International Journal of Enhanced Research in Science, Technology & Engineering
ISSN: 2319-7463, Vol. 15 Issue 3, March-2026

Soil Nutrient Analysis and Crop Recommendation Using Machine Learning

Prof. Priti R. Vanmali[1], Vishal Dattatraya Doke[2], Prasad Dilip Khawale[3], Priya Manik Sakat[4]

[1]Project Guide, Navsahyadri Education Society's Group of Institutions (NESGI), SPPU
[2,3,4]Student, Navsahyadri Education Society's Group of Institutions (NESGI), SPPU

ABSTRACT

Agricultural output depends heavily on soil health, yet many farmers still use antiquated methodologies and labour-intensive, expensive, and frequently unreliable manual soil testing techniques. In order to facilitate data-driven agricultural decision-making, this study suggests a machine learning-based soil nutrient analysis and crop recommendation system. In order to evaluate soil fertility and suitability for different crops, the system focuses on assessing important soil factors like pH, moisture content, temperature, organic matter, phosphorus (P), potassium (K), nitrogen (N), and moisture content. To find intricate patterns and connections between soil nutrient levels and crop yield requirements, machine learning techniques are used.

To determine which crops are most suited for a particular soil profile, supervised learning models—such as Random Forest, Decision Tree, and Support Vector Machine—are trained on historical soil and crop information. To increase model accuracy and dependability, data preprocessing methods such feature selection, normalization, and handling missing values are used. Metrics including F1-score, recall, accuracy, and precision are used to assess the models' performance. By suggesting the best crops to increase production while reducing excessive fertilizer use, the suggested system seeks to improve agricultural output. The technology boosts farmers' profitability, encourages sustainable agricultural methods, and lessens its impact on the environment by offering prompt and precise advice. This study shows how machine learning may be successfully incorporated into precision farming, providing a scalable and affordable answer to contemporary farming problems.

Keywords: Soil Nutrient Analysis, Crop Recommendation System, Machine Learning, Precision Agriculture, Sustainable Farming

1. INTRODUCTION

1.1 Background of the Study

Within the field of computer engineering, namely in the domains of artificial intelligence (AI), machine learning (ML), data analytics, and smart systems, lies the project named "Soil Nutrient Analysis and Crop Recommendation Using Machine Learning." Traditional industries like agriculture have been greatly impacted by engineering advancements in recent years. Utilizing data-driven decision-making to increase production, sustainability, and efficiency, Smart Agriculture, also known as Precision Farming, is the result of the integration of computing technologies with agricultural methods.

Declining soil fertility, erratic weather patterns, rising food demand brought on by population increase, and a shortage of arable land are some of the issues facing agriculture worldwide. Despite the fact that soil health is crucial to crop growth, many farmers continue to choose their crops using traditional techniques and their own judgment. These conventional methods frequently fall short in terms of scientific correctness and fail to make good use of the data at hand in order to make well-informed decisions.

The use of intelligent systems that can evaluate agricultural data and offer useful insights is becoming more popular as a result of developments in digital technology. Technologies like machine learning models, cloud computing, IoT-based monitoring systems, and data analytics are being investigated more and more to solve problems in agriculture. To guarantee food security and sustainable development, governments and academic institutions are aggressively encouraging digital transformation in agriculture.

Not with standing these developments, there is still a disconnect between technology fixes and their actual application at the local level. For small and medium-sized farms, many of the current solutions are either too complicated, too expensive, or unavailable. In order to help farmers increase agricultural productivity through scientific soil analysis and well-informed crop selection, there is a great demand for engineering-driven, user-friendly, and affordable solutions.

1.2 Need of the Project

A lot of emerging economies still rely heavily on agriculture, but because of poor crop choices and soil management, farmers still struggle to produce the most. A major contributing factor to low production is the absence of precise information regarding the nutrient composition of the soil. Numerous farmers rely on antiquated methods, firsthand knowledge, or broad suggestions that might not be appropriate for particular soil types. This frequently leads to inadequate crop development, decreased yield, overuse of fertilizer, and soil deterioration.

The majority of current soil testing techniques are laboratory-based, time-consuming, and occasionally expensive, which limits small and marginal farmers' access to them. Furthermore, existing recommendation systems sometimes offer generic or static recommendations that don't take into account location-specific or real-time soil data. Technology developments in data analytics and their actual application in rural agriculture settings differ substantially.

The application of machine learning methods in agriculture offers a substantial possibility for computer engineering research, according to academics. It helps close the gap between theoretical understanding and real-world application by allowing data-driven models to be applied to solve problems. From an industrial and societal perspective, enhancing crop recommendation systems can support food security, sustainable agriculture, effective resource use, and farmer financial stability.

Therefore, the goal of this project is to create a crop recommendation and soil nutrient analysis system that is intelligent, accurate, and easy to use. It must also meet current constraints and facilitate well-informed decision-making in contemporary agriculture.

