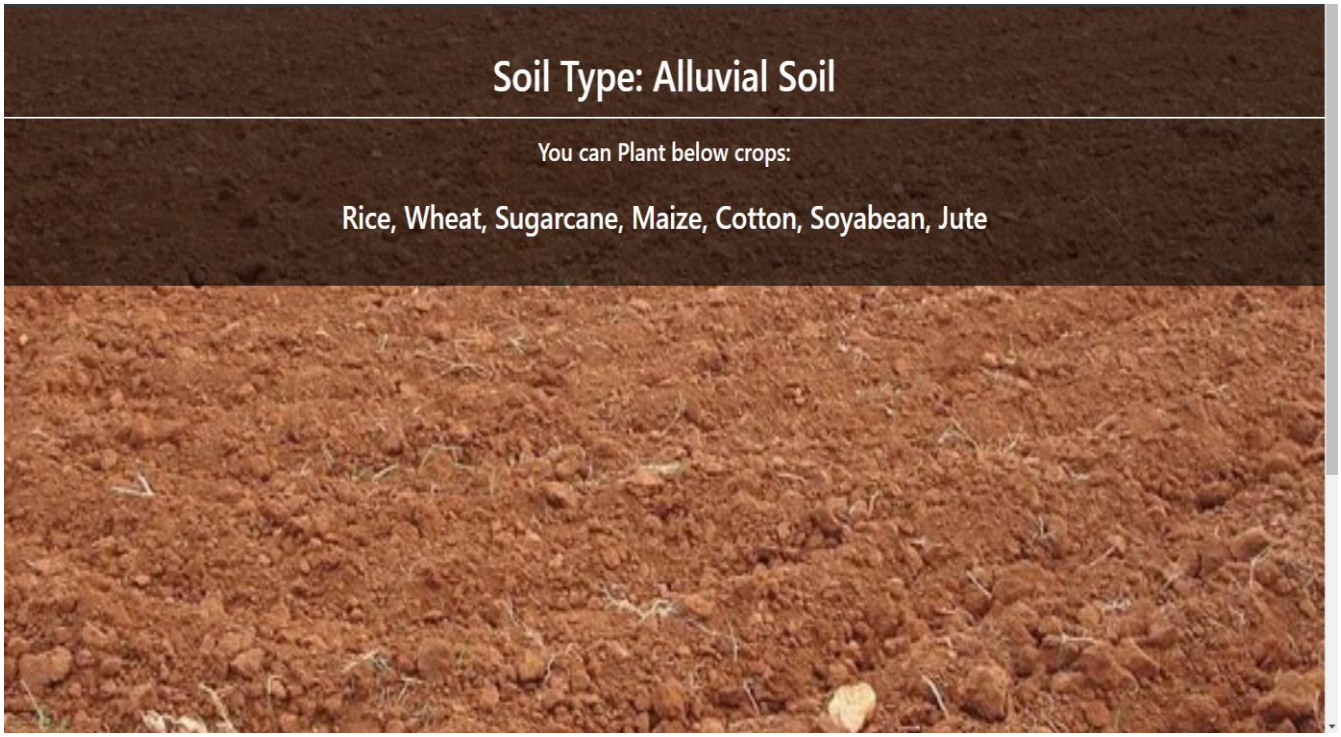


Image 8.10: Image prediction page

## 8.8 Soil Health Analysis

- **Soil Health Analysis: -** Soil Health Analysis is the most important part of this project. Extraction of soil properties is the main task for the farmers because on the soil properties the type of crop that can be grown is depended also on how much soil I healthy and what fertilizers they need to add to the soil so that soil can become healthier. We have added some homemade methods from which the farmer can do the methods by sitting at home, its methods take 1-24 hours' time. We have implemented assistance or user can also access it by clicking on the button in the android application. We have added the image of soil health analysis below to get more about it. Currently, we have added the 5 methods as you can see in the image Screenshot of Soil health Analysis, the methods currently working are the peanut butter jar soil test for sand, silt, and clay, the pH test for soil acidity or alkalinity, the earthworm test to gauge organic matter, soil aggregate stability test and soil compaction test as shown in the below images.

