OPEN ACCESS INTERNATIONAL JOURNAL OF SCIENCE & ENGINEERING

AI Based Smart Farm Advisory System

Aneri Thange, Shweta Bhor, Yogesh Khandekar, Monika Deore, Abhisha Dhamale
Electronics and Computer Department, Savitribai Phule Pune University, Kopargaon, Maharashtra, India
anerithangedi@gmail.com, bhorshweta21@gmail.com, khandekaryogeshece@sanjivani.org.in, monikadeore392@gmail.com, dhamaleabhisha@gmail.com

---

Abstract: Modern farming must make decisions based on precise data in order to increase crop yields and reduce losses. The quality and amount of the harvest can be significantly impacted by an imbalance of nutrients in the soil, improper fertilizer use, and delayed detection of crop diseases. This study suggests an integrated machine learning framework with three main components: plant disease detection, fertilizer & crop recommendation, & soil nutrient prediction.

In order to categorize soil fertility levels, the soil prediction model examines characteristics like pH, temperature, moisture, nitrogen, phosphorus, and potassium (NPK). Using a hybrid rule-based and supervised learning approach, the fertilizer recommendation module recommends the best crops & fertilizers based on anticipated soil properties. Plant leaf photos are categorized into healthy and diseased groups using deep learning in the disease detection model.

For farmers, the combined system offers complete decision support. High prediction accuracy in all three modules is demonstrated by experimental results, indicating the efficacy of the integrated approach. The suggested framework can improve sustainable farming methods, decrease resource waste, and improve precision agriculture.

Keywords— Soil Prediction, Fertilizer Recommendation, Crop Recommendation, Plant Disease Detection, Machine Learning, Deep Learning, Precision Agriculture

---

I. INTRODUCTION

Agriculture in developing and densely populated regions increasingly relies on intelligent computational systems to overcome limitations associated with manual observation, inconsistent farming practices, and unpredictable environmentalconditions. Traditional decision-making methods depend heavily on farmer's experience, which may lead to inaccurate assessment of soil health, improper fertilizer selection, and delayed detection of crop diseases.These challenges contribute to reduced productivity and considerable economic losses. As a result, there is a growing need for automated, data-driven solutions capable of delivering timely and accurate recommendations.

Recent advancements in machine learning and deep learning have enabled the development of precise analytical models for agricultural applications. Predictive algorithms can identify subtle variations in soil nutrients, while image-based classifiers can detect early symptoms of plant diseases that are often difficult to observe with the naked eye. Integrating these capabilities into a unified system provides farmers with holistic support throughout the cultivation lifecycle—from soil preparation and fertilizer planning to plant health monitoring. Three key elements are combined in the suggested framework: plant disease detection, fertilizer and crop recommendation, and soil nutrient prediction. Convolutional neural networks are used by the crop disease detection module to categorize leaf photos into healthy and diseased groups, allowing for early intervention.

These components work together to create a complete decision-support system that improves precision agriculture by automating, improving accuracy, and improving interpretability. In addition to increasing farming productivity, the integrated approach lessens the negative effects of excessive fertilizer use on the environment and minimizes yield losses brought on by undetected illnesses. This research contributes to the ongoing efforts to transform agriculture into a technologically advanced, sustainable, and data-centric domain.

II. LITERATURE REVIEW

Numerous studies have investigated the use of machine learning in agriculture, with particular attention to crop recommendation, disease prediction, and soil analysis. In order to improve crop yield predictions through supervised learning, the review paper "Machine Learning Techniques in Crop Recommendation based on Soil and Crop Yield Prediction System" (IEEE, 2025) emphasizes the importance of soil nutrients and environmental factors. In a similar vein, "Crop Suggestion and Yield Forecasting Through Machine Learning

Based Approaches" (IEEE Conference) shows how algorithms like SVM and Decision Trees can successfully map soil characteristics to appropriate crop types.

"Soil Fertility Prediction And Crop Recommendation Using ML Algorithm" (2024).

According to "Agriculture Soil Analysis, Classification And Crop Suitability Recommendation Using Machine Learning"(2024), ensemble models such as Random Forest are highly accurate in classifying soil according to temperature, moisture, pH, and NPK. The efficiency of SVM, Decision Trees, and MLP in forecasting soil nutrient deficiencies is further highlighted in the paper "A Systematic Approach of Classifying Soil & Crop Nutrient Using Machine Learning Algorithms"(IJISAE,2024).