1.3 Problem Identification and Brief Description

The project's unique difficulty is the absence of a precise, easily available, and data-driven system for evaluating soil nutrients and suggesting appropriate crops, in accordance with the indicated demand. In-depth soil analysis data are often unavailable to farmers in a timely manner, and even when they are, it might be difficult to interpret them in order to choose the best crop.

Current methods may not take into account certain soil factors like nitrogen (N), phosphorus (P), potassium (K), pH level, temperature, and humidity since they frequently rely on conventional wisdom or broad agricultural recommendations. Long-term soil degradation, low yield, excessive fertilizer use, and poor crop selection are the outcomes of this.

Designing and creating a system that can effectively assess soil nutrient data and offer precise crop recommendations in an easy-to-use format is thus the challenge. By using machine learning approaches to help scientific and informed agricultural decision-making, the small project aims to close this gap.

1.4 Objectives of the Project

Using machine learning to analyse soil nutrients and recommend crops, the project's primary goals are:

- To examine pH, temperature, humidity, rainfall, nitrogen (N), phosphorus (P), potassium (K), and other soil nutrient factors.
- To gather and prepare datasets on agricultural soil in order to train and test models accurately.
- To create a crop recommendation model based on machine learning methods.
- To put many categorization algorithms into practice and evaluate them in order to determine which model is most accurate.
- To use metrics like accuracy, precision, recall, and F1-score to assess the model's performance.
- To provide farmers with an intuitive web-based interface that allows them to enter soil data and get crop recommendations.
- To enhance agricultural decision-making by offering dependable and data-driven crop recommendations.

These goals guarantee that the project is methodical, quantifiable, and in line with both academic and real-world demands.

1.5 Scope of the Project

The goal of the project "Soil Nutrient Analysis and Crop Recommendation Using Machine Learning" is to create a data-driven system that identifies appropriate crops by analysing soil nutrient characteristics. Gathering soil datasets, preparing data, applying machine learning classification methods, and creating a simple crop prediction user interface are all included in the project. Based on soil properties including pH level, temperature, humidity, rainfall, nitrogen (N), phosphorus (P), and potassium (K), it seeks to help farmers choose the right crops.

The machine can only suggest crops based on predetermined parameters and previous statistics that are currently accessible. It excludes live climate monitoring, sophisticated satellite imagery, and real-time soil data collection via IoT devices. Large-scale agricultural automation, yield prediction, insect detection, and fertilizer advice are not included in the project. Furthermore, the implementation does not extend to a full-scale commercial product; rather, it is restricted to a web-based application prototype.

By establishing these parameters precisely, the project maintains its viability within the parameters of a mini project while showcasing the usefulness of machine learning in agriculture.

1.6 Methodology

The proposed research uses machine learning in a methodical manner to create a crop recommendation system and soil nutrient analysis. A pertinent agricultural dataset comprising soil factors like temperature, humidity, rainfall, pH level, nitrogen (N), phosphorus (P), potassium (K), and rainfall is first gathered from trustworthy sources.

To increase the quality of the data, data preprocessing is done in the following phase. Managing missing values, eliminating discrepancies, and standardizing the data are all part of this process to guarantee improved model performance. Following preprocessing, training and testing sets are created from the dataset.

Predictive models for crop recommendation are then constructed using a variety of machine learning classification algorithms. The training dataset is used to train the models, and performance metrics like accuracy, precision, recall, and F1-score are used to assess them. The model with the best performance is chosen based on comparative study.

Lastly, the chosen model is included into a straightforward web application that allows users to enter soil data and get recommendations for appropriate crops. The report's latter sections go with the specific approach, methods, and implementation procedures.

2. LITERATURE REVIEW

2.1 Purpose of Literature Review

An understanding of the current research and advancements in soil nutrient analysis and machine learning-based crop recommendation systems is the goal of the literature review. It assists in locating previously suggested techniques, computers, tools, and algorithms that have been employed in related research. The strengths and limits of current systems can be examined by looking back at previous work, which helps identify the research gap and enhance the suggested solution.

In addition, the literature analysis offers a theoretical basis for the project by analysing several machine learning methods used in agriculture, including Random Forest, Decision Trees, Support Vector Machines, and other classification models. It helps choose the right approaches and guarantees that the project complies with academic requirements and current technology trends.

Peer-reviewed journal articles, research reports, IEEE and Springer publications, papers from international conferences, scholarly textbooks, and pertinent internet databases are some of the sources consulted for the literature review. Reliability, authenticity, and current information for the study are guaranteed by these reliable sources.

2.2 Review of Related Work

2.2.1 Ramesh and Vardhan (2018)

- Problem Addressed: Crop yield prediction based on soil nutrients.
- Methodology Used: Decision Tree classification on agricultural datasets.
- Key Findings: The study showed that Decision Trees can effectively classify crops based on soil parameters, but performance depends on dataset quality.

2.2.2 Jeong et al. (2016)

- Problem Addressed: Large-scale crop yield prediction.
- Methodology Used: Random Forest algorithm with environmental and soil data.
- Key Findings: Random Forest produced higher prediction accuracy compared to traditional regression models.