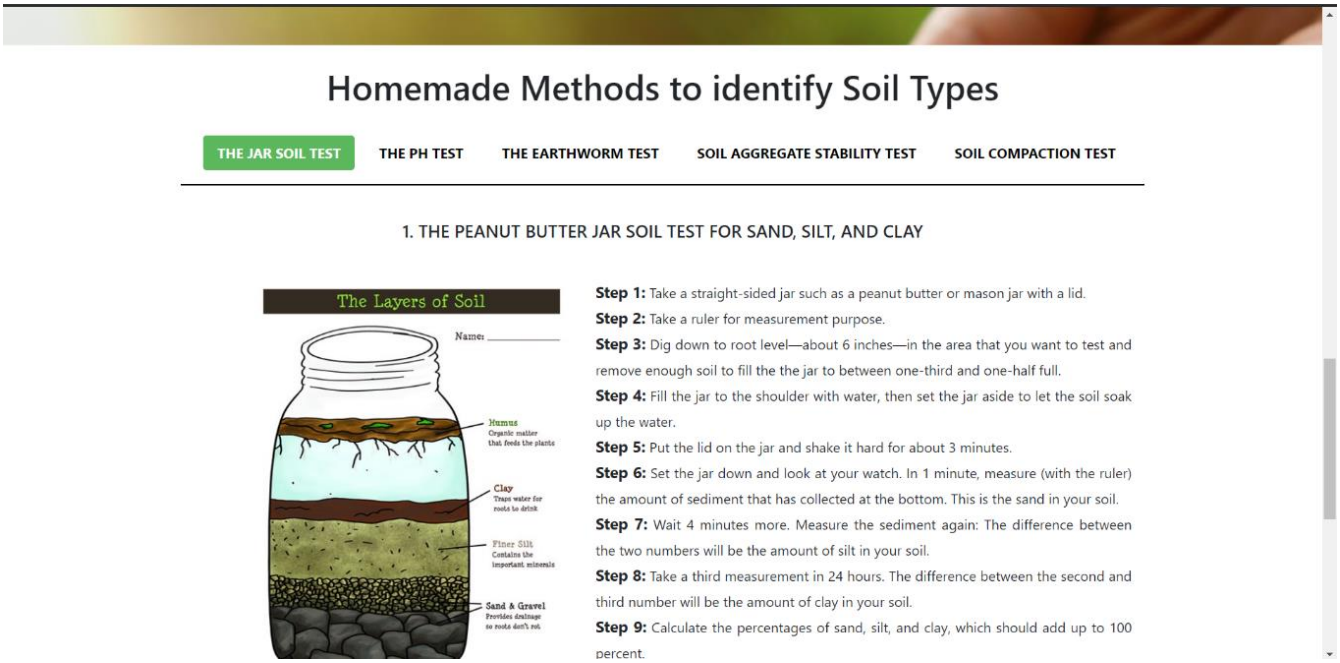


Image 8.11: Soil Health Analysis Desktop view

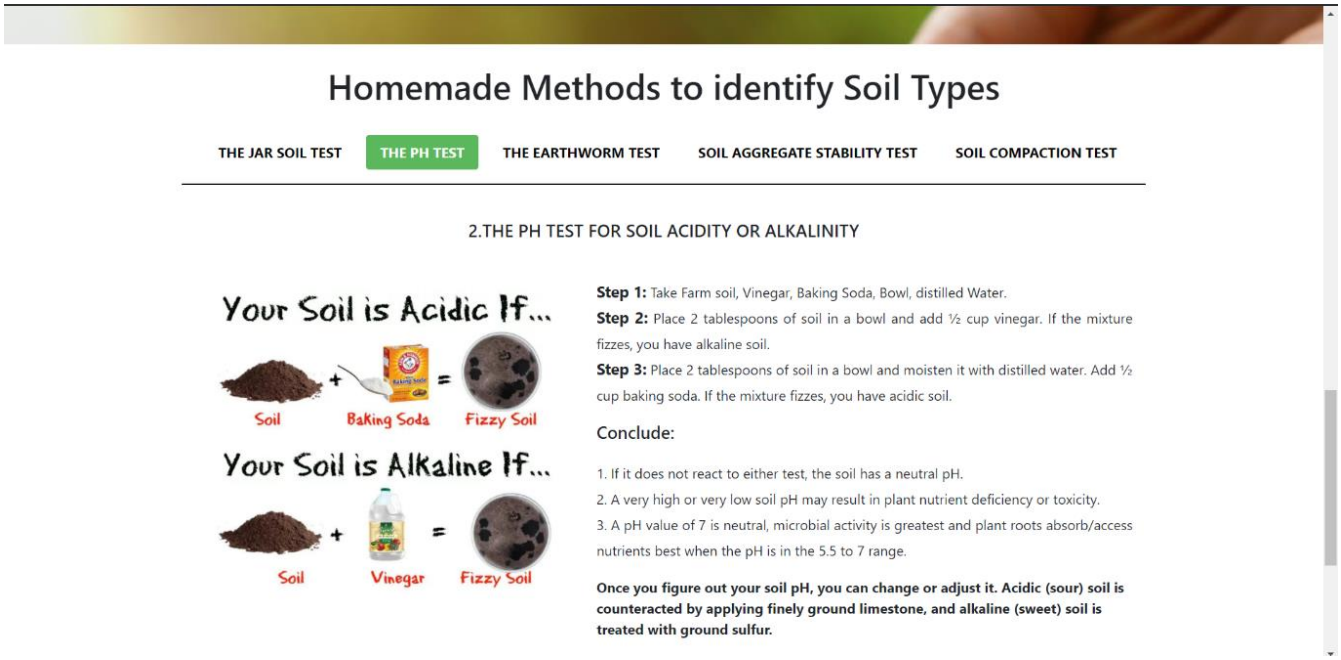


Image 8.12: Soil Health Analysis Desktop view

# CHAPTER 9
# CREATION OF ML MODEL FOR CROP PREDICTION

## 9.1 ML Model

Now-a-day, with technological enhancement, many Organizations using Machine learning, data science on a large scale. This model works with predefined data points available in the form of datasets. This set contains information about previous events. Organizing these datapoints before it is given to the model is very important these are where we use data Analysis. If the data given to the ml model is not well structured, it gives out false or undesired output. This can cause major loss to the business. Hence it is very important to do proper data analysis.

## 9.2 About Dataset

The data that we are going to use in this example is about Crop. Specifically containing various information about Crops, such as a Sowing Time, Ph value, min temp, max temp, soil type, etc. Here raw data isn't useful, so we need to understand that simply collecting data isn't enough. Here data analysis plays a vital role in unlocking the information that we require and gaining new insights from this raw data.
Let's think as data scientists and clearly define some of the problems For example, is there data on the crops of other crops and their characteristics? What features of soil affects the crop? or Ph value? Or Min temp? Does water availability also affect the crop, or perhaps, something else?
We collected the dataset, having the following fields sowing time, min temp, max temp, Ph value, soil type etc.

