Enhancing Sustainable Agriculture through Machine Learning-Based Crop Recommendation and NPK Sensor Integration

Supervisor:

Dr. Zineb Hassan Ali

Assistant Professor at Faculty of Artificial Intelligence, KSF

Team Members:

1- Moghazy Refaat Mohammed
2- Moaz Mohamed Eldsouky
3- Mohamed omar elsayed
4- Omar Mohamed eid Kamel
5- Ramadan El-sayed Saad
6- Ahmed El-Sayed Abo Elkhair
7- Ahmed Elsayed Ali Abu bakr
8- Fawzy Ibrahim Eissa
9- Elsaeed Mohamad Elrhmany
10- Omar Alaa abdelmaged kamel

2024-2025

Acknowledgment

First and foremost, we extend our deepest gratitude and praise to God Almighty for His abundant blessings and guidance throughout the course of our project, enabling us to complete this work successfully.

We are especially honored to express our sincere appreciation to our esteemed supervisor Dr. Zineb Hassan, for granting us the opportunity to undertake this project and for her invaluable support and direction throughout its duration. Her insightful feedback, expert guidance, and unwavering encouragement were instrumental in shaping the quality and direction of our work. It was a privilege to study and work under her supervision, and we remain deeply thankful for her mentorship, kindness, and positive spirit.

We would also like to express our gratitude to Dr. Tamer Medhat, for his continuous support and for facilitating our participation in academic competitions through his kind endorsement. His efforts also enabled us to connect with the Faculty of Artificial Intelligence, allowing us to further explore interdisciplinary studies that bridged our project with engineering sciences.

We are truly grateful for the contributions of both individuals to our academic journey.

Abstract

Agriculture is a cornerstone of human sustenance and economic development. However, farmers often face significant challenges in selecting the most suitable crops for their land, leading to suboptimal yields and resource inefficiencies. This project, "Crop Recommendation" aims to harness the power of data science and machine learning to provide actionable crop recommendations tailored to specific soil, weather, and environmental conditions.

By leveraging a comprehensive dataset that includes parameters such as soil composition, temperature, humidity, rainfall, and pH levels, the project develops a robust predictive model capable of suggesting crops with the highest potential for success. Advanced techniques, including exploratory data analysis, feature engineering, and machine learning algorithms, are employed to build an efficient recommendation system.

The proposed solution not only enhances agricultural productivity but also promotes sustainable resource utilization by reducing wastage of water, fertilizers, and labor. This research holds immense potential for empowering farmers, improving food security, and supporting global agricultural innovation. The project's findings highlight the transformative role of technology in modern farming and provide a scalable blueprint for integrating AI into agricultural practices worldwide.

Contents:

Chapter 1: Introduction	
1. Overview…………………………………………………………	7
2. A general introduction to agriculture and its importance….........	8
3. The problem of selecting appropriate crops…………………………..	8
4 The importance of crop recommendations using data………………..	8
5 Project objectives……………………………………………………..	8
Chapter 2: Literature Review	
1. Overview………………………………………………………..	10
2. Related Works………………………………………………………..	10
Chapter 3: The Proposed Methodology	
1. Overview………………………………………………………	14
2 Role of Artificial Intelligence in Agriculture…………………	14
3 Role of Sensors in Precision Farming……………………..	15
4 Hardware Setup and Sensor Integrations……………………..	17
5 Understanding Soil Parameters………………………………	39
5.1 Soil Moisture……………………………………………………….	39
5.2 Humidity………………………………………………………..	40
5.3 Temperature……………………………………………………	41
5.4 Electric Conductivity (EC)……………………………………………..	42
5.5 pH…………………………………………………………………	43
5.6 Nitrogen (N), Phosphorus (P), Potassium (K)….................................	44
2 .Role of AI in Crop Recommendation…………………………………	45
2.1 Importance of Soil Analysis for Agriculture…………………..	46
2.2 Challenges in Traditional Farming Methods………………….	46
2.3 Fields Where Crop Recommendation Systems Are Used……	47
2.4 How AI Can Help with Crop Recommendation……………..	47
3. Data Acquisition and Preprocessing………………………………..	50
3.1 Methods for collecting sensor data…………………………..	50
3.2 Storing data (CSV, Database, Cloud integration)……………	50
3.3 Data cleaning and handling missing value……………………	51
3.4 Feature scaling and normalization……………………………	52

3.5	Overview of ML algorithms for crop recommendation……….	52
3.6	Performance evaluation (Accuracy, Precision, Recall, F1-Score) ……………………………………………………….	54
3.7	Model optimization techniques	54
Chapter 4: Applications and Benefits of the Soil Analysis Device		61
1. Overview…………………………………………………………		62
2 .Applications of the Soil Analysis Device ……………………………		62
2.1	Precision Agriculture ……………………………………	62
2.2	Water Conservation ………………………………………	62
2.3	Soil Health Management ……………………………………	62
2.4	Decision Support for Farmers……………………………..	63
3. Societal and Environmental Benefits ……………………………….		63
3.1	Increased Agricultural Productivity ……………………..	63
3.2	Reduction in Environmental Pollution ……………………	63
3.3	Sustainable Land Use ……………………………………..	63
3.4	Economic Empowerment …………………………………..	62
3.5	Climate Resilience ………………………………………	62
4. Enhancing Agricultural Sustainability Through Technology …………		62
4.1	Advancing Smart Farming Techniques …………………….	62
4.2	Supporting Eco-Friendly Practices ………………………..	62
4.3	Encouraging Data-Driven Agricultural Policies ……………..	63
4.4 Summary ……………………………………………………………		63
Chapter5: Results and Methods		64
5.1 Machine Learning Model performance ……………………………..		65
5.2 Crop Recommendation Device Evolution ………………………….		66
5.3 Comparative Analysis……………………………		66
5.4 Summary………………………………………………..		66
6. Comparisons of 5 different model's evaluation metrics…………………		69
7. Analysis of results…………………………………………………….		72
8. Proposed of the project……………………………………………….		72
Chapter 6: Conclusion and Future Works		71
6.1. Conclusion ………………………………………………………….		70
6.2. Future Works……………………………………………………….		73
6.3. Challenges……………………………………………………		74
6.4 Web pages…………………………………………………………		74
List of Publication		75
List of Table		76
List of Figure		77
Reference		78

Chapter 1

Introduction

1. Overview

Agriculture is one of the main pillars of the global economy and the provision of food for humanity. However, farmers face many challenges in choosing the most suitable crops to grow under different environmental conditions such as soil type, weather, and the amount of water available. Inappropriate crop selection can lead to poor productivity, increased costs, and loss of resources. The data will be analyzed and processed to extract important characteristics that affect crop success. Using machine learning algorithms, a model will be trained that can predict the most suitable crops for each set of input conditions.

2. A general introduction to agriculture and its importance

Agriculture has been the backbone of human civilization for thousands of years, serving as the primary source of food, raw materials, and economic activity. It is a vital sector that not only sustains human life but also supports livelihoods, particularly in rural areas where it remains the primary source of income for millions of people.

The importance of agriculture extends beyond food production. It provides raw materials for industries such as textiles, biofuels, and pharmaceuticals. Additionally, it plays a critical role in maintaining ecological balance by supporting biodiversity and contributing to carbon sequestration.

With the growing global population and increasing demand for food, innovations in agriculture, such as precision farming, smart irrigation, and crop recommendation systems, are becoming increasingly necessary.

3. The problem of selecting appropriate crops

Selecting the right crops for cultivation is a critical decision that significantly impacts agricultural productivity, resource utilization, and economic returns. However, this decision is often fraught with challenges due to the complexity of factors that must be considered, such as:

1- Soil Characteristics
2- Climatic Conditions
3- Resource Availability
4- Market Demand
5- Environmental Sustainability

4. The importance of crop recommendation using data

Data-driven crop recommendation is crucial for modern agriculture as it helps farmers make informed decisions based on scientific analysis rather than intuition. By analyzing factors such as soil quality, climate conditions, and resource availability, data-driven systems provide tailored suggestions for the most suitable crops, resulting in:

1. Increased Productivity
2. Resource Efficiency
3. Economic Benefits
4. Sustainability
5. Risk Mitigation

5. Project Objectives

The primary objective of this project is to develop an intelligent crop recommendation system designed to assist farmers in making data-driven decisions regarding crop selection. By leveraging sensor data and machine learning algorithms, the system aims to enhance overall agricultural productivity through precise, timely recommendations. A secondary goal is to promote efficient utilization of natural resources such as soil, water, and fertilizers, thereby minimizing waste and optimizing input. The project also supports the broader goal of sustainable farming by encouraging environmentally friendly practices that preserve soil health and biodiversity. Furthermore, by providing accurate and adaptive recommendations, the system is expected to help mitigate various agricultural risks related to crop failure, market demand fluctuations, and climate variability. The objective of this work can be summarized as follows:

1. Develop a Crop Recommendation System
2. Enhance Agricultural Productivity
3. Promote Resource Efficiency
4. Support Sustainable Farming
5. Mitigate Risks

Chapter 2
Literature Review

1. Overview

The Internet of Things (IoT) frameworks play a crucial role in modern agriculture by collecting data from large and remote agricultural regions through sensor networks. This data is then utilized to forecast suitable crops using various advanced machine learning algorithms. Agriculture, being a fundamental pillar of economic development, relies heavily on efficient crop production. However, many farmers continue to rely on traditional farming methods, which are often imprecise and lead to reduced productivity and inefficient use of time and resources. By identifying critical agricultural practices and aligning them with optimal seasonal conditions, productivity in precision agriculture can be significantly improved. The primary objective of this work is to advance precision agriculture through the development of a robust IoT-based framework integrated with state-of-the-art machine learning techniques. The proposed system aims to assist farmers by offering actionable insights, such as crop recommendations based on soil compatibility and seasonal timing. It further supports decision-making by forecasting the most suitable crops to cultivate, thereby enhancing yield and contributing to sustainable agricultural development.

