Crop Disease Detection and Soil Analysis

Department : Information Technology
Group ID: 18

Group Members:
1. Shreeraj Rane (16010420086)
2. Surin Shah (16010420087)
3. Mayur Bhore (16010420090)
4. Shreyas Panchikattil (16010420096)

Group Mentor:
Prof. Chirag Desai

SOMAIYA VIDYAVIHAR UNIVERSITY
K J Somaiya College of Engineering

Somaiya TRUST

Problem Definition

- This project aims to develop a computer vision-based website that can automate the identification of diseases of crop and fruit. The objective is to provide farmers with an efficient and accurate tool to protect their crops and minimize potential losses.
- Additionally, the project proposes the use of machine learning algorithms to estimate agricultural output, which can assist farmers in selecting suitable crops based on soil type and weather conditions.
- The website aims to provide a user-friendly portal that offers farmers a one-stop-shop for information about crop selection and soil type, facilitating informed decision-making.
- Ultimately, the project aims to improve crop yield and increase profitability for farmers by providing them with advanced technology and data-driven insights.

Background Work

Sr. No	Title	Authors	Year	Abstract
1	OPNN: Optimized Probabilistic Neural Network based Automatic Detection of Maize Plant Disease Detection	E. Akanksha, N. Sharma and K. Gulati.	2021	The proposed method uses an optimized probabilistic neural network (OPNN) to classify the maize plant images into healthy or diseased categories. The OPNN model is optimized using a particle swarm optimization (PSO) algorithm to improve its accuracy. Experimental results demonstrate that the proposed OPNN model achieves high accuracy in detecting maize plant diseases, outperforming traditional machine learning algorithms.
2	Plant diseases detection using image processing techniques	S. K. Tichkule and D. H. Gawali	2016	This project provides a survey of various methods used for plant leaf disease detection using image processing techniques. The k-nearest-neighbor method is simple but slow, while neural networks are tolerant to noise but have complex structures. Support vector machines (SVM) were found to be competitive with other machine learning algorithms for classifying high-dimensional data sets.

Background Work

Sr. No	Title	Authors	Year	Abstract
3	Identification of Maize Leaf Diseases Using Improved Deep Convolutional Neural Networks	X. Zhang, Y. Qiao, F. Meng, C. Fan and M. Zhang.	2018	This research paper proposes an improved CNN model to identify maize leaf diseases. The model consists of convolutional and pooling layers, fully connected layers, and a soft-max classifier. To enhance performance, techniques like data augmentation, dropout regularization, and transfer learning were applied. Experimental results show that the proposed model achieved an accuracy of 96.5%, outperforming several deep learning and traditional machine learning models.
4	Soil Analysis Using Visible and Near Infrared Spectroscopy	Wetterlind, Johanna & Stenberg, Bo & Viscarra Rossel, Raphael.	2013	Soil nutrient content, including nitrogen, phosphorus, and potassium were estimated using near-infrared (NIR) spectroscopy and machine learning algorithms. The authors achieved good prediction accuracy using partial least squares regression (PLSR) and random forest algorithms.

Background Work

Sr. No	Title	Authors	Year	Abstract
5	Soil analysis and crop fertility prediction using machine learning	Jagdeep Yadav, Shalu Chopra, Vijayalakshmi M	2021	A model for predicting soil fertility, crop prediction, and crop yield based on soil and crop datasets from the Indian region using machine learning algorithms. The model was tested using different algorithms and found that the Artificial Neural Network (ANN) had the highest accuracy in all three predictions. .
6	Soil moisture prediction using machine learning	S. Prakash, A. Sharma and S. S. Sahu	2018	This paper presents a study on the use of machine learning methods to predict soil moisture in advance. The study utilized multiple linear regression, support vector regression, and recurrent neural network models for the prediction. The results showed that multiple linear regression outperformed support vector regression and recurrent neural network.

Scope (Functional Requirements)

Disease Detection and Diagnosis

The website should be able to analyze images of crops and accurately identify the presence of diseases.

Preventive Options

The website should be able to provide information on the type of disease identified and recommend appropriate preventive options.

Smarter Crop Selection

The website should provide farmers with information on crop selection based on soil type and weather conditions.