## 9.3 Modules Needed

numpy: numpy is the package for scientific computing with Python. Numpy can be used as an efficient multi-dimensional container of generic data.
pandas: Panda is an open-source library that allows you to perform data manipulation in Python.
matplotlib: Matplotlib is a Python 2D plotting library.
seaborn: Seaborn is a Python data-visualization library based on matplotlib. Seaborn provides a high-level interface for drawing attractive and informative statistical graphics.

Steps for installing these packages:

If you are using anaconda- jupyter / spyder or any other third-party software's to write your python code, make sure to set the path to the "scripts folder" of that software in the command prompt of your pc.
Then type – pip install [package-name]
Example: pip install NumPy
Then after the installation is done. (make sure you are connected to the internet!!) Open your IDE, then import those packages.

Import packages
Set the path to the crop.csv file

Perform various data cleaning and data visualization operations on your data. These steps are illustrated beside each line of code in the form of comments for better understanding.

## 9.4 How to Create ML Model

ML Model takes 8 parameters to predict the output the software used to develop this model in jupyter note, and some python libraries such as NumPy pandas and sklearn. steps are as follows:

Step 1: Import the modules needed.

```
In [1]: import numpy as np
        import pandas as pd
        import matplotlib.pyplot as plt
```

Step 2: check for first five entries of dataset.

```
In [2]: dataset = pd.read_csv("crop-data.csv")
        dataset.head()
```