2. Related Works

Soil nutrient management and figuring out the right crops to use are essential for keeping farming healthy and productive. Typically, farmers use old methods that involve self-analysis and a lot of labor to decide which crops to grow and what soil nutrition to use. As a result, they do not achieve the best results or high crop yields. However, new ML systems have helped us find better solutions to these problems. This literature review looks at what research and new ideas are out there that use machine learning to help keep track of soil nutrients and suggest the right crops to grow. Systems for monitoring soils using machine learning are now capturing much interest. These systems have sensors that check for how much moisture is in the soil, how acidic it is, or what nutrients there are, and help gather data about the soil in real time. The data collected goes to an office in the middle, where it's checked and used to make decisions. For example, an IoT-based system with sensors for soil moisture and nutrients can check the condition of the soil as it happens in real time. The collected data was then used to help figure out better ways to water the plants, give them nutrients, and save money on water and fertilizer.

ML techniques have been found to work well when looking at soil data and making guesses about how much nutrients are in the land. Machine learning algorithms can see patterns and connections in big data sets, which helps them, make better predictions and help people make good, forward-thinking choices. Established an ML model that looked at past data, nearby weather, and what crops were needed to determine how specific nutrients were likely to be in the soil. The model did a good job of deciding whether plants were low or too high on nutrients, making it easier to change how much fertilizer was used when needed. Crop recommendation systems help farmers choose the best crops for their fields based on the soil type and what people are most likely to want to buy. ML algorithms have been used to make crop suggestions that check soil's nutrients, the weather, and what farm activities are popular in the market.

The recommendation of fertilizers is essential because they help farmers get better harvests and ensure that farming methods don't harm the environment in the long run. Most of the time, recommending fertilizers is done by experts who analyze the data with their experience. It is known to take considerable time and may not always be accurate. Thanks to more advanced ML techniques, there has been a rise in using them to help improve the methods for recommending fertilizers .

Represents a way of using ML to help suggest which crops and fertilizers are best to plant in each area. They applied various approaches from ML. Many studies show that people have progressed greatly with devices that help farmers watch soil nutrients and pick which crops to grow. Internet of Things technology and ML help offer real-time soil information and recommend individual solutions for farmers. They offer a great deal of potential to increase the number of crops grown, make better use of resources, and support sustainable ways of farming. Future research should work on making ML systems more accurate and able to work for larger groups, investigate how to keep food safe for everyone once it leaves the farm, and find new data sources to help the technology work well. Below is a simple table that shows the main points of the latest studies in Table 1.

Table1: Summary of Agricultural Applications Using Data-Driven Techniques

Application	Method	Dataset	Key Findings	Limitation
An automated remote field monitoring system	Lora WAN	Time series data	Visualization in the cloud	Data analysis techniques are not mentioned
Soil monitoring system	IoT	Time series data	Efficient identification of soil type with real-time display	Data analysis techniques are not mentioned
Soil classification based on micronutrients	ELM	Private	Achieved 94% accuracy	Dataset is limited to the Tamil Nadu region
Crop recommendation	MLP	Kaggle Dataset	Accuracy of 98.22%	Lacks detailed analytical evaluation
Crop recommendation platform for farmers	RF	Kaggle Dataset	Random Forest achieved 97.18% accuracy	No platform was developed for user deployment
Ongoing crop and field information support	MSVM-DAG-FFO	Own dataset	Achieved 97.3% accuracy	No interactive platform available for farmers
Correct selection of crop	SCS	Dataset from Pakistan	Accuracy of 97.4%	Dataset covers only two soil types; limited crop applicability

Chapter3

Methodology and Component

1. Overview

The development of crop recommendation systems has undergone a substantial transformation, evolving from traditional knowledge-based practices to sophisticated, data-driven solutions. Historically, farmers relied heavily on experience, seasonal intuition, and trial-and-error methods to determine suitable crops, with minimal reliance on scientific soil analysis. The Green Revolution marked a significant milestone by introducing chemical fertilizers and high-yield crop varieties, which improved productivity yet lacked tailored soil-specific recommendations. With the emergence of precision agriculture, technologies such as GPS mapping, remote sensing, and advanced soil testing were incorporated to support data-informed decision-making. The integration of Big Data, ML, and IoT technologies has further advanced the field, enabling real-time monitoring of soil characteristics and automated crop suggestion systems. Tools like modern NPK 7-in-1 soil sensors now provide comprehensive insights into soil moisture, temperature, pH, and nutrient levels. These inputs feed into AI-based models to deliver accurate and context-aware crop recommendations. Looking ahead, the adoption of deep learning, block chain technology, and advanced predictive analytics is expected to further enhance disease detection, resource optimization, and market-responsive crop selection, fostering more sustainable and productive agricultural systems.

2. Role of Artificial Intelligence in Agriculture

Artificial Intelligence (AI) is revolutionizing modern agriculture by significantly enhancing productivity, minimizing resource wastage, and enabling precise, data-driven decision-making. Through its integration into various aspects of the agricultural value chain, AI supports farmers in achieving higher yields, improved efficiency, and sustainable practices.

One of the key applications of AI is in crop recommendation and precision farming, where ML algorithms analyze data from advanced soil sensors such as NPK 7-in-1 devices that measure critical parameters including nitrogen (N), phosphorus (P), potassium (K), pH, moisture content, temperature, and electrical conductivity. These insights are processed alongside historical agricultural data and weather patterns to provide tailored crop recommendations that align with current soil health and environmental conditions.

In smart irrigation and water management, AI-enabled IoT systems dynamically regulate water distribution by responding to real-time data from soil moisture sensors and meteorological inputs. This adaptive approach prevents both overwatering and under watering, contributing to effective water conservation and improved crop health. AI also plays a pivotal role in pest and disease detection. Through advanced image recognition techniques, AI systems can identify signs of plant stress, pest infestations, and disease outbreaks from drone or mobile imagery. Early detection facilitates prompt intervention, reducing crop losses and minimizing excessive pesticide usage.

In the area of yield prediction and market analytics, AI algorithms synthesize data from soil characteristics, climate trends, and previous harvest records to generate reliable forecasts. These predictions assist farmers, agricultural planners, and policymakers in optimizing production strategies and aligning outputs with market demands. Automated farming and robotics is another frontier where AI demonstrates its potential. Intelligent machines, including autonomous robots and drones, perform labor-intensive tasks such as seeding, harvesting, weeding, and spraying with high precision, thereby reducing human labor dependency and operational costs.

3. Role of Sensors in Precision Farming

Precision farming represents a transformative approach in modern agriculture, leveraging advanced sensor technologies to monitor soil conditions, weather patterns, and crop health in real time. These technologies enable farmers to collect accurate, site-specific data, facilitating informed decision-making aimed at maximizing agricultural yields, optimizing resource utilization, and promoting environmental sustainability.

A wide array of sensor types is employed to support the objectives of precision agriculture. Soil sensors, such as the NPK 7-in-1 soil sensors, are fundamental components in crop recommendation systems. These sensors measure critical soil attributes including N, P, K, pH level, moisture content, temperature, and EC, providing real-time insights into soil fertility and composition. The collected data serves as input to AI-based models that generate tailored crop recommendations aligned with specific field conditions.

Moisture and irrigation sensors play a crucial role in water management. Soil moisture sensors detect current water levels within the soil, enabling the implementation of intelligent irrigation systems that autonomously regulate water supply to prevent both overwatering and drought stress. Similarly, weather sensors, including temperature, humidity, rainfall, and wind sensors help farmers adjust planting schedules, pesticide application, and irrigation strategies based on changing climatic variables.

Plant health and imaging sensors, such as multispectral and hyperspectral cameras mounted on drones or satellites, are used to monitor crop health, detect diseases, and assess chlorophyll content. Additionally, AI-powered imaging systems enable the early detection of pest infestations, fungal infections, and nutrient deficiencies, thereby supporting timely interventions. pH and EC sensors further contribute by measuring soil acidity and nutrient availability, guiding precise fertilizer application.

The integration of these sensor technologies significantly enhances precision farming by enabling real-time soil monitoring, optimized fertilization and irrigation practices, early disease and pest detection, and data-driven crop recommendation. Collectively, these advancements contribute to increased productivity, reduced input costs, and the promotion of sustainable farming practices.

In recent years, there has been an increasing interest in integrating sensor technology with ML to develop precision agriculture solutions. Various studies have focused on using sensors to monitor soil conditions and provide recommendations to improve crop yield and sustainability. A review of the relevant literature reveals several trends and key developments in this area.

A. Soil Monitoring Systems and Sensors

Several studies have utilized sensors to measure soil properties such as moisture, pH, temperature, and NPK. These systems provide real-time data, which is critical for making informed decisions about irrigation, fertilization, and crop management. For instance, research by [Moghazy ,Omar,Mohmmed ,Moaz,Elsaeed., 2025] explored the use of NPK sensors for monitoring soil nutrients and proposed a system for optimizing fertilization strategies. Similarly[Moghazy,Omar,Elsaeed,Mohmmed,2025] used soil moisture and temperature sensors in combination with satellite data to create a system for predicting optimal planting times.