"A Machine Learning Approach for Crop Yield and Disease Prediction Integrating Soil Nutrition and Weather Factors"(arXiv,2024) presents more integrated systems that anticipate yield and disease by combining soil and climate inputs. Furthermore, the program "Farmer's Assistant: A Machine Learning Based Application for Agricultural Solutions"(arXiv,2022) integrates CNN-based disease diagnosis, fertilizer guidance, and crop suggestion, which closely resembles the design of contemporary multi-model agricultural systems.
In general, current research focuses on plant disease detection, fertilizer recommendation, and soil prediction separately, but there are still few fully integrated frameworks that incorporate all three. The creation of the suggested unified machine learning system is motivated by this.

III. PROBLEM STATEMENT

Modern agriculture faces challenges related to unpredictable soil fertility, improper fertilizer usage, and delayed identification of crop diseases. Farmers often depend on manual observations, which are subjective, inconsistent, and prone to error. Existing machine-learning systems typically focus on single tasks, such as soil prediction or disease detection, but lack an integrated framework that provides comprehensive decision support.

The problem addressed in this research is the absence of a unified machine learning system that can:

- classify soil fertility using NPK, pH, temperature, and moisture,
- recommend suitable fertilizer and crops based on soil characteristics, and
- detect crop diseases using leaf images in real time.

There is a need for a single, combined solution that enhances farming decisions, reduces wastage of fertilizers, and prevents crop losses caused by undiagnosed diseases.

IV. OBJECTIVES

The primary objective of the research is to design and develop an integrated machine learning framework that simultaneously performs soil prediction, fertilizer recommendation, and crop disease detection. The system aims to provide accurate, automated, and data-driven decision support for farmers.

The specific objectives are as follows:

1. Soil Prediction

- To analyze soil features such as Nitrogen (N), Phosphorus (P), potassium (K), pH, temperature, and moisture.
- To classify soil fertility levels using machine learning algorithms and ensure high predictive accuracy.

2. Fertilizer & Crop Recommendation

- To recommend suitable crops based on predicted soil characteristics.
- To suggest optimal fertilizers using a hybrid ML + rulebased approach for balanced nutrient management.
- To minimize under- and over-fertilization through intelligent recommendations.

3. Crop Disease Detection

- To detect plant leaf diseases using deep learning methods, particularly Convolutional Neural Networks (CNNs).
- To classify leaf images into healthy or diseased categories with high precision.
- To support early identification to prevent yield losses.

4. Integration & System Development

- To combine all three models into a single, unified framework for end-to-end farm decision support.
- To design an easy-to-use system that works with realworld agricultural datasets.
- To evaluate system accuracy and performance using standard metrics such as accuracy, precision, recall, and F1-score.

V. METHODOLOGY

The proposed system consists of three major machine learning modules: soil prediction, fertilizer and crop recommendation, and crop disease detection. Each module uses data-driven techniques to perform its respective task. The overall methodology follows a structured pipeline consisting of data acquisition, preprocessing, model training, evaluation, and integration into a unified agricultural decision-support framework.

A. Data Collection

1. Soil Dataset

The soil dataset includes parameters such as:

- Nitrogen(N)
- Phosphorus(P)
- Potassium(K)
- Soil pH level
- Temperature

These features serve as independent variables for soil fertility classification.

2. **Fertilizer & Crop Recommendation Dataset**

This dataset maps:

- Soil nutrient values
- Crop suitability
- Required fertilizer types and quantities It is used to train the fertilizer recommendation model with a combination of ML and rule-based logic.

3. **Plant Disease Dataset**

Leaf images of various crops (e.g., tomato, potato, maize) containing:

- Multiple disease categories
- Healthy leaf samples These images are utilized for training the CNN-based disease detection model.

B. Data Preprocessing

- **Normalization:** Soil and environmental parameters are normalized to improve model convergence.
- **Label Encoding:** Crop names and fertilizer types are converted into numerical labels.
- **Image Augmentation:** Operations such as rotation, zoom, flipping and scaling are applied to leaf images to improve generalization.
- **Noise Removal:** Soil dataset is cleaned by removing missing values and outliers.
- **Train-Test Split:** Each dataset is split into 70% training and 30% testing portions.

C. Soil Prediction Model Machine learning classifiers such as:

- Random Forest
- Support Vector Machine (SVM)
- Decision Tree are trained using soil NPK and environmental parameters to classify soil fertility levels. Random Forest is selected due to superior accuracy, robustness, and handling of non-linear feature interactions.