Out[2]:

| | srno | Sowing_Time | Region | Min_Temp | Max_Temp | Soil_pH | Soil_Type | Min_Rainfall | Max_Rainfall | Rainfall | Temperature | Crop |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | june | andhra pradesh | 20 | 31 | 6.0 | black cotton soil | 25.0 | 60.0 | 42.5 | 25.5 | bajra |
| 1 | 1 | july | andhra pradesh | 20 | 31 | 6.0 | black cotton soil | 25.0 | 60.0 | 42.5 | 25.5 | bajra |
| 2 | 2 | january | andhra pradesh | 20 | 31 | 6.0 | black cotton soil | 25.0 | 60.0 | 42.5 | 25.5 | bajra |
| 3 | 3 | february | andhra pradesh | 20 | 31 | 6.0 | black cotton soil | 25.0 | 60.0 | 42.5 | 25.5 | bajra |
| 4 | 4 | march | andhra pradesh | 20 | 31 | 6.0 | black cotton soil | 25.0 | 60.0 | 42.5 | 25.5 | bajra |

Step 3: Finding the missing value/null if any.

```
In [3]: # Finding the missing values
        dataset.isna().any()

        # Finding if missing values
        dataset.isnull().any()
```

```
Out[3]: srno            False
        Sowing_Time     False
        Region          False
        Min_Temp        False
        Max_Temp        False
        Soil_pH         False
        Soil_Type       False
        Min_Rainfall    False
        Max_Rainfall    False
        Rainfall        False
        Temperature     False
        Crop            False
        dtype: bool
```

Step 4: Drop unnecessary columns from data set

```
In [4]: dataset = dataset.drop(['srno'] ,axis=1)
```

```
In [5]: dataset.head()
```

Out[5]:

| | Sowing_Time | Region | Min_Temp | Max_Temp | Soil_pH | Soil_Type | Min_Rainfall | Max_Rainfall | Rainfall | Temperature | Crop |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | june | andhra pradesh | 20 | 31 | 6.0 | black cotton soil | 25.0 | 60.0 | 42.5 | 25.5 | bajra |
| 1 | july | andhra pradesh | 20 | 31 | 6.0 | black cotton soil | 25.0 | 60.0 | 42.5 | 25.5 | bajra |
| 2 | january | andhra pradesh | 20 | 31 | 6.0 | black cotton soil | 25.0 | 60.0 | 42.5 | 25.5 | bajra |
| 3 | february | andhra pradesh | 20 | 31 | 6.0 | black cotton soil | 25.0 | 60.0 | 42.5 | 25.5 | bajra |
| 4 | march | andhra pradesh | 20 | 31 | 6.0 | black cotton soil | 25.0 | 60.0 | 42.5 | 25.5 | bajra |

Step 5: Here, sowing is of object type(string), it should be int or float, so we need to change it and same for region and soil type.

```
In [6]: # Import LabelEncoder
from sklearn import preprocessing
#creating labelEncoder
le = preprocessing.LabelEncoder()
# Converting string labels into numbers.
dataset['Sowing_Time']=le.fit_transform(dataset['Sowing_Time'])
print(dataset['Sowing_Time'])
```

```
0         6
1         5
2         4
3         3
4         7
        ..
4919      5
4920      1
4921     11
4922     10
4923      9
Name: Sowing_Time, Length: 4924, dtype: int32
```

```
In [7]: dataset['Region']=le.fit_transform(dataset['Region'])
print(dataset['Region'])
```

```
0         1
1         1
2         1
3         1
4         1
        ..
4919     40
4920     40
4921     40
4922     40
4923     40
Name: Region, Length: 4924, dtype: int32
```

```
In [8]: dataset['Soil_Type']=le.fit_transform(dataset['Soil_Type'])
        print(dataset['Soil_Type'])
```

```
0       1
1       1
2       1
3       1
4       1
       ..
4919    5
4920    5
4921    5
4922    5
4923    5
Name: Soil_Type, Length: 4924, dtype: int32
```

## Step 6: Classification

```
In [9]: # X is independent variable
        X = dataset.iloc[:, range(0, 10)].values
        X
```

```
Out[9]: array([[  6. ,   1. ,  20. , ...,  60. ,  42.5,  25.5],
               [  5. ,   1. ,  20. , ...,  60. ,  42.5,  25.5],
               [  4. ,   1. ,  20. , ...,  60. ,  42.5,  25.5],
               ...,
               [ 11. ,  40. ,  10. , ..., 400. , 300. ,  12.5],
               [ 10. ,  40. ,  10. , ..., 400. , 300. ,  12.5],
               [  9. ,  40. ,  10. , ..., 400. , 300. ,  12.5]])
```

```
In [10]: # Y is Dependent Variable
         y = dataset.iloc[:, 10].values
         y
```

```
Out[10]: array(['bajra', 'bajra', 'bajra', ..., 'turnip', 'turnip', 'turnip'],
               dtype=object)
```