B. ML Models for Crop Recommendations

ML has emerged as a powerful tool for processing large datasets obtained from sensors. Various algorithms, such as decision trees, support vector machines (SVM), and neural networks, have been employed to analyze soil data and generate crop recommendations. For example, [Moghazy,Omar,Elsaeed,Mohmmed,2025] developed a crop recommendation system using soil and weather data, achieving high accuracy in suggesting crops suitable for specific regions. Other studies, such as[Moghazy,Omar,Saied,Mohmmed,2025] have employed deep learning models to predict crop yield based on soil and environmental parameters.

C. Integrated Systems for Precision Agriculture

There has also been significant work on integrating multiple technologies, such as IoT, cloud computing, and big data analytics, with soil sensors and ML models to create intelligent crop recommendation systems. For instance, [Moghazy,Omar,Saied,Mohmmed,2025] implemented an IoT-based soil monitoring system combined with a cloud platform to analyze soil conditions and provide crop suggestions via a mobile app. Their system demonstrated a high level of accuracy in real-time monitoring and decision-making.

D. Challenges and Gaps in Existing Research

Despite the advancements in soil monitoring and crop recommendation systems, several challenges remain. Many systems still lack scalability and are often designed for specific crops or regions. Moreover, most existing studies focus on one or two soil parameters, limiting the range of recommendations. Additionally, the integration of multiple sensor types and ML models often presents challenges in terms of data fusion, real-time processing, and model accuracy.

E. Relevance to Current Research

The reviewed works highlight the importance of using comprehensive sensor data and advanced ML models for improving crop management. The proposed system in this research aims to bridge the gap by integrating seven critical soil parameters (soil moisture, humidity, temperature, EC, pH, and NPK values) into a unified platform for crop recommendations. By leveraging a more extensive dataset and exploring new machine learning approaches, this work aims to enhance the accuracy and applicability of crop recommendation systems across diverse farming environments.

4. Hardware Setup and Sensor Integrations

In this project, we use artificial intelligence and machine learning techniques to measure the percentage of moisture in the soil and the work of an automatic irrigation system for plants with knowledge of the type of plant that is best for cultivation in this soil through its basic elements.

A. NPK 5-pin sensor: Overview

As shown in Figure 1, NPK 7 in 5-pin soil sensors are designed to measure key soil properties such as N, P, K, moisture, temperature, and electrical conductivity. The 5-pin configuration enables more precise measurements, ensuring that farmers get accurate data about soil nutrients and conditions. This data can be used in crop recommendation systems to optimize fertilization and irrigation strategies for improved crop growth. The sensor helps to monitor and maintain soil health, promoting sustainable farming practices Cost approximately .120$.

Figure 1: A. NPK 5-pin sensor.

B. Max Rs485-TTL

As shown in Figure 2, the MAX485 TTL (Transistor-Transistor Logic) to RS-485 interface module which is used to connect the Soil NPK Sensor with the Arduino as this interface module can be easily powered up using the Arduino's 5 Volts. The max485 interface module is ideal for serial communications over long distances of up to 1200 meters or in electrically noisy environments, this is the reason it is commonly used in industrial environments. It supports up to 2.5MBit/Sec data rates, but as the distance increases, the maximum data rate that can be supported comes down. The RS-485 could communicate with multiple devices (up to 32) on the same Bus/cable when used in master and slave configuration.

We have already written a detailed article on how to use the MAX485 interface module with Arduino and communicate with multiple controllers. The term "Max RS485-TTL" refers to a communication standard used to transmit data over long distances, typically in industrial and agricultural sensor systems. RS485 is a differential bus standard that supports multiple devices on a single network. TTL refers to a logic level used for simpler, lower-voltage communication. In soil sensors or IoT applications, this communication standard is useful for sending sensor data to controllers or other devices with stable and long-distance connectivity Cost approximately 5$.

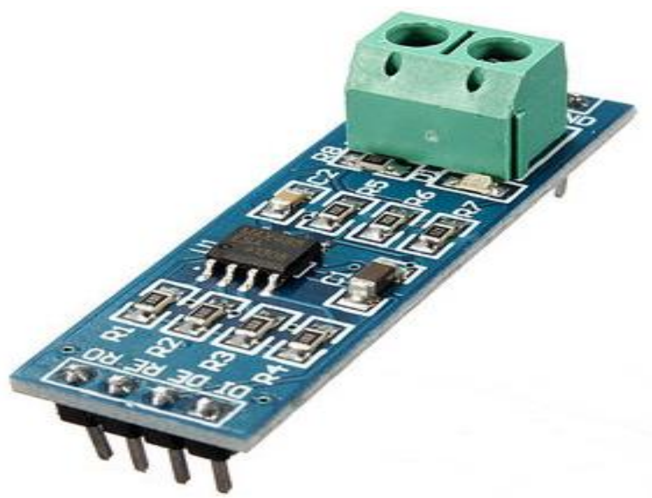


Figure 2: MAX485 TTL to RS-485 interface module.

Key Features of MAX485 TTL to RS-485 interface module:

1. Use MAX485 Interface chip.
2. Uses differential signaling for noise immunity.
3. Distances up to 1200 meters.
4. Speeds up to 2.5Mbit/Sec.
5. Multi-drop supports up to 32 devices on the same bus.
6. Red power LED.
7. 5V.

C. Rs485 USB

As shown in Figure 3, RS485 to USB converters are used to interface RS485-based devices, such as sensors or controllers, with a computer or microcontroller that uses USB ports for communication. The converter allows data to be transmitted between the RS485 devices (like your soil sensor) and a computer, enabling easier data collection, configuration, and analysis through software. These converters typically handle protocol conversion, enabling seamless communication across different types of systems cost approximately 6$.

Figure 3: RS485 to USB converters.

D. Raspberry Pi 5 Model B+

As shown in Figure 4, Raspberry Pi 5 Model B+ is a single-board computer featuring a quad-core ARM Cortex-A53 CPU running at 1.4 GHz. It has 8GB RAM, integrated Wi-Fi, Bluetooth 4.2, and Ethernet. The B+ model enhances performance with improved network speeds and power management, making it suitable for IoT applications, automation, and interfacing with sensors. It’s commonly used in projects like crop recommendation systems, where it can process data from soil sensors and manage communication between components cost Approximately 100$.

Figure 4: Raspberry Pi 5 Model B+.

E. Arduino Uno

As shown in Figure 5, Arduino Uno is a popular microcontroller board based on the ATmega328P. It features 14 digital I/O pins, 6 analog inputs, a USB connection, and a power jack. It is widely used in DIY electronics projects, allowing users to control sensors, motors, and other devices. In systems like crop recommendation models, the Arduino Uno can collect sensor data, interface with other devices (like an NPK soil sensor), and send it to a central controller like a Raspberry Pi for further processing cost approximately 10$.

Figure 5: Arduino Uno.

F. 12V battery

As shown in Figure 6, a battery is a device that stores chemical energy and converts it into electrical energy costs approximately 10$.

Figure 6: 12V battery.

G. LCD

As shown in Figure 7, a 24x2 LCD is used to display such as the one in the image, it's typically for displaying textual information or simple graphics in embedded systems and electronic projects cost approximately 6$.

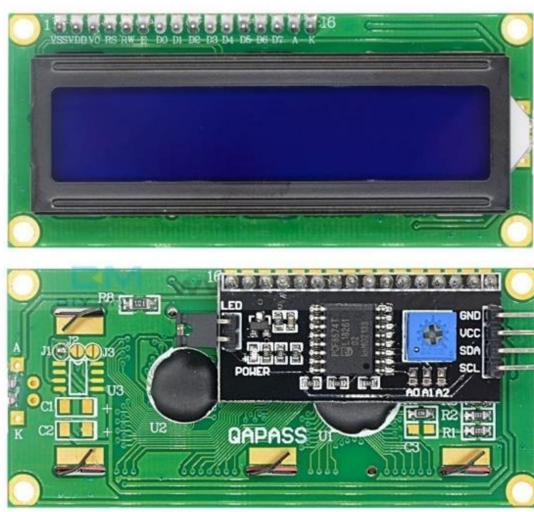


Figure 7: G 24x2 LCD

Here are some common scenarios where you might use such a display:

1. Displaying sensor readings: Showing temperature, humidity, pressure, or other data from connected sensors.
2. User interfaces: Presenting menus, status messages, or prompts for user interaction.
3. Debugging and monitoring: Displaying system status, error codes, or variable values during development.
4. Time and date displays: Creating simple clocks or displaying calendar information.
5. Output from microcontrollers: Showing results of calculations, control system states, or other information generated by a microcontroller (like Arduino or Raspberry Pi Pico).
6. Industrial control panels: Displaying machine status, setpoints, or fault conditions.
7. DIY electronics projects: Incorporating a visual output for a wide range of hobbyist creations.

Technical Specifications and Working Principles

A. NPK 5-Pin Sensor Specifications

Pin 1: VCC (Power Supply) - Typically 5V or 3.3V depending on the sensor.

Pin 2: GND (Ground) - Common ground for all components.

Pin 3: Soil Moisture - Measures the water content in the soil, typically using a capacitive or resistive sensing method.

Pin 4: pH - Measures the acidity/alkalinity of the soil.

Pin 5: NPK (Nitrogen, Phosphorus, Potassium) - These are the macronutrients critical for plant growth. The sensor measures the concentration of these nutrients in the soil.

B. Operating Voltage:

The sensor typically operates at a voltage range of 3.3V to 5V, making it compatible with a range of microcontrollers such as Arduino, Raspberry Pi, etc.

C. Communication Protocol: Analog output (for direct reading of values) or I2C/SPI (if using digital sensors) for communication with the microcontroller or data logger.