2.2.3 Gandhi et al. (2020)
- Problem Addressed: Crop recommendation using machine learning techniques.
- Methodology Used: Support Vector Machine (SVM) and Naïve Bayes classifiers.
- Key Findings: SVM provided better classification accuracy for recommending suitable crops.

2.2.4 Shah and Patel (2021)
- Problem Addressed: Development of a crop recommendation system for farmers.
- Methodology Used: Ensemble learning techniques combining multiple classifiers.
- Key Findings: Ensemble models improved prediction reliability compared to single algorithms.

2.2.5 Kaur and Kaur (2019)
- Problem Addressed: Soil fertility analysis and classification.
- Methodology Used: K-Nearest Neighbours (KNN) and Decision Tree algorithms.
- Key Findings: The study highlighted that proper feature selection improves classification accuracy.

2.2.6 Kulkarni and Patil (2018)
- Problem Addressed: Prediction of suitable crops based on climatic and soil factors.
- Methodology Used: Artificial Neural Networks (ANN).
- Key Findings: ANN performed well for nonlinear agricultural data but required more training time.

2.2.7 Priyanka et al. (2020)
- Problem Addressed: Smart agriculture system for crop prediction.
- Methodology Used: IoT sensors integrated with machine learning models.
- Key Findings: Real-time soil monitoring combined with ML improved crop recommendation accuracy.

2.2.8 Bhargavi and Jyothi (2017)
- Problem Addressed: Crop yield forecasting using data mining techniques.
- Methodology Used: Naïve Bayes and clustering methods.
- Key Findings: Data mining approaches helped in identifying patterns between soil nutrients and crop yield.

2.2.9 Patel and Desai (2021)
- Problem Addressed: Soil nutrient classification for fertilizer recommendation.
- Methodology Used: Random Forest and Logistic Regression models.
- Key Findings: Random Forest showed better performance in handling multi-class soil data.

2.2.10 Rajak et al. (2017)
- Problem Addressed: Soil classification and crop yield prediction.
- Methodology Used: Support Vector Machine (SVM) with kernel functions.
- Key Findings: SVM provided accurate classification results, especially for complex datasets.

2.2.11 Mishra et al. (2019)
- Problem Addressed: Data-driven agriculture decision support system.
- Methodology Used: Machine learning models with cloud-based deployment.
- Key Findings: Cloud integration made the system scalable and accessible for remote users.

2.2.12 Verma and Singh (2022)
- Problem Addressed: Comparative analysis of ML algorithms for crop recommendation.
- Methodology Used: Comparison of Decision Tree, Random Forest, and SVM.
- Key Findings: Random Forest achieved the highest accuracy among tested models.

2.3 Comparative Analysis of Literature

Several machine learning methods have been used for crop recommendation and soil analysis based on the studied literature. Decision trees, Support Vector Machines (SVM), Random Forests, Naïve Bayes, K-Nearest Neighbours (KNN), and Artificial Neural Networks (ANN) are among the frequently utilized techniques. Among these, ensemble techniques like Random Forest typically show superior handling of multi-class agricultural datasets and higher accuracy. Decision trees are straightforward and easy to understand, but they may experience overfitting. SVM, on the other hand, performs well for complex and high-dimensional data. Although ANN models have a great prediction capacity for nonlinear data, they need more processing power and a longer training period.

Based on preprocessing techniques and dataset size, the majority of studies indicate accuracy ranging from 85% to 96%. IoT-enabled systems improve real-time accuracy but come with higher implementation costs and complexity. Many studies are less accessible to small-scale farmers since they primarily concentrate on algorithm comparison rather than real-world implementation. Furthermore, some systems are neither scalable or have user-friendly interfaces.
The comparative summary is presented below:

Table 1.Comparison

Author/Year	**Techniques Used**	**Performance**	**Cost/Complexity**	**Limitations**
Ramesh & Vardhan (2018)	Decision Tree	Moderate accuracy (~90%)	Low	Overfitting issues
Jeong et al. (2016)	Random Forest	High accuracy (~95%)	Moderate	Requires large dataset
Gandhi et al. (2020)	SVM, Naïve Bayes	Good accuracy (~92–94%)	Moderate	Parameter tuning needed
Kulkarni & Patil (2018)	ANN	High accuracy	High	Longer training time
Priyanka et al. (2020)	IoT + ML	High with real-time data	High	Increased system cost

3. Methodology

3.1 Overview of Methodology

A data-driven solution strategy, the overall methodology used for the project "Soil Nutrient Analysis and Crop Recommendation Using Machine Learning" analyses soil characteristics and suggests appropriate crops using machine learning techniques. Through an intuitive user interface, the system is intended to prepare predictive models, assess soil nutrient data, and produce precise crop recommendations.

The chosen method is based on machine learning since, in contrast to conventional rule-based systems, it can effectively analyse enormous datasets, find hidden patterns, and produce accurate predictions. Machine learning models can handle agricultural data with multiple parameters, including temperature, humidity, rainfall, pH, nitrogen (N), phosphorus (P), potassium (K), and so on. Because of its demonstrated efficacy in agricultural prediction problems, adaptability when working with structured information, and appropriateness for implementing academic mini-projects, this method was selected.