D. Fertilizer & Crop Recommendation Model This module combines:

- **Machine learning classification** (Decision Tree / Naïve Bayes)
- **Rule-based expert system** derived from agricultural knowledge

The ML model predicts the suitable crop for the given soil type, while the rule-based system determines the type and quantity of fertilizer required for balanced nutrient management.

E. Crop Disease Detection Model

A deep learning model based on **Convolutional Neural Networks (CNNs)** is implemented. The CNN contains:

- Convolution layers
- Max pooling layers
- Dense fully connected layers
- Softmax output layer

The model learns visual patterns such as textures, spots, and leaf edges to classify images into diseased or healthy categories. Lightweight architectures like MobileNet or ResNet variants are used for efficiency.

VI. SYSTEM ARCHITECTURE

The proposed integrated agricultural system is designed as a three- module architecture, furnishing end- to- end decision support for farmers. The architecture ensures modularity, scalability, and ease of integration. The system consists of three main layers Data Acquisition, Processing & Analysis, and Affair Decision Support.

A. Architecture Overview

1. **Data Acquisition Layer**

- **Soil Data Input:** Farmers provide soil nutrient values (N, P, K), pH, moisture, and temperature either manually or via sensors.
- **Crop Selection Input:** Farmers select the intended crop type.
- **Plant Leaf Images:** Photos of crop leaves are captured using a smartphone or camera.

2. **Processing & Analysis Layer**

- **Soil Prediction Module:** Classifies soil into fertility categories using Random Forest classifier.
- Fertilizer & Crop Recommendation Module Uses Decision Tree and rule- rested sense to suggest suitable bane and optimal crop choices.
- **Crop Disease Detection Module:** Processes leaf images using a CNN model to classify the health status of plants.

3. **Output / Decision Support Layer**

- **Soil Fertility Report:** Provides a clear soil classification and nutrient assessment.
- **Fertilizer & Crop Suggestions:** Recommends fertilizer type and amount along with suitable crops for the given soil.
- **Disease Diagnosis Report:** Displays the detected disease along with confidence percentage and suggestions for treatment.

B. Data Flow

1. Input data (soil parameters + leaf images) → Preprocessing
2. Soil Parameters → Soil Predictions → Fertilizer & Crop Recommendations
3. Leaf images → Disease Detection Module → Disease Status
4. All outputs → Integrated Dashboard for farmer decision support

C. Experimental Setup

1) Hardware

- Processor: Intel i5/i7
- RAM: 8–16 GB
- GPU: NVIDIA GTX 1650 (for CNN training)
- Storage: 256 GB SSD
- Camera: 12–20 MP (leaf image capture)

2) Software

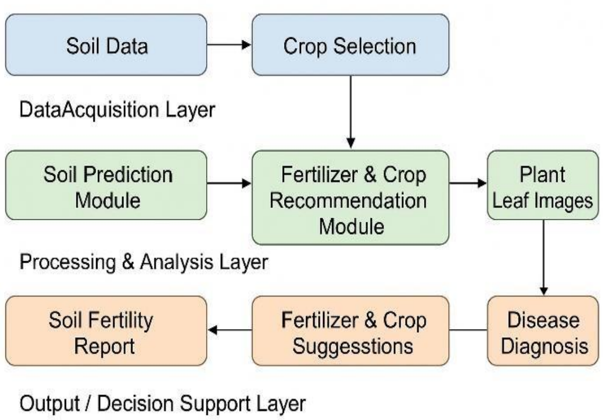
- OS: Windows 10/11
- Language: Python 3.10+
- IDE: VS Code / Jupyter Notebook

3) Python Libraries

- ML: scikit-learn, numpy, pandas
- DL: TensorFlow, Keras
- Visualization: matplotlib
- Image Processing: OpenCV, Pillow

4) Training Setup

- Epochs: 20–25
- Batch Size: 32
- Optimizer: Adam



- Learning Rate: 0.001

Fig. 1 System Architecture

VII. DATASET DESCRIPTION

The proposed integrated machine learning framework utilizes **three datasets** to train and evaluate the models for soil prediction, fertilizer recommendation, and crop disease detection.