D. Measurement Range**:** Soil Moisture: Typically, from 0% to 100%.

pH: Typically, from 0 to 14 (most soil types are within 5-8).

NPK: Concentration levels of Nitrogen, Phosphorus, and Potassium can range from low to high, usually measured in ppm (parts per million).

E. Accuracy: Accuracy can vary based on the model of the sensor, typically ±5% to ±10% for the nutrient levels, ±0.1 for pH, and ±5% for moisture content.

Working Principle of the NPK Sensor

A. Soil Moisture Measurement:

The moisture content is typically detected via a resistive or capacitive measurement method. A probe is inserted into the soil, and the resistance or capacitance is measured, which correlates to the amount of water present in the soil.

B. Soil pH Measurement:

A potentiometric method is usually used where a probe inserted into the soil generates an electric potential that is proportional to the hydrogen ion concentration, giving the pH value.

C. NPK Measurement:

Electrochemical sensors are employed to measure the levels of N, P, and K in the soil. These sensors typically use specific electrodes that respond to the ions of each nutrient and provide a voltage signal proportional to the nutrient concentration.

3.2. Working Principle of the Crop Recommender System

A. Data Collection:

The NPK 5-pin sensor is installed in the soil and continuously sends data about moisture, pH, and nutrient levels to a microcontroller (such as a Raspberry Pi or Arduino).

B. **Data Preprocessing:**

The raw sensor data is processed to correct any errors or noise. For example, soil moisture values may be adjusted based on calibration factors.

C. **Machine Learning Model:**

The processed data is sent to an ML model (e.g., decision trees, SVM, or neural networks), which uses historical crop performance data to suggest the best crop based on the soil's nutrient levels, pH, and moisture. The model takes inputs like current soil conditions and environmental factors, and it outputs crop recommendations tailored to the specific conditions of the field.

D. Feedback Mechanism:

The system could be integrated with IoT devices, allowing real-time feedback and adjustments. For example, irrigation systems could be controlled based on soil moisture levels, and fertilizers can be suggested based on NPK levels.

E. Key Features:

1. Real-time Data Monitoring: The sensor continuously monitors the soil, providing real-time data on important parameters.
2. Scalability: The system can be scaled to monitor large fields or multiple plots by using multiple sensors and integrating them into a central system for processing.
3. Accuracy and Precision: The sensor's ability to provide accurate and reliable data ensures that the recommendations made by the ML model are based on solid, real-time information.
4. Energy Consumption: The system is designed to be energy-efficient, typically powered by low-power devices such as Raspberry Pi or low-power microcontrollers.
5. Solar panels or rechargeable batteries can be used to power the system in field applications.

By combining these sensor specifications and the operational principles, you will give readers a detailed understanding of how the system works, making it easier for them to replicate or implement a similar solution

F. Connecting the sensor to Raspberry Pi/Arduino

Connecting the NPK 5-Pin Sensor to Arduino:

Materials Needed:

- NPK 5-pin soil sensor
- Arduino board (e.g., Arduino Uno, Arduino Nano, etc.)
- Jumper wires
- Breadboard (optional)

1. Step 1: Pinout of the NPK 5-Pin Sensor

The NPK 5-pin sensor typically has the following pins:

2. VCC (Power Supply)
3. GND (Ground)
4. Soil Moisture Output
5. pH Output
6. NPK Output
7. Step 2: Wiring the NPK Sensor to Arduino

To connect the NPK sensor to the Arduino, you'll follow this wiring diagram:

For Soil Moisture, pH, and NPK Data:

1. VCC (Pin 1):
   - Connect this to the 5V pin on the Arduino. (Note: Some sensors may operate at 3.3V, so check your sensor's specifications).
2. GND (Pin 2):
   - Connect this to the GND pin on the Arduino.
3. Soil Moisture Output (Pin 3):
   - Connect this to one of the analog input pins on Arduino (e.g., A0). This will be used to read the moisture data.
4. pH Output (Pin 4):
5.
   - Connect this to another analog input pin (e.g., A1) to read pH levels.
6. NPK Output (Pin 5):
   - Connect this to another analog input pin (e.g., A2) to measure NPK levels.
   - Step 3: Writing the Arduino Code

Once the wiring is done, you can use the following sample code to read the sensor values and display them on the Serial Monitor.

Detailed Code Explanation (Section 1: Initialization and Setup)

This initial segment of the code establishes the fundamental environment for the microcontroller (e.g., an Arduino board) before any complex operations commence. It meticulously covers the inclusion of necessary software libraries, the declaration of essential variables, and subsequently, the comprehensive initialization of all hardware and software components within the setup() function. The code begins by integrating crucial libraries that provide pre-built functionalities for interacting with peripheral devices. This includes the Wire.h library, fundamental for I2C communication, a widely used protocol for interfacing with various devices, notably LCD screens equipped with I2C converters to simplify wiring. Additionally, the hd44780.h and hd44780_I2Cexp.h libraries are included; these are specifically tailored for controlling HD44780-based Liquid Crystal Displays, with the latter focusing on I2C-connected variants. The SoftwareSerial.h library is also incorporated, enabling the creation of virtual UART (Universal Asynchronous Receiver-Transmitter) serial ports on any pair of digital pins. This feature is particularly advantageous when the hardware serial port is occupied or insufficient, and it is likely employed here for communication with the soil sensor. Following the library inclusions, global variables are declared: hd44780_I2Cexp lcd; instantiates an lcd object, serving as the programmatic interface for sending commands and data to the LCD. SoftwareSerial mySerial(5, 2); creates a software serial object named mySerial, dedicating digital pin 5 as the receive (RX) pin and digital pin 2 as the transmit (TX) pin for sensor communication. Finally, int DE = 3; and int RE = 4; define control pins for an RS485 converter module, used to manage data flow direction. The setup() function then executes upon microcontroller startup, performing all essential initializations. It begins by configuring the 16x2 LCD, verifying successful initialization (halting the program and reporting an error if unsuccessful), and activating its backlight. An "Initializing..." message is displayed temporarily. Concurrently, both the standard hardware serial communication (for computer monitoring) and the mySerial software serial communication (for the sensor) are started at a baud rate of 9600. The RS485 control pins (DE and RE) are configured as outputs and set to a low logical state, typically placing the RS485 module in receive mode, ready to acquire data from the sensor. A brief 2-second delay is introduced to allow the initialization message to be seen, after which the LCD is cleared, preparing it for live sensor data display. In essence, this initial code segment diligently prepares the entire electronic system—LCD, serial communications, and RS485 module—ensuring everything is properly configured before the main program loop commences the continuous cycle of requesting, reading, and displaying soil parameters.

Step 4: Testing the Setup & Step 5: Calibration and Tuning

With the hardware connections complete and the Arduino code prepared, the next crucial phase involves Testing the Setup to verify its functionality. This begins by connecting the Arduino board to your computer via USB and then Uploading the Code using the Arduino IDE. Once the code is successfully transferred to the board, you should Open the Serial Monitor (accessible via "Tools > Serial Monitor" or by pressing Ctrl+Shift+M). If the setup is working correctly, you will immediately start to see real-time readings for Soil Moisture, pH, and NPK values being continuously printed to the Serial Monitor, typically updated every second, indicating that the sensor is actively acquiring data. Following initial validation, Calibration and Tuning become essential for ensuring the accuracy and reliability of the sensor readings. For Soil Moisture, direct correspondence between raw sensor values and actual percentage moisture may not be precise. Therefore, calibration is recommended by comparing readings against known reference soil moisture levels or by testing the sensor in various controlled moisture conditions (e.g., dry, moist, saturated soil) to establish a conversion curve. Similarly, Calibration for pH and NPK levels is critical. This can be achieved by utilizing standard reference solutions with known pH or nutrient concentrations, or by cross-referencing sensor outputs with laboratory test results from soil samples. This meticulous calibration process is vital for translating raw sensor data into meaningful and accurate insights for effective crop management.

1. Required Components for the NPK 5-Pin Sensor Setup

To successfully set up and interface the NPK 5-pin soil sensor for measuring crucial parameters such as soil moisture, pH, and NPK levels, a specific set of components is essential. At the core of the system is the Microcontroller: either an Arduino Uno, which is well-suited due to its inherent analog input capabilities, or a Raspberry Pi, which would necessitate the addition of an Analog-to-Digital Converter (ADC) Module (like the MCP3008) to process the analog signals from the sensor. Naturally, the NPK 5-Pin Soil Sensor itself is indispensable. Connections between these various components are facilitated by Jumper Wires, including both male-to-male and male-to-female types. For easy prototyping and circuit assembly, a Breadboard is an optional but highly recommended addition. Powering the system requires a dedicated Power Source: typically USB power for the Arduino, or a 5V adapter or power bank for the Raspberry Pi. Lastly, Resistors might be necessary for voltage level shifting or division, ensuring compatibility between the sensor's output voltage and the microcontroller's input specifications (e.g., 10kΩ resistors could be used for voltage division if required). Each of these components plays a vital role in establishing a functional and accurate soil sensing system.