Scope (Non-Functional Requirements)

Usability

The system should have an intuitive and easy to use interface, allowing farming farmers to navigate and interact with the system without extensive training.

Scalability and Performance

The website should be able to handle a large volume of data and user requests without experiencing downtime or performance issues.

Speed and Responsiveness

The website should be fast and responsive to ensure a seamless user experience.

Availability

The system ashould have a high level of availability , ensuring minimal downtime to provide uninteruppted access to farmsers.

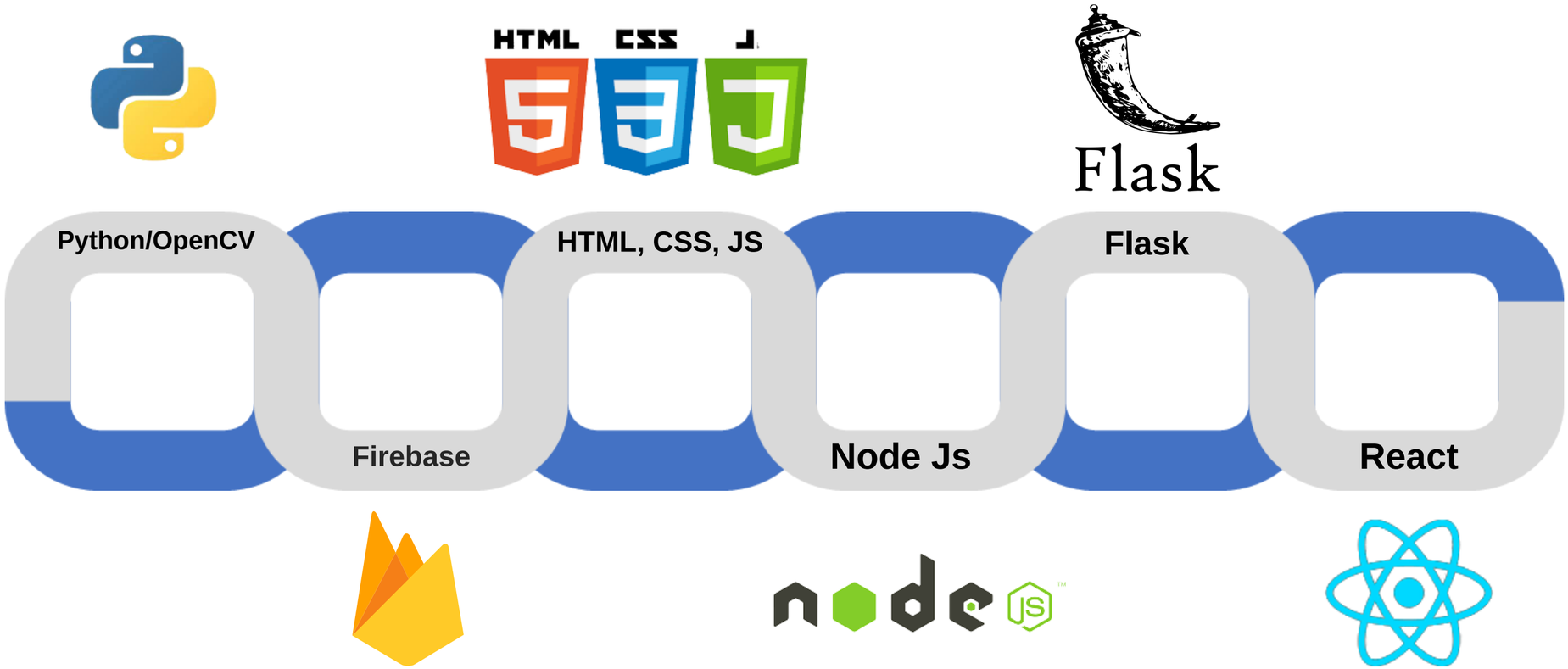
Data Set Description

- The Maize and Corn Disease dataset, available on Kaggle, contains images of healthy maize and corn plants, as well as plants infected with four different diseases: Gray Leaf Spot, Common Rust, Northern Leaf Blight, and Healthy. The dataset consists of a total of 4,000 images, with 1,000 images for each class. The images were collected from different regions in Uganda, Africa. The dataset can be used for developing and testing machine learning models for the detection and classification of maize and corn diseases.
  Link:https://www.kaggle.com/datasets/smaranjitghose/corn-or-maize-leaf-disease-dataset