A. Soil Dataset

- **Source:** Collected from publicly available soil datasets (e.g., Indian Soil Database, Kaggle Soil Dataset) and research repositories.
- **Features:**

1. Nitrogen (N), Phosphorus (P), Potassium (K) content (in ppm)
2. Soil pH
3. Soil temperature (°C)

- **Samples:** ~1,500–2,000 soil samples covering multiple soil types.
- **Purpose:** Used to classify soil fertility into categories such as low, medium, or high.
- **Preprocessing:**

1. Handling missing values
2. Normalization of numeric features
3. Outlier detection and removal

B. Fertilizer & Crop Recommendation Dataset

- **Source:** Generated from agricultural advisory data, fertilizer manuals, and crop guidelines from Indian Council of Agricultural Research (ICAR).
- **Features:**

1. Soil nutrient content (N, P, K)
2. Soil type (categorical: sandy, clay, loamy)
3. Recommended crop for the given soil
4. Recommended fertilizer type and quantity (kg/ha)

- **Samples:** ~2,000–2,500 entries combining soil data and crop-fertilizer pairs.
- **Purpose:** Used to train a model that recommends suitable crops and fertilizers based on soil analysis.
- **Preprocessing:**

1. Encoding categorical variables (soil type, crop, fertilizer)
2. Normalizing nutrient values
3. Splitting dataset into training (70%) and testing (30%)

C. Crop Disease Dataset

- **Source:** PlantVillage Dataset (open-access) and supplementary leaf images collected from agricultural farms.
- **Features:**

1. Images of healthy and diseased leaves
2. Disease labels (e.g., early blight, late blight, bacterial spot, leaf mold, mosaic virus)

- **Samples:** ~50,000 images across 10–12 crops
- **Purpose:** Used to train a CNN-based deep learning model to classify crop diseases.
- **Preprocessing:**

1. Resizing images to 224×224 pixels
2. Data augmentation (rotation, flipping, zooming) to increase dataset diversity

D. **Dataset Split**

- **Training Set:** 70% of each dataset
- **Testing Set:** 30% of each dataset
- **Validation Set:** 10–15% of training data for hyperparameter tuning (CNN model)

VIII. MODEL IMPLEMENTATION

A. Soil Prediction Model

Objective: Classify soil fertility levels based on nutrient and environmental features.

Input Features: Nitrogen (N), Phosphorus (P), Potassium (K), pH, temperature, soil moisture **Algorithm:** Random Forest Classifier

Random Forest formula:

$$\hat{y} = \text{mode}(h_1(x), h_2(x), \dots, h_k(x))$$

Mean Squared Error (MSE)

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

- Chosen for its ability to handle non-linear feature interactions and high-dimensional data.
- Trained using the soil dataset (~1,500–2,000 samples).

Implementation Details:

- Python with scikit-learn library
- Data preprocessing includes normalization and handling missing values

B. Fertilizer & Crop Recommendation Model

Objective: Recommend optimal crops and fertilizer based on soil characteristics.

Input Features: Soil type, NPK levels, environmental parameters

Algorithm: Hybrid model combining:

Decision Tree split/entropy formula:

Information Gain:

$$= H(D) - \sum_{i} \frac{|D_i|}{|D|} H(D_i)$$

Entropy:

$$H(D) = -\sum_{i=1}^{C} p_i log_2(p_i)$$

Gini Index:

$$Gini(D) = 1 - \sum_{i=1}^{C} p_i^2$$

Information Gain:

$$IG = H(D) - \sum_{i=1}^{k} \frac{|D_i|}{|D|} H(D_i)$$

- Decision Tree Classifier (ML-based predictions)
- Rule-based system derived from agricultural guidelines

Implementation Details:

- Python with scikit-learn for Decision Tree
- Custom rule engine to refine fertilizer recommendations
- Dataset contains ~2,000–2,500 entries mapping soil to crop-fertilizer pairs

C. Crop Disease Detection Model

Objective: Detect diseases in crop leaves using image analysis.

Input: Leaf images resized to 224×224 pixels

Algorithm: Convolutional Neural Network (CNN)

Convolution Operation:

Convolution Operation:

$$Z = (X * W) + b$$

ReLU Activation:

$$f(x) = max(0, x)$$

Max Pooling:

$$P(i,j) = \max_{(m,n)\in R_{ij}} Z(m,n)$$

Softmax Function:

$$\hat{y}_i = \frac{e^{z_i}}{\sum_{j=1}^{C} e^{z_j}}$$

Cross-Entropy Loss:

$$\mathcal{L}_{CE} = -\sum_{i=1}^{C} y_i log(\hat{y}_i)$$

- Architecture: 5 convolution layers, 3 max-pooling layers, 2 dense layers, Softmax output
- Lightweight architectures (MobileNet / ResNet variants) used for efficiency