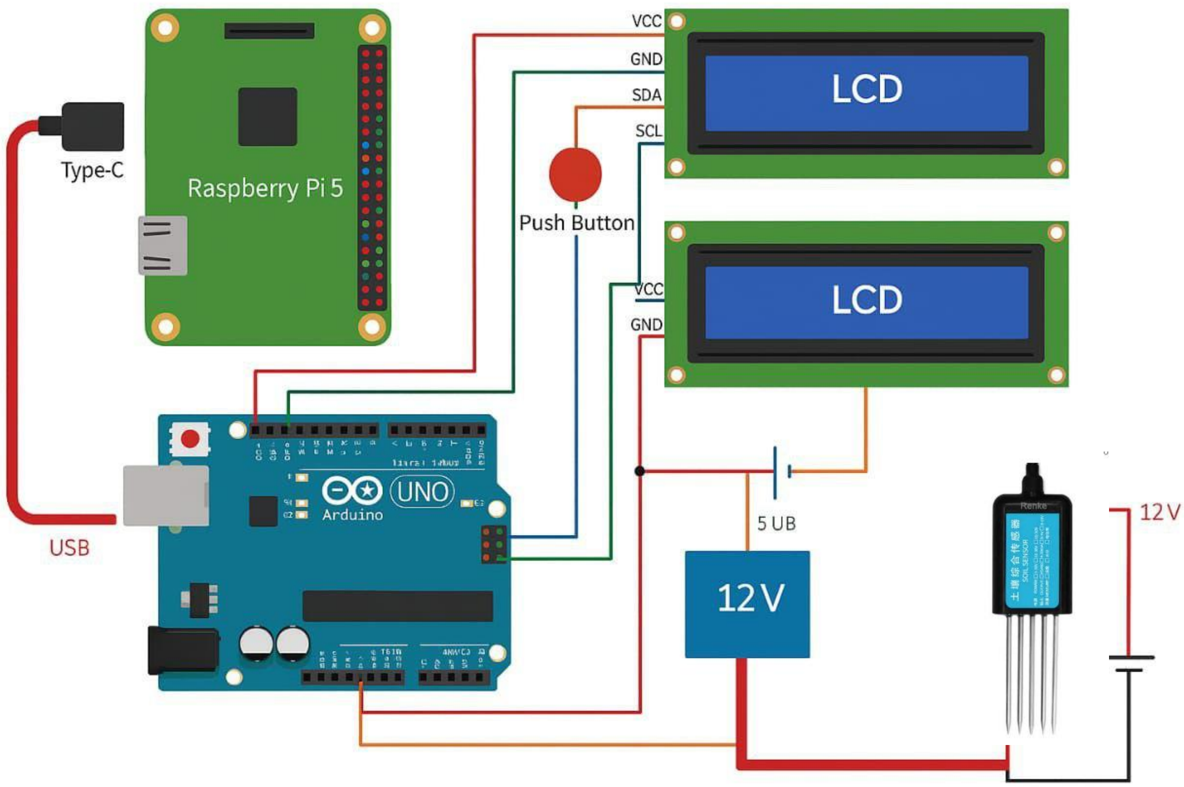


Figure 8: Circuit Diagram for hardware connection

Soil NPK Sensor Specifications:
NPK Sensor Pinout:

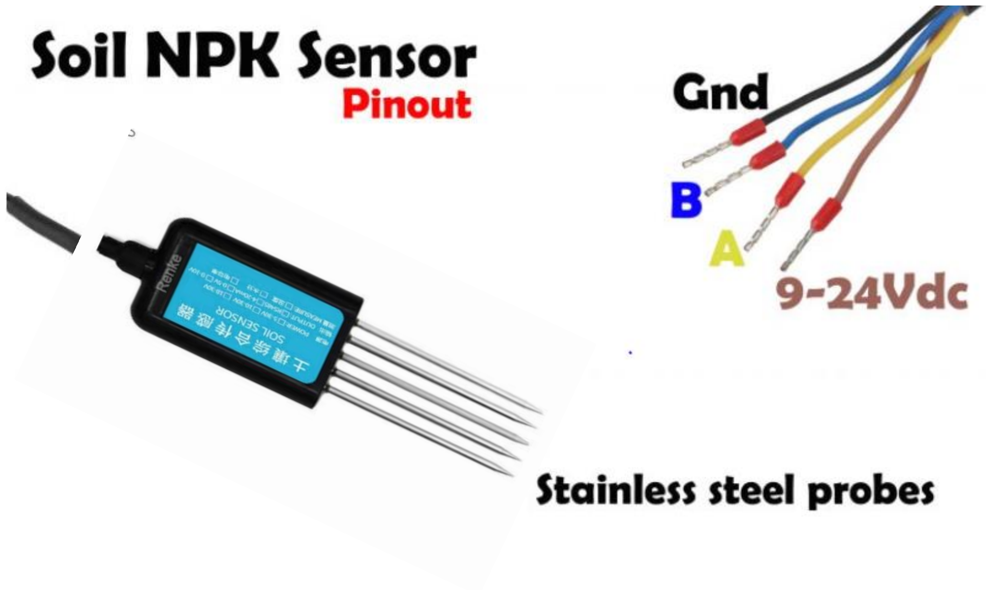


Figure 9: NPK five pins sensor

The Soil NPK Sensor has a total of 4 wires. The brown wire is the VCC wire and it should be connected with 9V-24Vdc Power Supply. The Black wire is the GND wire and it should be connected with the Arduino's GND. The remaining two wires which are the Blue and Yellow wires these are the B and A wires and these two wires should be connected with the B and A pins of the Max485 Modbus module which I will explain in a minute.

So, You will need 9 to 24Vdc to power up this Soil NPK Sensor. The NPK Sensor supports 2400, 4800, and 9600 baud rates, due to which it can be used with different microcontroller boards like 8051 family of microcontrollers, PIC microcontrollers, Arduino boards, and so on. In this tutorial, I will use the Soil NPK Sensor with the Arduino board. The Soil NPK Sensor is provided with the Modbus communication port RS485 due to which it can be easily interfaced with the Arduino board using the Modbus module like MAX485/RS485 module. The working temperature is from 5 to 45 Celsius. The Nitrogen, phosphorus, and Potassium resolution is 1mg/kg or 1mg/liter. The measuring range of the Soil NPK Sensor is 0 to 1999mg/kg, and the working humidity is from 5 to 95%. The maximum power consumption is ≤ 0.15W.

A. Voltage:
   9V-24V DC
B. Maximum Power Consumption: ≤ 0.15W

C. Baud Rate:
   2400/4800/9600
D. Working Temperature:
   5 to 45 ° C
E. Resolution:
   1mg/kg (mg/l)
F. Measuring Range:
   0-1999mg/kg
G. Working Humidity:
H. 5 to 95% (relative humidity), no condensation
 Measurement Accuracy:
I. ±2%F.s
   Communication Port:
J. RS485
   Protection Class:
K. IP68

2. Circuit Diagram

The circuit connections for both Arduino and Raspberry Pi setups are described below.

Table 2.Connecting NPK Sensor to Arduino Wiring:

NPK Sensor Pin	Arduino Pin
VCC (Power)	5V
GND (Ground)	GND
Soil Moisture	A0 (Analog Input)
pH Output	A1 (Analog Input)
NPK Output	A2 (Analog Input)

B. Connecting NPK Sensor to Raspberry Pi

Since the Raspberry Pi does not have analog input pins, an ADC module (MCP3008) is required.

Table 3 :Wiring with MCP3008 ADC Module:

NPK Sensor Pin	MCP3008 ADC Pin	Raspberry Pi Pin
VCC (Power)	VCC (3.3V)	3.3V
GND (Ground)	GND	GND
Soil Moisture	CH0	SPI Data Input (MOSI)
pH Output	CH1	SPI Data Input (MOSI)
NPK Output	CH2	SPI Data Input (MOSI)

Note :

 For Arduino: Direct connection to A0, A1, A2 (analog pins).
 For Raspberry Pi: Use MCP3008 ADC to convert analog signals to digital.

 Both setups require 5V power and ground connections.

1. Data Collection from the Sensor

1. Overview

The NPK 5-pin soil sensor provides real-time soil parameter data, including:

- Soil Moisture
- pH Level
- Nutrient Content (NPK - Nitrogen, Phosphorus, Potassium)

This data can be collected using Arduino or Raspberry Pi and processed for crop recommendation and soil health analysis.

2. Data Collection with Arduino

Since Arduino has analog input pins, it can directly read sensor values.

A. Reading Sensor Data

The sensor outputs analog voltage corresponding to soil conditions. These readings can be collected using the analog Read() function.

B. Arduino Code for Data Collection

Upload the following Arduino sketch to read and display sensor values

C. Data Interpretation

- Soil Moisture**:** Values typically range from 0 (dry soil) to 1023 (fully wet)**.**
- pH Level: A neutral soil pH should read around **7**, acidic soil will be lower, and alkaline soil will be higher.
- **NPK Values**: Each nutrient (Nitrogen, Phosphorus, Potassium) will have different reference voltage levels that may require calibration.

3. Data Collection with Raspberry Pi

The Raspberry Pi lacks analog input pins, so an Analog-to-Digital Converter (ADC) like MCP3008 is needed.

A. Setting Up the MCP3008 ADC

To read analog values from the NPK sensor, connect it to the MCP3008 and then to the Raspberry Pi's SPI (Serial Peripheral Interface) pins.

B. Python Code for Data Collection on Raspberry Pi:

This system aims to acquire a set of vital data from a specialized soil sensor, including moisture level, temperature, electrical conductivity, pH, and the levels of essential nutrients such as nitrogen, phosphorus, and potassium. After collecting this data, it is clearly displayed on a Liquid Crystal Display (LCD) for direct on-site monitoring, as well as on the computer's Serial Monitor for logging and analysis purposes.