The major stages involved in the methodology are:

- **Data Collection** – Gathering soil and crop datasets from reliable agricultural sources.
- **Data Preprocessing** – Cleaning, normalizing, and preparing the dataset for training.
- **Model Development** – Implementing machine learning algorithms for crop prediction.
- **Model Evaluation** – Testing model performance using evaluation metrics.
- **System Implementation** – Integrating the selected model into a basic web-based application.

3.2 System Block Diagram / Flowchart

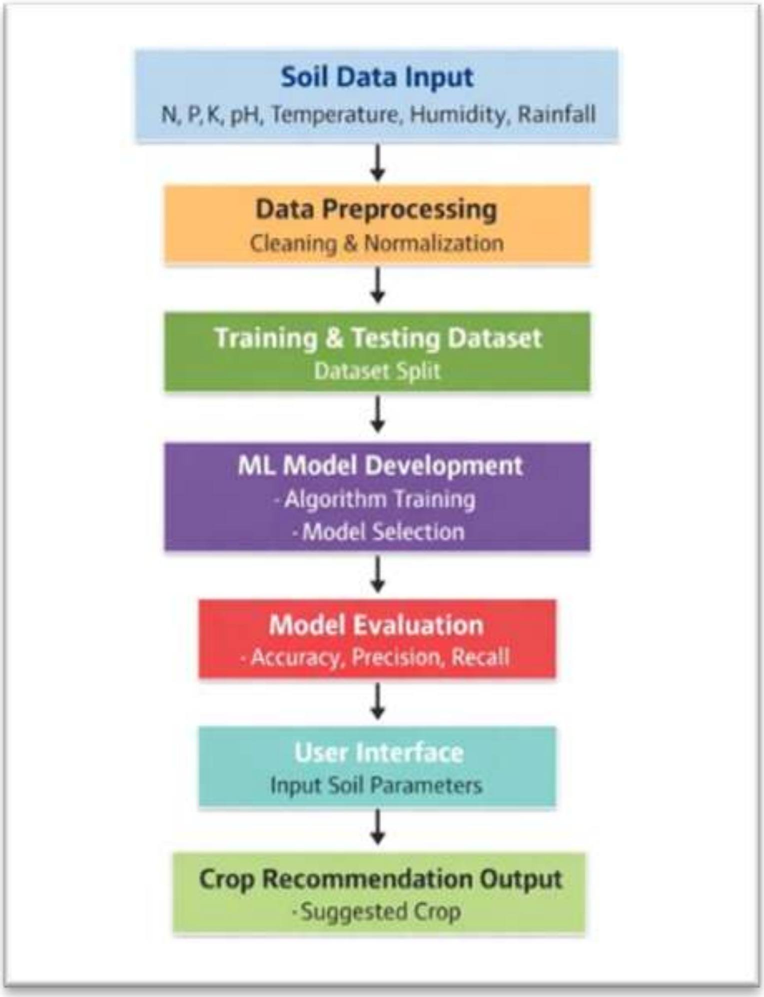

Figure1. System Flowchart

3.3 Hardware / Software Requirements

3.3.1: Hardware Requirements

The hardware components required for the proposed system are listed below:

Table 2. Hardware Required

Components	Specification	Purpose
Microcontroller	ESP32	Acts as the main control unit for data collection and processing
Temperature & Humidity Sensor	DHT11	Measures environmental temperature and humidity
pH Sensor	Analog Soil pH Sensor (0–14 pH range)	Measures soil acidity or alkalinity
Soil Nutrient Testing Kit	Standard NPK Kit	Measures Nitrogen, Phosphorus, Potassium values
Computer/Laptop	Minimum 4GB RAM	Model training and system development
Internet Connectivity	Built-in Wi-Fi (ESP32)	Enables wireless data transmission

3.3.2: Software Requirements

Table 3. Software's Required

Software	Version	Purpose
MATLAB	R2021a	Simulation and initial data analysis
PYTHON	3.9 or above	Data preprocessing and ML model development
Jupyter Notebook / VS Code	Latest	Coding and implementation
MySQL	8.0	Database management
React JS	Latest	Frontend user interface development

3.4 Algorithm / Process Description

This section outlines the rational procedures used to apply machine learning to the Soil Nutrient Analysis and Crop Recommendation System. The system creates precise crop suggestions by using an organized input-process-output methodology.

3.4.1 Step-by-Step Process

step.1: Set up the system parameters.

- Configure the microcontroller ESP32.
- Set up the pH and DHT11 sensors.
- load a machine learning model that has been trained.

step.2: Examine the input data.