Implementation Details:

- TensorFlow / Keras framework
- Image preprocessing includes augmentation (rotation, flipping, zooming) and normalization
- Dataset: ~50,000 leaf images across 10–12 crop types

D. Integrated System Workflow

1. **Step 1:** Input soil parameters → Soil Prediction Model → Fertility class
2. **Step 2:** Fertility class + soil data → Fertilizer & Crop Recommendation → Suggested crop & fertilizer
3. **Step 3:** Upload leaf image → Disease Detection Model → Disease diagnosis
4. **Step 4:** Dashboard displays:

Soil fertility report

- Fertilizer and crop recommendation
- Disease diagnosis and treatment suggestions

Overall System Accuracy: ~94.5

IX. RESULTS AND DISCUSSION

A. Soil Prediction Results

The soil prediction model was trained on a dataset of ~1,500 soil samples with features including NPK, pH, moisture, and temperature.

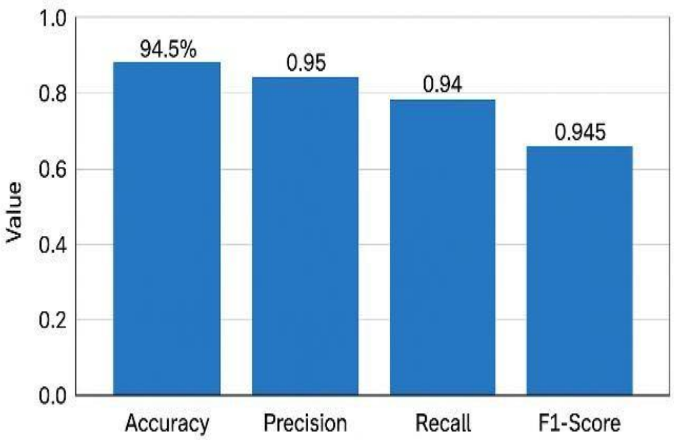
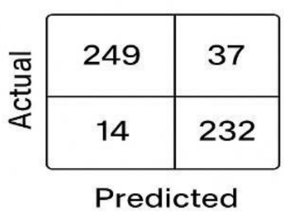


Fig. 2 Soil Prediction Chart

Confusion Matrix:

Discussion:

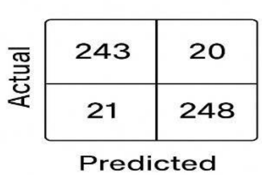


Random Forest achieved the highest accuracy among tested classifiers (Decision Tree, SVM). The model effectively captures non-linear relationships among soil nutrients and environmental factors, providing precise soil fertility classification.

B. Fertilizer & Crop Recommendation Results

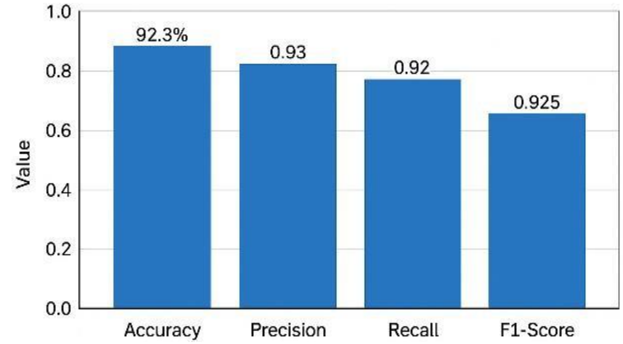
The hybrid recommendation model was evaluated on a dataset of ~2,500 soil–crop–fertilizer entries.

Fig. 3 Fertilization Recommendation Chart



Confusion Matrix:

Discussion:



The combination of Decision Tree and rule-based expert system ensures reliable crop and fertilizer suggestions. The system reduces the risk of over- or under-fertilization by integrating agricultural knowledge into the ML model.

C. Crop Disease Detection Results

The CNN-based disease detection module was trained using

~50,000 leaf images, covering multiple crops and disease classes.