1. Initial Setup and Library Inclusions

At the beginning of any Arduino program, this section forms the fundamental groundwork. Here, the necessary programming libraries are invoked, providing the functionalities required to interact with hardware components. This includes libraries for controlling LCD screens that communicate via the I2C protocol, along with a library for Software Serial communication, which enables communication with the sensor through specific digital pins. This section also defines the essential variables that will later be used to specify the pin numbers connected to external components, such as the control pins for the RS485 module responsible for directing data flow.

2. The setup() Function (Initialization)

The setup() function is considered the beating heart of the program at startup, executing only once. Its primary role is to configure and initialize all components before the main operations begin. Within this function, the following actions are performed:

- **LCD Screen Initialization:** The dimensions of the screen (e.g., 16 columns x 2 rows) are defined, and the success of the initialization process is confirmed. In case of any issues, the program halts and displays an error message on the Serial Monitor.
- **Backlight Activation:** The LCD screen's backlight is turned on to ensure clear visibility.
- **Displaying a Startup Message:** An "Initializing..." message is shown on the screen to inform the user that the system is preparing.
- **Serial Communication Setup:** Serial communication is initiated with both the computer (for monitoring purposes) and the sensor (for data exchange) at the same specified baud rate.
- **Control Pin Configuration:** The control pins for the RS485 module are defined as outputs, and their initial state is set to "receive mode," preparing them to receive incoming data.
- **Delay and Screen Clear:** A brief delay is introduced to allow the initialization message to be displayed, after, which the screen is cleared in preparation for displaying the actual data.

3. The loop() Function (Main Operating Loop)

This function serves as the primary engine of the system, executing continuously after the setup() function has completed. This loop contains the entire programmatic logic for data acquisition, processing, and display, and is divided into several stages:

- **Sending Data Request:** A specific data packet (query) is constructed containing the necessary commands to request readings from the soil sensor.
- **Setting Transmit Mode:** The state of the RS485 module's control pins is changed to "transmit mode," allowing the query to be sent to the sensor.
- **Transmitting the Query:** The query packet is sent via the software serial connection to the sensor.
- **Setting Receive Mode:** Immediately after sending the query, the RS485 control pins are reset to "receive mode," preparing them to receive the sensor's response.
- **Waiting and Receiving Data:** The program waits for a specified period to allow the sensor to process the request and send its response. Once a sufficient amount of received data is available, it is read and stored in a designated array.
- **Extracting and Analyzing Data:** The received data is raw and therefore needs to be parsed and analyzed to extract numerical values. This involves combining consecutive bytes to form complete values for humidity, temperature, conductivity, pH, as well as nitrogen, phosphorus, and potassium levels. Some of these values are then processed further by dividing them by a conversion factor (e.g., 10.0) to obtain precise decimal readings (e.g., 25.0 instead of 250).
- **Displaying Data on the Serial Monitor:** All extracted values are neatly printed on the computer's Serial Monitor. This step is crucial for debugging and detailed value monitoring.
- **Displaying Data on the LCD Screen (First Screen):** The LCD screen is cleared, and then the first set of data, including humidity, temperature, conductivity, and pH, is displayed. These values are shown in a clear format with abbreviated labels and remain visible for a specified duration.
- **Displaying Data on the LCD Screen (Second Screen - Nutrients):** After the display time for the first screen ends, the screen is cleared again. Then, the second set of data, comprising nitrogen, phosphorus, and potassium levels, is displayed. These values are shown in the same clear format and remain visible for a longer duration.
- **Repetition:** After both screens have been displayed, the loop function returns to its beginning, starting a new cycle of data requests and display updates, ensuring continuous monitoring of soil conditions.

C. Data Interpretation

- Soil Moisture: The ADC will return values between 0-1023, which should be calibrated to match real moisture levels.
- pH Level**:** Needs calibration against known pH solutions**.**
- NPK Levels: Interpretation depends on sensor calibration data**.**

4. Data Storage and Transmission

Once collected, the data can be:

- Stored locally (e.g., CSV file, SQLite database).
- Sent to a cloud service for remote monitoring.
- Used for machine learning in a crop recommendation system.

Example: Storing Data in a CSV File (Python)

```
import csv
from datetime import datetime

with open("soil_data.csv", "a") as file:
    writer = csv.writer(file)
    writer.writerow([datetime.now(), soil_moisture, ph_value, npk_value])
```

5. Understanding Soil Parameters

Soil health plays a crucial role in plant growth, influencing water retention, nutrient availability, and overall crop productivity. The NPK sensor measures key soil parameters to assess soil conditions for better crop management.

5.1 Soil Moisture

Definition:
Soil moisture refers to the amount of water present in the soil, which is essential for plant hydration and nutrient transport.

Importance:

- Affects seed germination and root development.
- Determines irrigation needs.
- Impacts microbial activity and soil aeration.

Measurement:

- Measured using soil moisture sensors, which detect voltage changes based on water content.
- Expressed as percentage (%) or volumetric water content (VWC).

Ideal Soil Moisture Levels:

Table 4: Ideal Soil Moisture Levels

Soil Type	Ideal Moisture (%)
Sandy Soil	5 – 20%
Loamy Soil	20 – 50%
Clay Soil	30 – 70%

5.2 Humidity

Definition:
Humidity refers to the amount of water vapor in the air surrounding the soil and plants.

Importance**:**

- Affects plant transpiration and nutrient uptake.
- Influences soil evaporation rates.
- Helps in disease prevention (e.g., fungal infections thrive in high humidity).

Measurement:

- Measured in Relative Humidity (RH%) using DHT11/DHT22 sensors**.**

Optimal Humidity Levels for Crops:

Table 5: Optimal Humidity Levels for Crops

Crop Type	Ideal Humidity (%)
Vegetables	50 – 70%
Fruits	60 – 80%
Grains	40 – 60%

5.3 Temperature

Definition:
Soil temperature influences seed germination, root growth, and microbial activity.

Importance:

- Controls enzyme activity for nutrient availability.
- Impacts water retention and evaporation rates.
- Determines plant stress levels.

Measurement:

- Measured using DS18B20 temperature sensors.
- Expressed in Celsius (°C) or Fahrenheit (°F).

Optimal Soil Temperature Ranges:

Table 6: Optimal Soil Temperature Ranges

Crop Type	Ideal Temperature (°C)
Wheat	15 – 25°C
Corn	18 – 30°C
Potatoes	12 – 20°C

5.4 Electric Conductivity (EC)

Definition:
Electric Conductivity (EC) measures the salt concentration in soil, indicating nutrient availability.

Importance:

- High EC indicates excess salts or fertilizers, which can damage plant roots.
- Low EC suggests poor soil fertility.
- Helps determine irrigation water quality.

Measurement:

- Measured in dS/m (decisiemens per meter).
- Read using EC sensors that detect soil ion concentration.

Ideal EC Levels:

Table 7: Ideal EC Levels

Soil Type	Ideal EC (dS/m)
Sandy Soil	0.2 – 1.0
Loamy Soil	1.0 – 2.5
Clay Soil	2.5 – 4.0

5.5 pH

Definition:
pH measures the acidity or alkalinity of soil, affecting nutrient availability.

Importance:

- Determines plant nutrient absorption.
- Influences microbial activity in the soil.
- Affects fertilizer efficiency.

Measurement:

- Measured using pH sensors.
- pH range: 0 (acidic) – 14 (alkaline), with 7 being neutral.

Ideal pH Levels:

Table 8: Ideal pH Levels

Crop Type	Ideal pH Range
Rice	5.0 – 6.5
Tomatoes	5.5 – 6.8
Wheat	6.0 – 7.5

5.6 Nitrogen (N), Phosphorus (P), Potassium (K) – NPK Levels

(A) Nitrogen (N)

Function:

- Essential for leaf growth and photosynthesis.
- Major component of chlorophyll.

Deficiency Symptoms:

- Yellowing leaves.
- Stunted growth.

Ideal Levels:

Table 9: Nitrogen (N) Ideal Levels

Soil Type	Nitrogen Content (mg/kg)
Sandy Soil	20 – 40
Loamy Soil	40 – 60
Clay Soil	60 – 100

(B) Phosphorus (P)

Function:

- Promotes root development and flowering.
- Improves seed and fruit quality.

Deficiency Symptoms:

- Poor root growth.
- Dark green or reddish leaves.

Ideal Levels:

Table 10: Phosphorus (P) Ideal Levels

Soil Type	Phosphorus Content (mg/kg)
Sandy Soil	10 – 20
Loamy Soil	20 – 40
Clay Soil	40 – 80

(C) Potassium (K)

Function:

- Enhances drought resistance.
- Improves fruit quality and disease resistance.

Deficiency Symptoms:

- Brown leaf edges.
- Weak stems and slow growth.

Ideal Levels:

Table 11: Potassium (K)

Soil Type	Potassium Content (mg/kg)
Sandy Soil	50 – 100
Loamy Soil	100 – 150
Clay Soil	150 – 250

Note:

Understanding soil parameters is crucial for precision farming, allowing farmers to optimize irrigation, fertilization, and overall crop health management. Using real-time data from NPK sensors, farmers can make informed decisions, leading to better yields and sustainable farming practices.

2. Role of AI in Crop Recommendation

Artificial Intelligence (AI) is transforming modern agriculture by analyzing soil data, predicting crop suitability, and optimizing farming decisions**.** AI-powered **crop** recommendation systems use machine learning (ML) algorithms to analyze soil parameters and suggest the best crops for higher yield and sustainability.