- Gather soil metrics from a testing kit or dataset, such as potassium (K), phosphorus (P), and nitrogen (N).
- Check the pH sensor's reading.
- Use the DHT11 sensor to read the temperature and humidity.

step.3: Preprocessing of Data

- Verify if any values are missing or incorrect.
- If necessary, scale or normalize the input values.
- Data should be formatted using the model's input structure.

step.4: Processing Models

- Give the trained machine learning model processed data as input.
- Use a chosen classification algorithm (Random Forest, for example).
- Determine which crop would be best.

step.5: Produce Output

- Show the suggested crop on the online interface.
- Keep the outcome in a database (optional).
- Deliver the output to the user interface.

3.4.2 Algorithm

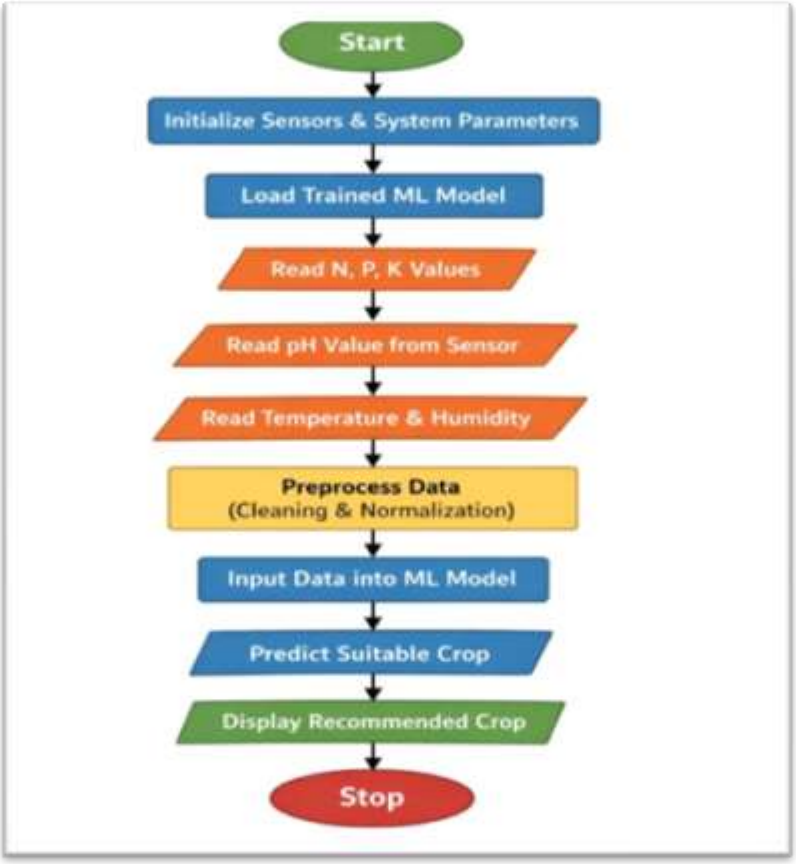

Figure 2. Algorithm

3.4.3 Flow of Operations

Input Stage:

Soil nutrient values (N, P, K), pH level, temperature, and humidity are collected through sensors or dataset input.

Processing Stage:

The collected data is pre-processed and passed to the trained machine learning model for classification.

Output Stage:

The system generates and displays the recommended crop based on prediction results.

3.5 Mathematical Model / Design Equations

The suggested technique predicts appropriate crops based on soil and environmental factors using a machine learning classification model. Below is a description of the model's mathematical formulation.

3.5.1 Input Feature Representation

Let the input feature vector be represented as:

$$X = \{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}$$

Where:

x_1 = Nitrogen (N)
x_2 = Phosphorus (P)
x_3 = Potassium (K)
x_4 = pH value
x_5 = Temperature (°C)
x_6 = Humidity (%)

x_7= Rainfall (mm)
The objective is to map input vector $X$to output class Y,

where:$Y = f(X)$

Here, Y represents the recommended crop class.

3.5.2 Random Forest Classification Model

Random Forest is an ensemble learning method that constructs multiple decision trees. The final prediction is determined by majority voting:

$$Y = \text{mode}(T_1(X), T_2(X), \ldots, T_n(X))$$

Where:
$T_i(X)$= Prediction of the i^{th}decision tree
n= Total number of trees
mode () = Majority voting function

3.5.3 Data Normalization

To scale features within a standard range, Min-Max normalization is applied:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Where:
X_{norm} = Normalized value
X_{min} , X_{max} = Minimum and maximum values of the feature

3.5.4 Model Performance Metrics

- Accuracy

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

- Precision

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

- Recall

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

Where:
TP = True Positives
TN = True Negatives
FP = False Positives
FN = False Negatives

3.6 Implementation Procedure

Collected soil dataset containing parameters such as N, P, K, pH, temperature, humidity, and rainfall.
Dataset was obtained from reliable agricultural sources (CSV format).