## Step 7: Splitting the data set into the Training set and Test set

```
In [11]: from sklearn.model_selection import train_test_split
         X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.25, random_state = 0)
```

## Step 8: Feature Scaling

```
In [12]: from sklearn.preprocessing import StandardScaler
         sc = StandardScaler()
         X_train = sc.fit_transform(X_train)
         X_test = sc.transform(X_test)
```

## Step 9: Fitting K-NN to the Training set

```
In [13]: from sklearn.neighbors import KNeighborsClassifier
         classifier = KNeighborsClassifier(n_neighbors = 5, metric = 'euclidean', p = 2)
         classifier.fit(X_train, y_train)
```

```
Out[13]: KNeighborsClassifier(algorithm='auto', leaf_size=30, metric='euclidean',
                              metric_params=None, n_jobs=None, n_neighbors=5, p=2,
                              weights='uniform')
```

Step 10: Predicting the Test set results

```
In [14]: y_pred = classifier.predict(X_test)
```

Step 11: Performance Analysis

Accuracy is totally depending on correct data and right algorithm.

```
In [15]: from sklearn.metrics import accuracy_score
         accuracy_score(y_test,y_pred)
Out[15]: 0.925264012997563
```

Step 12: Store this model for later use/implementation.

```
In [16]: import pickle
         filename ='Crop_Prediction_Model'
         pickle.dump(classifier, open(filename, 'wb'))
```

# CHAPTER 10
# Conclusion and Future Scope

We have built a portal and application for the Farmers in collaboration with maker's lab which would help them to get assistance regarding crop cultivation strategies, primarily for the services of prediction, recommendation, and information of crops, weather forecast and soil analysis. In Future we recommend optimum decisions for crops to be cultivated by farmers based on several parameters and help them make an informed decision throughout cultivation such as disease prediction, irrigation management, pesticides to be used, locust attacks, emergency weather alerts, and many more.

# APPENDIX A
# REFERENCES

[1] Multi-Temporal Satellite Images on Topsoil Attribute Quantification and the Relationship with Soil Classes and Geology, Department of soil Science, College of Agriculture Luiz de Queiroz, University of São Paulo, Rua Pádua Dias, 11, Piracicaba, Cx Postal 09, São Paulo, CEP 13416900, Brazil;

[2] Remote sensing applications, P. S. Roy, R. S. Dwivedi, D. Vijayen, "national RemoteSensingCenter,Isro". https://www.nrsc.gov.in/sites/default/files/pdf/ebooks/Chap_4_Soils.pdf

[3] Soil mapping approach in GIS using Landsat satellite imagery and DEM data. https://www.researchgate.net/publication/228671468_Soil_mapping_approach_in_GIS_using_Landsat_satellite_imagery_and_DEM_data

[4] Soil mapping approach in GIS using Landsat satellite imagery and DEM data Erturul Aksoy, Gökhan Özsoy* and M. Sabri Dirim Uludag University, Faculty of Agriculture, Department of Soil Science, 16059 Görükle-Bursa, Turkey.

[5] Soil Type Classification and Mapping using Hyperspectral Remote Sensing Data, Amol D. Vibhute1 *, K. V. Kale1 Senior Member IEEE 1. Dept. of Computer Science & IT, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad-431004 (MS), India *amolvibhute2011@gmail.com, kvkale91@gmail.com Rajesh K. Dhumal1, 2, S. C. Mehrotra1, 2 2. Srinivasa Ramanujan Geospatial Chair, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad-431004 (MS), India dhumal19@gmail.com, mehrotra.suresh15j@gmail.com

[6] Application of Satellite Remote Sensing to find Soil Fertilization by using Soil Colour Study Area Vellore District, Tamil Nadu, India. http://dx.doi.org/10.3991/ijoe.v9i2.2530 N. Suresh Kumar, S. Margret Anouncia, M. Prabu VIT University, Vellore, Tamil Nadu