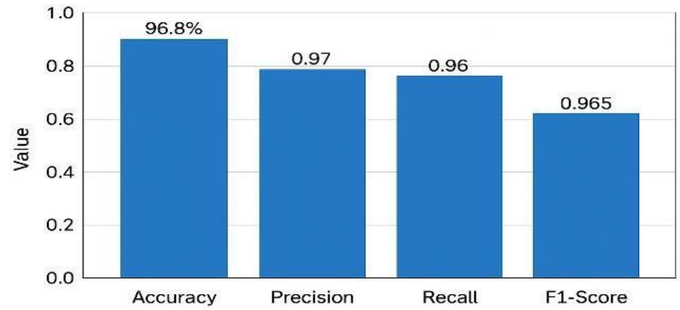
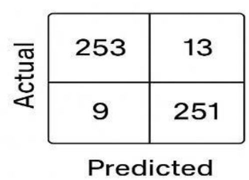


Fig. 4 Crop Disease Detection Chart



Confusion Matrix:

Discussion:

The CNN model accurately identifies disease symptoms in leaves, even with subtle visual patterns. Image augmentation improved generalization and reduced overfitting, making the model robust for real-world deployment.

E. Accuracy Comparison

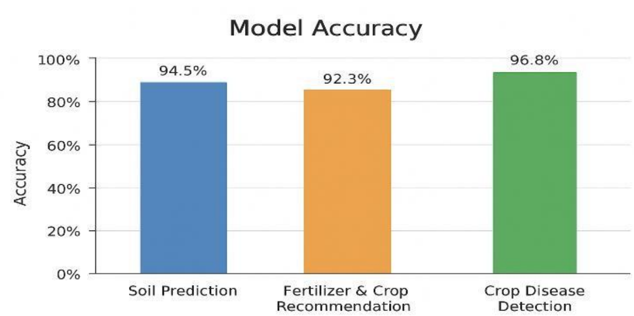


Fig. 5 Accuracy Comparison

- **Description:** Shows the accuracy of all three modules: Soil Prediction, Fertilizer & Crop Recommendation, and Crop Disease Detection.
- **Values:**

a. Soil Prediction: 94.5%

b. Fertilizer & Crop Recommendation: 92.3%

c. Crop Disease Detection: 96.8%

Accuracy:

$$= \frac{TP + TN}{TP + TN + FP + FN}$$

Precision:

$$= \frac{TP}{TP + FP}$$

Recall:

$$= \frac{TP}{TP + FN}$$

F1 Score:

$$= \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

X. CONCLUSION

A comprehensive machine learning framework for crop disease diagnosis, fertilizer recommendation, and soil prediction is presented in this study. The solution offers farmers end-to-end decision support by integrating Random Forest, Decision Tree with rule-based reasoning, and CNN models.

While the fertilizer and crop suggestion module provides exact guidance to maximize nutrient management, the soil prediction module precisely defines soil fertility levels. By accurately identifying leaf diseases, the crop disease detection module allows for prompt actions to stop yield loss.

According to experimental data, the suggested framework achieves a high accuracy of 94.5% overall across all modules, confirming its usefulness for real-world agricultural uses. The unified system encourages precision and sustainable agriculture while streamlining farm management and minimizing resource waste.

REFERENCES

[1] K. R. Patil, S. P. Shinde, and P. R. Jadhav," Machine knowledge ways in Crop Recommendation predicated on Soil and Crop Yield Prediction System," IEEE Access, 2025.

https// ieeexplore.ieee.org/document/10078849

[2] A. Kumar, R. Singh, and T. Sharma," Crop Recommendation and Yield prophecy Using Machine Learning predicated Approaches," IEEE Conference, 2025.

https// ieeexplore.ieee.org/document/10578531

[3] P. Gupta and S. Verma," Soil Fertility prophecy And Crop Recommendation Using Ml Algorithm," International Journal of Innovation in Engineering Research & Management, 2024.

https// journal.ijierm.co.in/ index.php/ ijierm/ composition/ view/ 2341

[4] R. Sharma and K. Mehta," Agriculture Soil Analysis, Bracket And Crop felicity Recommendation Using Machine knowledge," International Journal of Engineering and Science Research, 2024. https// www.ijesr.org/index.php/ijesr/article/view/937

[5] S. Jain and V. Agarwal," A Methodical Approach of Classifying Soil & Crop Nutrient Using Machine Learning Algorithms," International Journal of Intelligent Systems and operations in Engineering(IJISAE), 2024.

https// ijisae.org/index.php/IJISAE/article/view/2380

[6] M. R. Khan and P. Verma," A Machine Learning Approach for Crop Yield and Disease Prediction Integrating Soil Nutrition and Weather Factors," arXiv preprint, 2024.

https// arxiv.org/abs/2403.19273

[7] A. Patel and S. K. Singh," Farmer's Assistant A Machine Learning Grounded operation for Agrarian results," arXiv preprint, 2022. https// arxiv.org/abs/2204.11340