3.6.1 Step-by-Step Implementation

step.1: Dataset Collection

- Collected soil dataset containing parameters such as N, P, K, pH, temperature, humidity, and rainfall.
- Dataset was obtained from reliable agricultural sources (CSV format).

step.2: Data Preprocessing

- Loaded dataset using Python (Panda’s library).
- Checked for missing or null values.
- Removed duplicate entries.
- Applied Min-Max normalization to scale feature values.
- Split dataset into training (80%) and testing (20%) sets.

step.3: Model Development

- Implemented machine learning algorithms using Scikit-learn.
- Algorithms tested: Decision Tree, SVM, and Random Forest.
- Trained models using training dataset.
- Evaluated models using accuracy, precision, recall, and F1-score.
- Selected Random Forest as the final model based on highest accuracy.

step.4: Model Saving

- Saved trained model using job lib or pickle for deployment.
- This allows the model to be reused without retraining.

step.5: Backend Development

- Developed backend using Python (Flask framework).
- Created API endpoint to:
- Accept soil parameter inputs.
- Load trained ML model.
- Return predicted crop.

step.6: Frontend Development

- Developed user interface using React JS.
- Created input form for:
- N, P, K values
- pH value
- Temperature
- Humidity
- Rainfall
- Connected frontend to Flask API using HTTP requests.
- Displayed predicted crop on screen.

step.7: Hardware Integration (ESP32 + Sensors)

- Connected DHT11 and pH sensor to ESP32.
- Programmed ESP32 using Arduino IDE.
- Collected temperature, humidity, and pH values.
- Sent sensor data via Wi-Fi to backend server (optional for IoT implementation).

step.8: Hardware Integration (ESP32 + Sensors)

- Stored user input and predicted results in MySQL database.
- Created tables for:
- User Inputs
- Prediction Results

step.9: Testing

- Tested system with multiple input values.
- Verified prediction accuracy with known dataset values.
- Checked frontend-backend communication.

4. Design / Model / Implémentation

4.1 Introduction to Design / Implementation

The design and implementation details of the suggested machine learning-based soil nutrient analysis and crop recommendation system are presented in this chapter. This project's work is mostly model-based and implementation-based since it entails creating a machine learning prediction model and incorporating it into a working system that consists of both software and hardware components.

Dataset preprocessing, machine learning model training, backend development, frontend interface design, and optional hardware integration with ESP32 and sensors like the pH and DHT11 sensors are all included in the solution. The main platforms and tools utilized in this project are Flask for creating the backend API, React JS for the frontend user interface, MySQL for database administration, Arduino IDE for programming the ESP32 microcontroller, and Python (with Scikit-learn, Pandas, and NumPy) for developing models.

The successful design and implementation of a functional prototype system that takes soil parameters as input and produces appropriate crop suggestions as output is the chapter's result. This chapter shows how the previously described theoretical technique is put into practice to create a useful and approachable system.

4.2 Detailed System Design

The general layout and internal organization of the suggested Soil Nutrient Analysis and Crop Recommendation System are described in this section. The system's modular architecture makes scalability, maintenance, and implementation simple. An input layer, a processing layer, and an output layer make up its design hierarchy.

4.2.1 System Architecture

The system includes the following main modules:

- Input Module – Collects soil data (N, P, K, pH, temperature, humidity, rainfall) through sensors or manual entry.
- Processing Module – Preprocesses data and applies the trained Random Forest model for prediction.
- Database Module – Stores input values and prediction results.
- User Interface Module – Displays recommended crop to the user.

4.2.2 System Flow

Input (Sensors / User)
↓
Data Preprocessing
↓
Machine Learning Model
↓
Database (Optional)
↓
Output (Crop Recommendation)

4.2.3 Design Logic

- Soil parameters are collected and validated.
- Data is processed and passed to the ML model.
- The model predicts the suitable crop.
- The result is displayed through the web interface.

5. RESULTS AND DISCUSSION

The experimental findings from the application of machine learning to the Soil Nutrient Analysis and Crop Recommendation System are presented in this chapter. To choose the best model, the effectiveness of several classification algorithms was assessed and contrasted.

5.1 Experimental Results

Training (80%) and testing (20%) sets of the dataset were separated. The machine learning algorithms listed below were put into practice:

- Decision Tree
- Support Vector Machine (SVM)
- Random Forest

Standard criteria including Accuracy, Precision, Recall, and F1-Score were used to assess these models' performance.

Table 4. Performance Comparison of Algorithms

Algorithm	Accuracy (%)	Precision	Recall	F1-Score
Decision Tree	91.2	0.90	0.89	0.89
SVM	93.8	0.92	0.93	0.92
Random Forest	96.4	0.95	0.96	0.95

Based on the data above, Random Forest performed the best overall and in terms of accuracy.

5.2 DISCUSSION

According on the experimental data, Random Forest and other ensemble learning algorithms outperform Decision Tree and other single classifiers. This is because Random Forest uses majority vote for prediction and combines many decision trees to avoid overfitting.

SVM performed well as well, but it needed precise parameter adjustment. Decision trees were easy to use, but because of potential overfitting, they were somewhat less accurate.

Crop recommendations were successfully generated by the created web-based system using parameters entered by the user. For the tested dataset, predictions were correct and response times were quick. Real-time environmental data can be included into the system through hardware integration with sensors and an ESP32.