2.1 Importance of Soil Analysis for Agriculture

Soil analysis is the foundation of precision farming. It provides crucial insights into soil health**,** enabling farmers to make informed decisions about crop selection**,** fertilization, and irrigation**.**

Key Benefits of Soil Analysis:

1. Optimal Crop Selection: Determines which crops are best suited for a particular soil type.
2. Efficient Fertilization**:** Identifies nutrient deficiencies (NPK levels) and prevents over-fertilization.
3. Water Management**:** Helps in scheduling irrigation based on soil moisture levels.
4. Soil pH Balance: Ensures that crops grow in the right pH range to maximize nutrient uptake.
5. Prevention of Soil Degradation**:** Helps in sustainable land use and conservation of soil health.

2.2 Role of AI in Crop Recommendation

Artificial Intelligence (AI) is transforming modern agriculture by analyzing soil data, predicting crop suitability, and optimizing farming decisions. AI-powered crop recommendation systems use machine learning (ML) algorithms to analyze soil parameters and suggest the best crops for higher yield and sustainability.

Importance of Soil Analysis for Agriculture: Soil analysis is the foundation of precision farming. It provides crucial insights into soil health, enabling farmers to make informed decisions about crop selection, fertilization, and irrigation.

Key Benefits of Soil Analysis:

1. Optimal Crop Selection: Determines which crops are best suited for a particular soil type.
2. Efficient Fertilization: Identifies nutrient deficiencies (NPK levels) and prevents over-fertilization.
3. Water Management: Helps in scheduling irrigation based on soil moisture levels.
4. Soil pH Balance: Ensures that crops grow in the right pH range to maximize nutrient uptake.
5. Prevention of Soil Degradation: Helps in sustainable land use and conservation of soil health.

2.4 Challenges in Traditional Farming Methods

Traditional farming relies on experience-based decision-making, which often leads to inefficiencies and low productivity. Some of the major challenges include:

1. Lack of Precision:

- Farmers rely on trial and error instead of data-driven insights.
- Wrong crop selection leads to low yields and financial losses.

2. Unbalanced Use of Fertilizers:

- Overuse of Nitrogen (N), Phosphorus (P), and Potassium (K) depletes soil quality.
- Underuse leads to nutrient deficiency and poor plant growth.

3. Water Mismanagement:

- Over-irrigation can lead to waterlogging and soil erosion.
- Under-irrigation results in drought stress, reducing yield.

4. Climate Change Uncertainty:

- Unpredictable weather conditions affect soil moisture, temperature, and nutrient levels.
- Traditional methods do not adapt to changing climate patterns.

5. Labor-Intensive Processes:

- Manual soil testing and field inspections take time and effort.
- AI-powered automation can save labor and increase efficiency.

2.3 Fields Where Crop Recommendation Systems Are Used

AI-driven crop recommendation systems are widely applied in various agricultural domains to enhance productivity and sustainability.

1. Precision Farming

- AI helps analyze soil health, weather patterns, and crop suitability.
- Farmers get real-time insights for better crop management.

2. Smart Irrigation Systems

- AI models optimize water usage based on soil moisture and weather predictions.
- Reduces water waste and increases crop efficiency.

3. Sustainable Agriculture

- Prevents soil degradation by recommending crop rotation and organic farming practices.
- Ensures long-term soil fertility.

4. Agri-Tech and Digital Farming Platforms

- Mobile apps use AI to suggest crops based on soil test reports.
- Farmers can access data-driven recommendations remotely.

5. Government and Research Institutions

- AI-powered agricultural research helps in policy-making.

- Governments use AI to promote sustainable crop selection strategies.

2.4 How AI Can Help in Crop Recommendation

AI-powered crop recommendation systems analyze soil parameters, climate conditions, and historical data to suggest the best crops for a given field.

AI Techniques in Crop Recommendation:

1. Machine Learning (ML) Algorithms
    - AI models like Random Forest, Decision Trees, and Neural Networks analyze soil data and recommend crops.
    - Example: If a field has high nitrogen content and a pH level of 6.5, the system may suggest corn or wheat.
2. Deep Learning for Pattern Recognition
    - Identifies hidden relationships between soil nutrients, temperature, and rainfall.
    - Helps predict crop diseases and soil degradation trends.
3. Big Data Analytics in Agriculture
    - AI processes large datasets from soil sensors, satellites, and weather stations.
    - Farmers receive real-time predictions for crop selection.

4. Computer Vision for Soil Analysis

- AI-based image processing detects soil texture, color, and fertility.
- Used in automated soil testing.

5. Internet of Things (IoT) in Smart Farming

- IoT sensors collect soil moisture, pH, and NPK data in real time.
- AI integrates this data to provide personalized crop recommendations.

Example: AI-Powered Crop Recommendation System

Table 12: AI system that analyzes the following soil parameters

Parameter	Measured Value	AI Recommendation
pH Level	6.2	Wheat, Barley
Nitrogen (N)	High	Corn, Rice
Phosphorus (P)	Moderate	Soybean, Maize
Potassium (K)	Low	Potatoes, Beans
Soil Moisture	40%	Suitable for most crops

Using this information, the AI system suggests the most suitable crops for the given soil condition.

3. Data Acquisition and Preprocessing

Machine learning models for crop recommendation rely on high-quality data. The process starts with collecting sensor data, storing it efficiently, cleaning it, and preparing it for training.

3.1 Methods for Collecting Sensor Data

To train an ML model for crop recommendation, real-time and historical data are collected from various sources, including:

1. Soil Sensors (NPK, pH, Moisture, EC, etc.)

- Devices like 5-pin NPK sensors measure key soil parameters.
- Connected to Raspberry Pi or Arduino for real-time data transmission.

2. Weather APIs and Satellite Data

- Online APIs provide temperature, humidity, and rainfall data.
- Example: OpenWeatherMap API for real-time weather data.

3. IoT-Based Smart Farming Systems

- Wireless sensor networks (WSN) send data to cloud-based platforms.
- Example: LoRaWAN or MQTT-based IoT devices.

4. Government and Agricultural Databases

Open datasets from organizations like FAO, NASA, and local agricultural research centers.

3.2 Storing Data (CSV, Database, Cloud Integration)

Once the data is collected, it needs to be stored in a structured way. There are several options:

Data Storage Formats for Sensor Data

When managing sensor data, choosing an appropriate storage format is crucial for efficiency and scalability. Two common approaches are CSV files and SQL databases, each with distinct advantages and disadvantages. CSV (Comma Separated Values) files offer a simple and accessible storage format. Their primary benefit lies in their ease of handling, making them ideal for small datasets where quick viewing and basic manipulation are sufficient. For instance, sensor data can be effortlessly saved to a CSV file using libraries like Pandas in Python, as demonstrated by df.to_csv('soil_data.csv', index=False). However, CSVs are not scalable for very large datasets; querying specific information or managing complex relationships within vast amounts of data becomes inefficient. In contrast, SQL (Structured Query Language) databases provide a robust and structured storage solution. They are highly efficient for querying, managing, and retrieving information from large and complex datasets. This is evident in their ability to define tables with specific data types, such as the Soil Data table example, which meticulously organizes sensor readings like N, P, K, temperature, humidity, pH, and rainfall, alongside a crop label. While SQL databases excel in scalability and data integrity, their main drawback is the necessity for initial setup and ongoing maintenance, which can require more technical expertise compared to the simplicity of CSV files.

3. Cloud Integration (Real-Time Processing)

- Platforms like Google Firebase, AWS, or Azure store large-scale data.
- Ensures global access and scalability.

3.3 Data Cleaning and Handling Missing Values

.

Table 13: Raw sensor data may have errors, inconsistencies, or missing values.

Issue	Solution
Missing Values	Fill with **mean, median, or mode**
Duplicate Entries	Remove duplicates using Pandas
Outliers	Use **Z-score or IQR filtering**

Machine Learning Model Development

Once the data is preprocessed, we train an ML model to predict the best crop based on soil parameters.

3.5 Overview of ML Algorithms for Crop Recommendation

In the realm of agricultural technology, machine learning (ML) plays a pivotal role in intelligent crop recommendation systems. Collected soil sensor data, after undergoing crucial storage and preprocessing, serves as the foundation for training these ML models. Several diverse ML techniques are explored to recommend optimal crops based on various soil parameters. Decision Trees, for instance, operate by splitting data based on specific conditions (e.g., if Nitrogen levels exceed 50, then rice might be recommended). While highly interpretable, a single Decision Tree can be prone to overfitting, meaning it might perform well on training data but poorly on unseen data. To mitigate this, Random Forest models are employed, leveraging the power of multiple decision trees to enhance overall accuracy and significantly reduce overfitting. For datasets with numerous features, Support Vector Machines (SVM**)** are often preferred due to their effectiveness in high-dimensional spaces, though they can be computationally intensive. Finally, Neural Networks**,** a cornerstone of deep learning, are particularly adept at handling complex datasets and intricate patterns, but they typically demand substantial amounts of training data for optimal performance. Regardless of the chosen model, the general workflow involves three critical steps: Splitting Data into training and testing sets to evaluate model performance on unseen data, Model Training where the algorithm learns patterns from the training data, and Model Evaluation to assess the model's accuracy and effectiveness in making crop recommendations.

3.6 Performance Evaluation (Accuracy, Precision, Recall, F1-Score)

After training the ML model, it's crucial to evaluate its performance to ensure accurate crop recommendations.