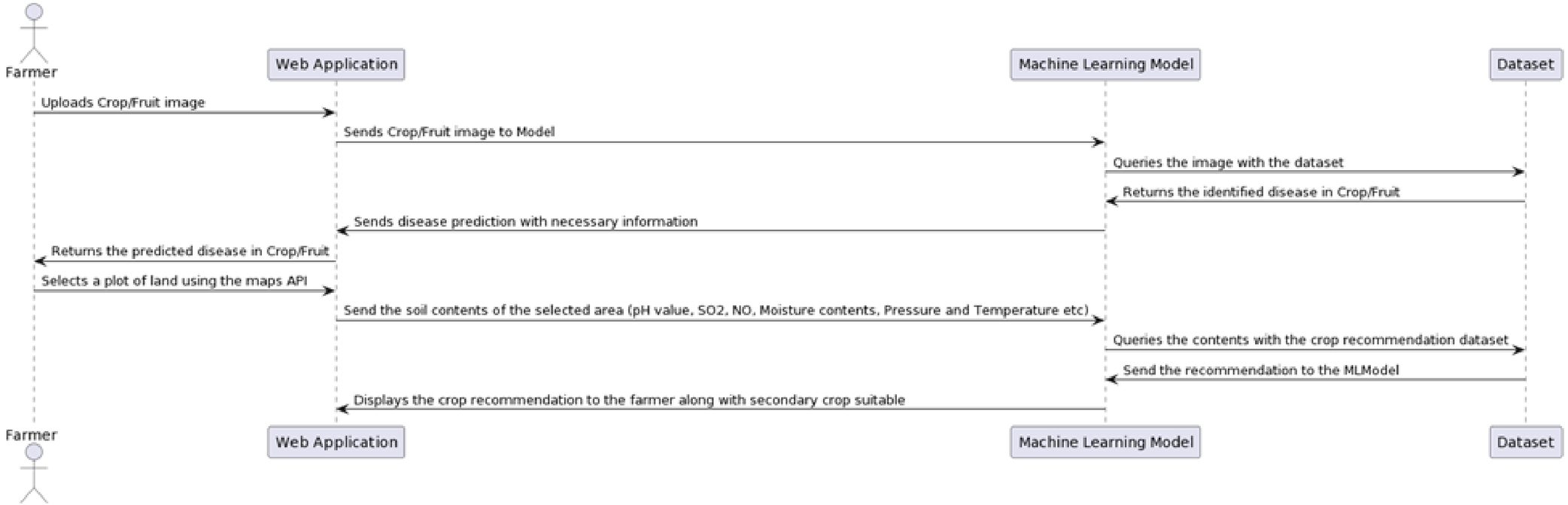
- The dataset is about precision agriculture and aims to help farmers make informed decisions about the most suitable crops to grow in a particular farm based on various parameters. The dataset includes parameters such as N, P, and K, which are ratios of Nitrogen, Phosphorous, and Potassium content in the soil, as well as temperature, humidity, pH value of the soil, and rainfall in millimeters. The data was built by augmenting datasets of rainfall, climate, and fertilizer data available for India.
  Link: https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset

SOMAIYA VIDYAVIHAR UNIVERSITY
K J Somaiya College of Engineering

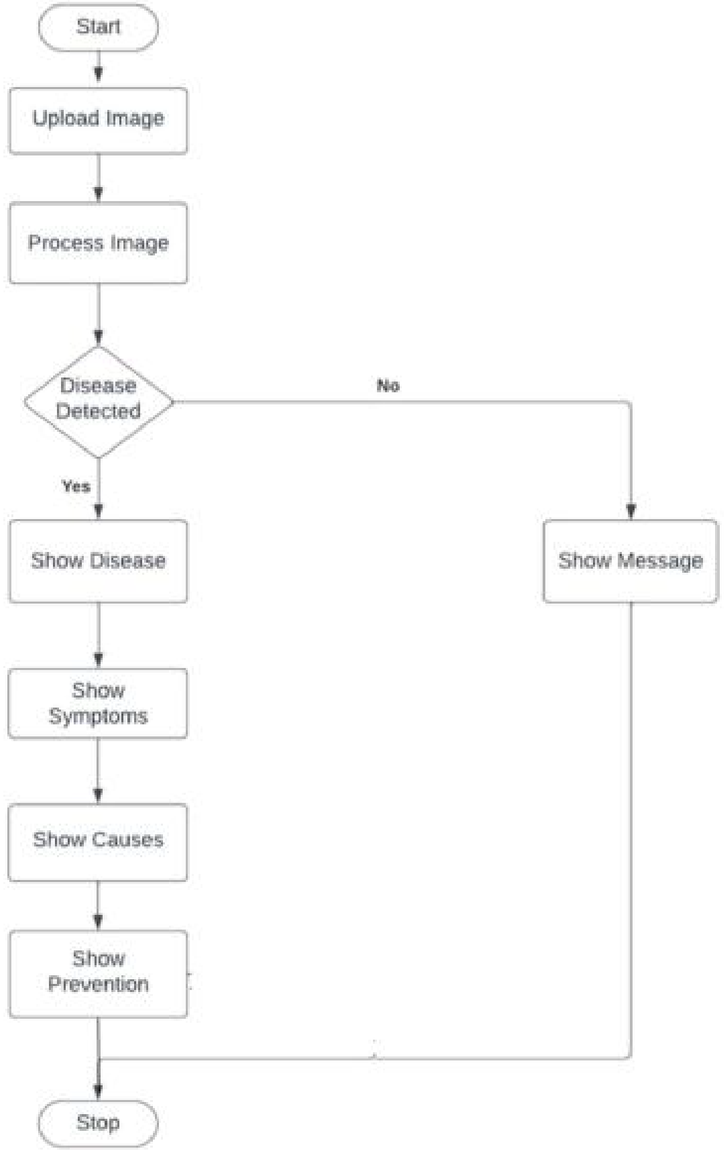
Technology used



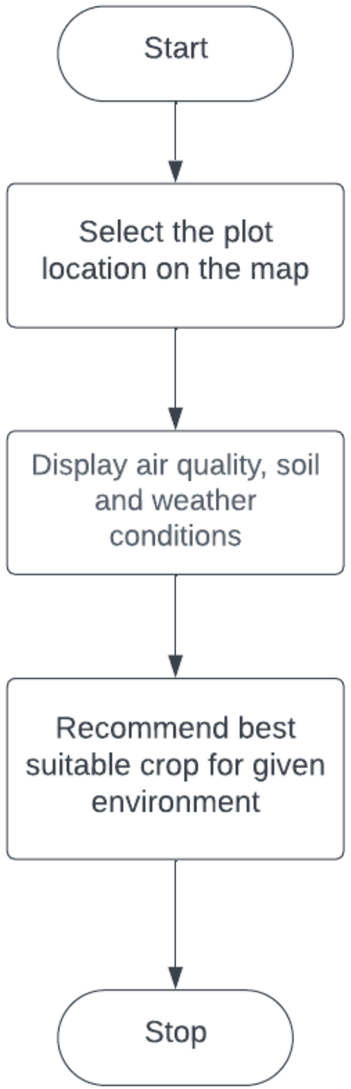
Overview of Implementation- Input and output of the system, Block Diagram/Work Flow Diagram/ Module Diagram