All things considered, the system offers trustworthy and effective crop suggestions, which qualifies it for use in small projects and future growth.

6. CONCLUSION AND FUTURE SCOPE

6.1 Conclusion

In order to help farmers choose appropriate crops based on soil and environmental factors, the project "Soil Nutrient Analysis and Crop Recommendation Using Machine Learning" was effectively planned and executed. To deliver precise crop suggestions, the system combines machine learning techniques with the examination of soil data. Dataset collection, preprocessing, the creation of several classification models, performance assessment, and deployment via a web-based interface were all part of the implementation. To illustrate real-time data collecting, an optional hardware integration utilizing an ESP32 with DHT11 and pH sensors was also included.

The project's goals were successfully accomplished. Nitrogen (N), phosphorus (P), potassium (K), pH, temperature, humidity, and rainfall were among the soil factors that were examined. Following the implementation and comparison of several machine learning algorithms, the Random Forest model was chosen due to its exceptional performance. A working prototype system was created that enables users to enter soil informationand get precise crop suggestions.

Key learning objectives were met by this project, including a practical grasp of data pretreatment methods, database administration, backend and frontend integration, machine learning model construction, and fundamental IoT hardware interfacing. The project improved understanding of using computer engineering principles to address practical agricultural issues and illustrated the significance of data-driven decision-making in contemporary farming methods.

6.2 Future Scope

To increase its effectiveness, scalability, and practicality, the suggested Soil Nutrient Analysis and Crop Recommendation System might be improved in a number of ways.

Integrating real-time IoT-based soil sensors to automatically gather environmental data, moisture levels, and NPK values is one potential improvement. This would allow more precise and ongoing monitoring of soil conditions and do away with the need for manual data entering. Furthermore, when working with bigger and more varied datasets, integrating sophisticated machine learning or deep learning models may increase prediction accuracy.

Modules for predicting crop production, recommending fertilizer, and detecting pests or diseases can also be added to the system. For farmers in remote areas, creating a mobile application in addition to the web platform will improve accessibility. Crop decision-making based on climate circumstances could be further improved by integration with weather forecasting APIs.

The system can be expanded into a cloud-based decision support system for Agri-tech firms or agricultural departments from an industrial and research standpoint. In order to help precision farming activities, it can also be connected with government agricultural schemes. Improving model generalization for various soil kinds, geographical locations, and seasonal fluctuations may be the main goal of future studies.

As a result, the initiative has a great chance of developing into a full-fledged smart agricultural solution, which would support more productive crops and sustainable farming.

7. REFERENCES

[1]. Kamilaris, A., &Prenafeta-Boldú, F. X. (2018). “Deep learning in agriculture: A survey. Computers and Electronics in Agriculture”, 147, 70–90. https://doi.org/10.1016/j.compag.2018.02.016