[7] https://ncert.nic.in/ncerts/l/kegy106.pdf

[8] https://www.tensorflow.org/tutorials/keras/classification

[9] https://www.analyticsvidhya.com/blog/2020/09/integrating-machine-learning-into-web-applications-with-flask/

[10]https://towardsdatascience.com/building-a-web-application-to-deploy-machine-learning-models-e224269c1331

[11] https://www.ijsr.net/archive/v4i4/SUB153091.pdf

[12] https://www.kaggle.com/thammuio/all-agriculture-related-datasets-for-india

[13] https://www.kaggle.com/aniketng21600/crop-damage-information-in-india

[14] https://www.ijsr.net/archive/v4i4/SUB153091.pdf

# APPENDIX B
# PLAGIARISM REPORT

Plagiarism Report: given document is 92% unique

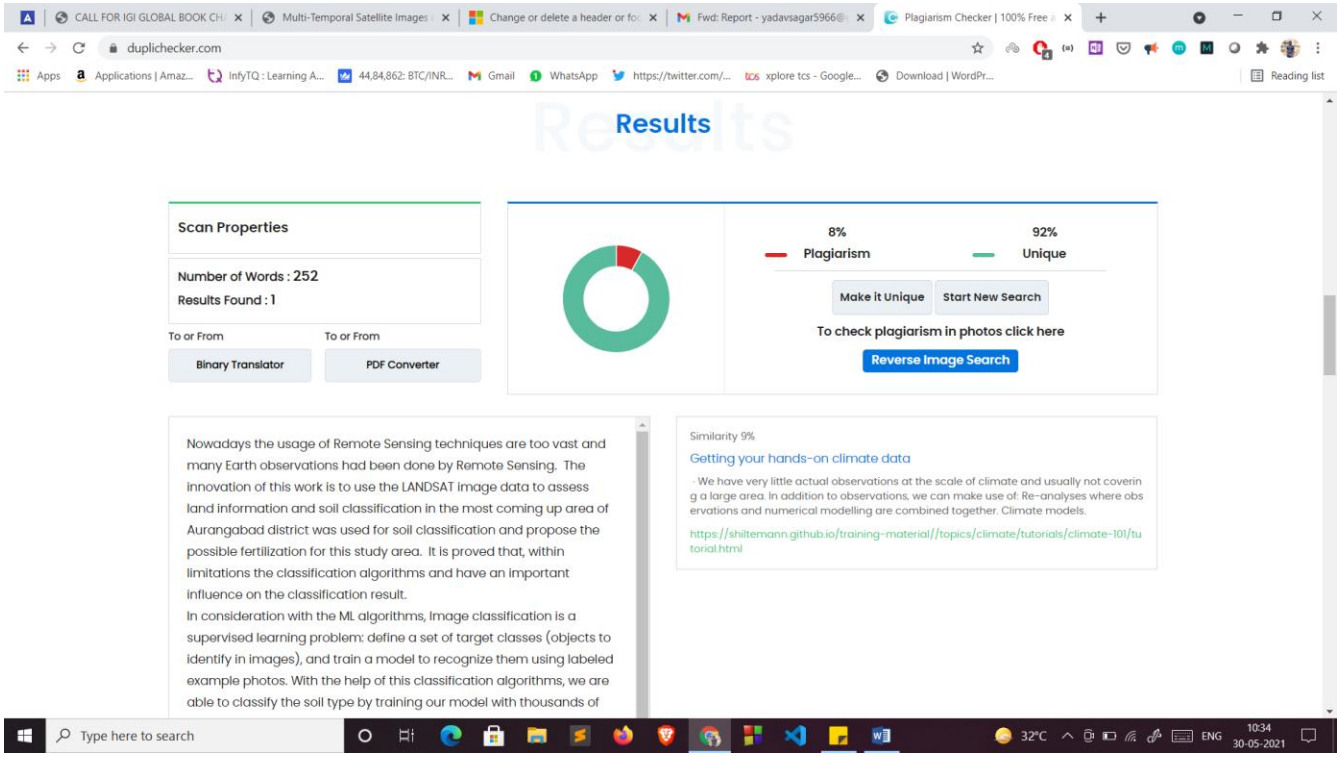


Figure a: Plagiarism Report