Disease Detection System Flowchart



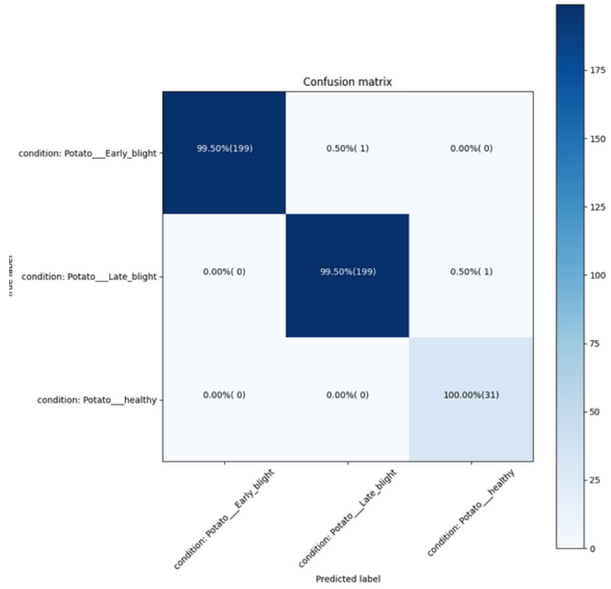
Soil Analysis Flowchart



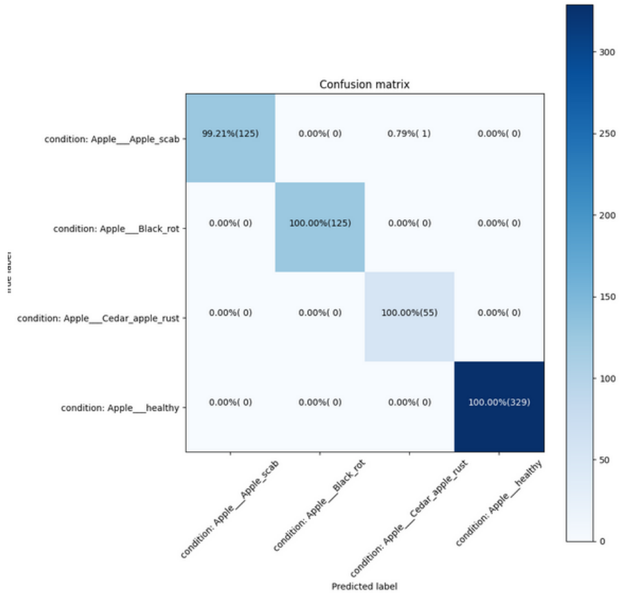
Result and Discussion

- **Accurate detection of crop diseases: The machine learning model has been taught to identify numerous diseases that affect the maize crop and fruits. To analyze crop imagery and display the results appropriately, we used computer vision. The model's high level of recall and precision enables farmers to make quick corrections.**

Corn/Maize



Potato



Apple

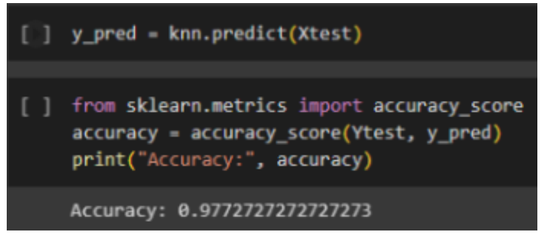
SOMAIYA VIDYAVIHAR UNIVERSITY
K J Somaiya College of Engineering

Somaiya TRUST

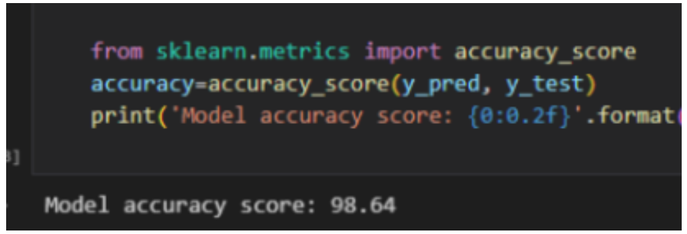
Result and Discussion

- **Soil analysis recommendations: The technology can recommend farmers on the best crops to plant on a particular area of land. The advice will be based on a number of factors, including the amount of soil moisture, the air quality index, the pH level, the presence of particular chemicals, etc. Farmers will have a range of options that, given the conditions, are suited for growing a variety of crops.**

We Compared two algorithms we tested which were KNN and Gradientboost Machine



KNN Accuracy = 97.72%



Gradientboost machine
Accuracy = 98.64%

SOMAIYA VIDYAVIHAR UNIVERSITY
K J Somaiya College of Engineering

Somaiya TRUST

Conclusion

- The project of designing a computer vision-based website that automates disease identification in maize cultivation and using machine learning to estimate agricultural output has the potential to revolutionize the way farmers manage their crops, improve food security, and contribute to sustainable agriculture.