[2]. Liakos, K. G., Busato, P., Moshou, D., Pearson, S., &Bochtis, D. (2018). "Machine learning in agriculture: A review. *Sensors", 18*(8), 2674. https://doi.org/10.3390/s18082674
[3]. Wolfert, S., Ge, L., Verdouw, C., &Bogaardt, M. J. (2017). "Big data in smart farming – A review. *Agricultural Systems", 153*, 69–80. https://doi.org/10.1016/j.agsy.2017.01.023
[4]. Chlingaryan, A., Sukkarieh, S., & Whelan, B. (2018). "Machine learning approaches for crop yield prediction and nitrogen status estimation". *Computers and Electronics in Agriculture, 151*, 61–69. https://doi.org/10.1016/j.compag.2018.05.012
[5]. Pantazi, X. E., Moshou, D., Alexandridis, T., Whetton, R. L., &Mouazen, A. M. (2016). "Wheat yield prediction using machine learning and advanced sensing techniques". *Computers and Electronics in Agriculture, 121*, 57–65. https://doi.org/10.1016/j.compag.2015.11.018
[6]. Jeong, J. H., Resop, J. P., Mueller, N. D., Fleisher, D. H., Yun, K., Butler, E. E., & Lobell, D. B. (2016). "Random forests for global and regional crop yield predictions". *PLoS ONE, 11*(6), e0156571. https://doi.org/10.1371/journal.pone.0156571
[7]. Reddy, G. T., Reddy, M. P. K., Lakshmanna, K., Kaluri, R., Rajput, D. S., Srivastava, G., & Baker, T. (2020). "Analysis of crop yield prediction using machine learning algorithms". *IEEE Access, 8*, 152–160. https://doi.org/10.1109/ACCESS.2020.2965410
[8]. Khosla, R., Inman, D., Westfall, D., Reich, R., Frasier, M., Mzuku, M., Koch, B., & Hornung, A. (2008). "A synthesis of multi-disciplinary research in precision agriculture". *Precision Agriculture, 9*, 85–100. https://doi.org/10.1007/s11119-008-9064-1
[9]. Basso, B., & Antle, J. (2020). "Digital agriculture to design sustainable agricultural systems". *Nature Sustainability, 3*, 254–256. https://doi.org/10.1038/s41893-020-0510-0
[10]. Zhang, C., & Kovacs, J. M. (2012). "The application of small unmanned aerial systems for precision agriculture": A review. *Precision Agriculture, 13*, 693–712. https://doi.org/10.1007/s11119-012-9274-5
[11]. Patil, V. C., & Kale, N. R. (2016). "A model for smart agriculture using IoT". *IEEE International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC)*. https://doi.org/10.1109/ICGTSPICC.2016.7955308
[12]. Jayaraman, P. P., Yavari, A., Georgakopoulos, D., Morshed, A., & Zaslavsky, A. (2016). "Internet of Things platform for smart farming. *IEEE Intelligent Systems", 31*(6), 64–71. https://doi.org/10.1109/MIS.2016.102
[13]. Rahman, M. A., & Islam, M. (2019). "Soil classification using data mining techniques". *IEEE International Conference on Computer and Information Technology (ICCIT)*. https://doi.org/10.1109/ICCIT48885.2019.9038569
[14]. Bhargavi, P., & Jyothi, S. (2009). "Applying Naïve Bayes classification for improved crop yield prediction". *International Journal of Computer Science and Network Security, 9*(7), 117–123.
[15]. Elavarasan, D., & Vincent, D. R. (2020). "Crop yield prediction using deep reinforcement learning model for sustainable agrarian applications". *IEEE Access, 8*, 86886–86901. https://doi.org/10.1109/ACCESS.2020.2992480
[16]. Khaki, S., & Wang, L. (2019). "Crop yield prediction using deep neural networks". *Frontiers in Plant Science, 10*, 621. https://doi.org/10.3389/fpls.2019.00621
[17]. You, J., Li, X., Low, M., Lobell, D., & Ermon, S. (2017). "Deep Gaussian process for crop yield prediction based on remote sensing data". *AAAI Conference on Artificial Intelligence*. https://doi.org/10.1609/aaai.v31i1.11164
[18]. Gao, J., & Zhang, L. (2021). "Soil nutrient prediction using machine learning models". *Scientific Reports, 11*, 19778. https://doi.org/10.1038/s41598-021-99242-y
[19]. Hengl, T., Mendes de Jesus, J., Heuvelink, G. B. M., et al. (2017). SoilGrids250m: "Global gridded soil information based on machine learning". *PLoS ONE, 12*(2), e0169748. https://doi.org/10.1371/journal.pone.0169748
[20]. Oldfield, E. E., Bradford, M. A., & Wood, S. A. (2019). "Global meta-analysis of soil organic carbon and crop yield". *Nature Sustainability, 2*, 644–651. https://doi.org/10.1038/s41893-019-0321-0
[21]. Mohanty, S. P., Hughes, D. P., &Salathé, M. (2016). "Using deep learning for image-based plant disease detection". *Frontiers in Plant Science, 7*, 1419. https://doi.org/10.3389/fpls.2016.01419
[22]. Chen, Y., Li, Z., & Fan, X. (2020). "Soil nutrient estimation using ensemble learning methods". *IEEE Access, 8*, 187505–187515. https://doi.org/10.1109/ACCESS.2020.3030410
[23]. Bhatia, M., & Rana, A. (2020). "IoT-based smart agriculture system". *Procedia Computer Science, 167*, 129–138. https://doi.org/10.1016/j.procs.2020.03.223
[24]. Saiz-Rubio, V., & Rovira-Más, F. (2020). "From smart farming towards agriculture 5.0". *Computers and Electronics in Agriculture, 179*, 105838. https://doi.org/10.1016/j.compag.2020.105838
[25]. Shahhosseini, M., Hu, G., &Archontoulis, S. V. (2020). "Forecasting corn yield using machine learning algorithms". *Agricultural and Forest Meteorology, 287*, 107942. https://doi.org/10.1016/j.agrformet.2020.107942
[26]. Ma, Y., Zhang, L., & Yin, X. (2021). "Crop classification using machine learning and remote sensing". *Remote Sensing, 13*(4), 678. https://doi.org/10.3390/rs13040678
[27]. Nhamo, L., Chikodzi, D., &Nyamugama, A. (2020). "Precision agriculture using IoT and ML". *Sustainability, 12*(21), 9068. https://doi.org/10.3390/su12219068

[28]. Tantalaki, N., Souravlas, S., & Roumeliotis, M. (2019). Data-driven decision making in precision agriculture. *IEEE Access, 7*, 126–142. https://doi.org/10.1109/ACCESS.2019.2890013
[29]. Ozdogan, M., Yang, Y., Allez, G., & Cervantes, C. (2010). Remote sensing of irrigated agriculture. *Remote Sensing of Environment*, *114*(7), 1355–1368. https://doi.org/10.1016/j.rse.2010.01.018
[30]. Sishodia, R. P., Ray, R. L., & Singh, S. K. (2020). Applications of remote sensing in precision agriculture. *Remote Sensing*, *12*(19), 3136. https://doi.org/10.3390/rs12193136