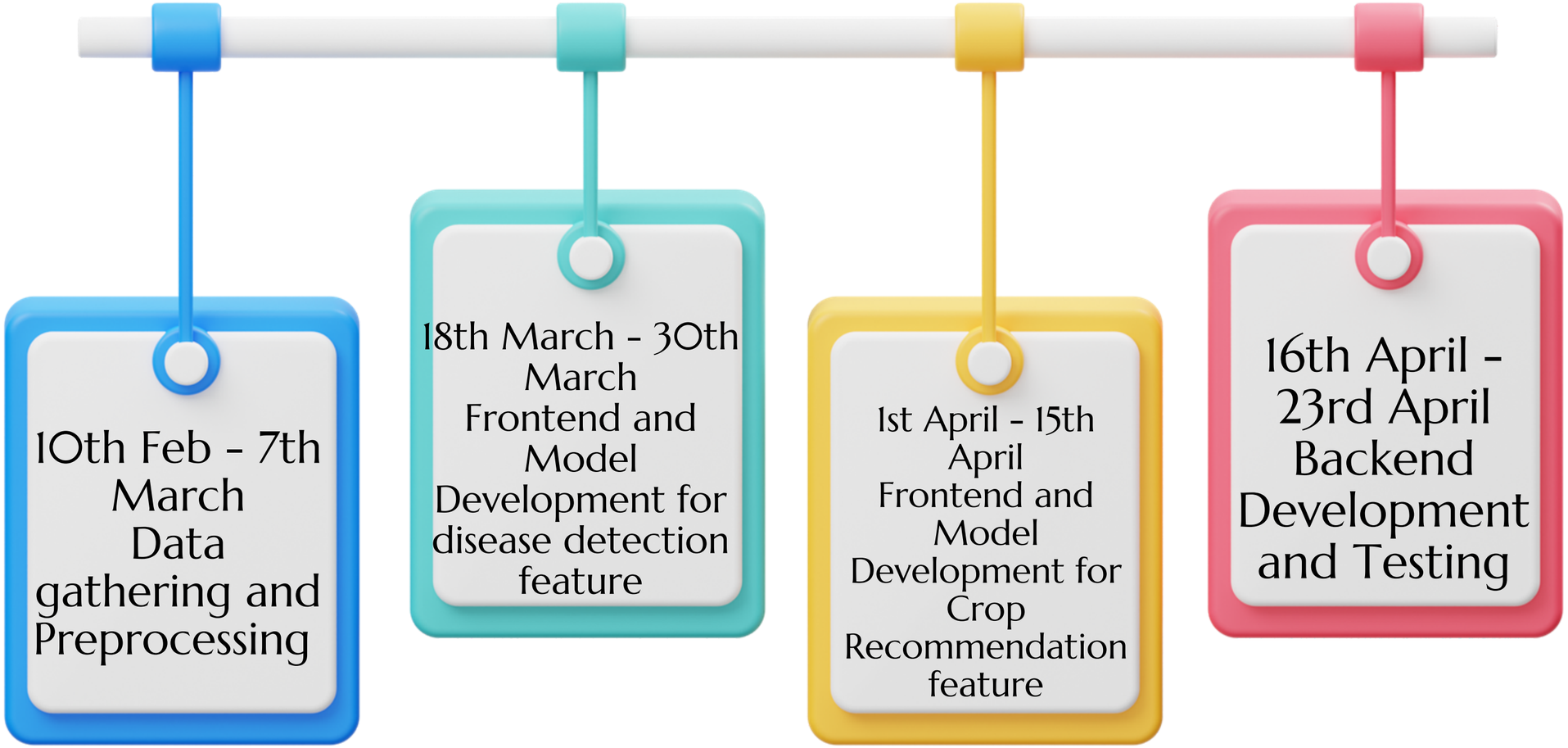
- The development of this website has several benefits, such as reducing the time and cost associated with traditional disease diagnosis methods, enabling timely and appropriate actions to protect crops, providing farmers with valuable information on crop selection and soil type, and improving agricultural output and profitability.

Learning

- In terms of learning, this project highlights the power of computer vision and machine learning in addressing real-world problems and how technology can be harnessed for the greater good. The project also underscores the importance of user-centered design, security, scalability, cross-platform compatibility, speed, and easy maintenance in building a successful and sustainable website.

- Overall, this project can serve as a blueprint for developing similar applications in other areas of agriculture, healthcare, and other industries, where technology can be used to solve complex problems and improve people's lives.

Implementation schedule of Mini-Project across Sem VI



SOMAIYA VIDYAVIHAR UNIVERSITY
K J Somaiya College of Engineering

Somaiya TRUST

Thank You!!!