Table 14: Key Performance Metrics

Metric	Definition	Formula
Accuracy	Measures the proportion of correctly classified crops.	$\frac{TP+TN}{TP+TN+FP+FN}$
Precision	Measures how many predicted crops are actually correct.	$\frac{TP}{TP+FP}$
Recall (Sensitivity)	Measures how well the model identifies the correct crop.	$\frac{TP}{TP+FN}$
F1-Score	Harmonic mean of Precision and Recall.	$2 \times \frac{Precision \times Recall}{Precision+Recall}$

Evaluating Machine Learning Model Performance: Beyond Simple Accuracy

When assessing the effectiveness of machine learning models, particularly in classification tasks like crop recommendation, understanding the nuances of prediction outcomes is crucial. These outcomes are categorized into four types: True Positives (TP) represent instances where the model correctly identified a crop that was indeed the correct recommendation. Conversely, True Negatives (TN) occur when the model accurately rejected an incorrect crop prediction. Errors in prediction fall into two categories: False Positives (FP), where the model incorrectly predicted a crop that was not suitable, and False Negatives (FN), which are critical cases where the model missed recommending a crop that was actually the correct choice. While overall prediction accuracy might seem like a straightforward metric, it can be misleading, especially when dealing with imbalanced datasets—situations where some crop types appear significantly more frequently than others. In such scenarios, a model might achieve high accuracy simply by correctly predicting the majority class, while performing poorly on the less represented, but potentially important, minority classes. This is precisely Why Use F1-Score? The F1-Score is a more robust and informative metric that considers both precision (the proportion of true positive predictions that were actually correct) and recall (the proportion of actual positive cases that were correctly identified). By harmonically averaging these two values, the F1-Score provides a balanced measure of a model's performance, making it a superior choice over simple accuracy when dealing with imbalanced datasets to ensure the model performs well across all classes, not just the dominant ones.

3.7 Model Optimization Techniques: Refining Performance

After a machine learning model has been trained and evaluated, the crucial next step is to optimize its performance to achieve the best possible results. This process primarily involves Hyper parameter Tuning, which is the strategic adjustment of a model's external parameters (hyper parameters) to significantly enhance its accuracy and generalization capabilities. There are several widely used methods for conducting hyper parameter tuning. Grid Search offers a systematic approach, exhaustively testing every possible combination of specified parameter values within a defined grid. This ensures that the optimal combination within that grid is found, as illustrated by the example where GridSearchCV is used to explore different n_estimators, max_depth, and min_samples_split for a Random-Forest-Classifier. While thorough, Grid Search can be computationally expensive for large search spaces. Alternatively, Random Search randomly samples parameter values from defined distributions, often proving more efficient than Grid Search by exploring a wider range of parameter values in the same amount of time. A more advanced technique is Bayesian Optimization, which employs a probabilistic model to intelligently guide the search for optimal parameters, aiming to find the best configuration with fewer evaluations. By meticulously tuning these hyper parameters, the model can be fine-tuned to extract maximum insight from the data, leading to more reliable and accurate crop recommendations.

2. Feature Selection

Beyond tuning hyper parameters, another critical aspect of model optimization is Feature Selection. This process recognizes that not all input features contribute equally to the accuracy and efficiency of crop prediction. Including irrelevant or redundant features can introduce noise, increase computational complexity, and potentially reduce the model's overall performance. Therefore, identifying and removing these less influential features is vital for building a more robust and parsimonious model. One highly effective method for achieving this, particularly with ensemble models, is Using Feature Importance in Random Forest. As demonstrated in the provided code snippet, a Random Forest classifier inherently calculates the importance of each feature during its training phase. This importance score quantifies how much each feature contributes to the reduction of impurity (or error) across all the trees in the forest. By accessing the model.feature_importances_ attribute, one can obtain these scores, which can then be visualized (e.g., using a bar plot with matplotlib.pyplot) to identify and prioritize the most impactful soil parameters for crop recommendation. Features with very low importance scores can subsequently be removed from the dataset, leading to a simpler model that is faster to train, less prone to overfitting, and potentially more accurate by focusing on the most relevant information.

3. Data Augmentation

When dealing with machine learning problems, particularly in domains where data collection can be challenging or costly, a common hurdle is the availability of a sufficiently large dataset. If the dataset is small, the trained model might struggle to generalize well to new, unseen data, leading to suboptimal performance. In such scenarios, Data Augmentation emerges as a valuable optimization technique. This involves increasing the size and diversity of the training dataset by generating synthetic data points. One powerful method for this, especially relevant for addressing class imbalance within a dataset, is SMOTE (Synthetic Minority Over-sampling Technique). SMOTE works by creating new synthetic examples for the minority class(es) based on existing examples. It does this by taking samples from the minority class and introducing synthetic examples along the line segments joining any of the k-nearest neighbors of that sample. This technique effectively balances the class distribution and provides the model with more data to learn from, ultimately leading to improved model performance, better generalization, and enhanced accuracy, particularly for less represented classes.

4. Regularization Techniques

Further enhancing model performance and preventing the common issue of overfitting—where a model learns the training data too well, failing to generalize to new data—is achieved through Regularization Techniques. These methods add a penalty to the model's loss function, discouraging overly complex models. L1 Regularization (Lasso), for instance, has the unique property of forcing some model coefficients to become exactly zero, effectively performing automatic feature selection by excluding less important features. In contrast, L2 Regularization (Ridge) works by shrinking the coefficients towards zero without necessarily making them exactly zero, which helps prevent excessively large weights that can contribute to overfitting. Beyond individual model refinements, Ensemble Learning offers a powerful strategy for boosting predictive accuracy by combining multiple models.

Two primary approaches exist: Bagging, exemplified by Random Forest, trains multiple models in parallel on different subsets of the data and averages their predictions to reduce variance. Conversely, Boosting, demonstrated by algorithms like XGBoost, trains models sequentially, with each new model learning to correct the mistakes of its predecessors, thereby focusing on misclassified instances and progressively improving overall performance. In summary, while Performance Metrics such as Accuracy, Precision, Recall, and F1-Score are vital for measuring a model's effectiveness, a comprehensive suite of Optimization Techniques—including Hyper parameter Tuning, Feature Selection, Data Augmentation methods like SMOTE, Regularization, and Ensemble Learning—are instrumental in significantly improving a model's accuracy, robustness, and ability to generalize well to unseen data.

Chapter 4

Applications and Benefits of the Soil Analysis Device

1. Overview

The modern agricultural sector is facing significant challenges due to soil degradation, water scarcity, and inefficient crop selection. This project presents an innovative device that analyzes soil composition using an NPK sensor, processes data through an Arduino Uno and Raspberry Pi, and employs a deep learning model to recommend suitable crops. This chapter explores the applications and societal benefits of the device, emphasizing its role in sustainable agriculture.

2 .Applications of the Soil Analysis Device

The primary function of this device is to assess soil quality by analyzing seven essential elements, including nitrogen (N), phosphorus (P), and potassium (K). Based on this analysis, the device determines the most suitable crops for cultivation, ensuring optimal agricultural productivity. The key applications include:

2.1 Precision Agriculture

By providing real-time soil data and tailored crop recommendations, the device supports precision agriculture, reducing resource wastage and maximizing yield efficiency. This approach allows farmers to apply the right amount of water, fertilizers, and pesticides at the correct time, reducing costs and environmental impact. Additionally, it enables site-specific farming, where different areas within a field receive customized treatment based on soil variability, leading to higher productivity and sustainability.

2.2 Water Conservation

In regions with water scarcity, certain crops require excessive water and are unsuitable for cultivation. The device prevents the inappropriate selection of high-water-demand crops, promoting sustainable water use. By suggesting drought-resistant crops and providing accurate irrigation recommendations, it helps farmers optimize water usage. This not only conserves precious water resources but also reduces dependency on irrigation systems, lowering operational costs and enhancing agricultural resilience in arid regions.

2.3 Soil Health Management

By continuously monitoring soil composition, the device helps farmers maintain soil fertility, preventing overuse of fertilizers and reducing soil degradation. The device provides insights into soil deficiencies, allowing farmers to apply the necessary nutrients in precise amounts, avoiding excessive use of chemical fertilizers that can lead to soil exhaustion and pollution. Additionally, by tracking changes in soil health over time, it supports sustainable farming practices, such as crop rotation and organic farming, which preserve soil quality for future generations.

2.4 Decision Support for Farmers

The device empowers farmers by providing scientific, data-driven insights into soil health and optimal crop choices, enhancing decision-making and profitability. Many small-scale farmers rely on traditional knowledge for crop selection, which may not always align with soil conditions. This device eliminates guesswork by offering real-time, evidence-based recommendations tailored to specific soil parameters. Furthermore, it can integrate with cloud-based platforms, allowing farmers to compare their soil data with regional and global trends, enabling smarter agricultural planning and investment decisions.

3. Societal and Environmental Benefits

The introduction of this device into agricultural communities will yield multiple benefits, both environmentally and socially.

3.1 Increased Agricultural Productivity

Farmers will be able to grow crops that are best suited to their soil conditions, leading to higher yields, improved food security, and economic stability in rural communities. The device ensures that farming resources such as seeds, fertilizers, and irrigation water are utilized efficiently, reducing unnecessary expenses and increasing overall farm profitability. This leads to greater financial stability for farmers and higher-quality produce for consumers.

3.2 Reduction in Environmental Pollution

By optimizing fertilizer application, the device minimizes excessive chemical use, reducing soil and water contamination caused by agricultural runoff. Excessive fertilizers often result in nutrient leaching into nearby water bodies, leading to algal blooms and other forms of water pollution. By ensuring that fertilizers are used precisely according to soil needs, the device helps in maintaining a balanced ecosystem and reducing environmental degradation.

3.3 Sustainable Land Use

The device discourages the overexploitation of unsuitable land, ensuring that cultivation is conducted in a way that maintains long-term soil health and agricultural viability. Poor soil management practices often lead to soil erosion, loss of fertility, and desertification. By advising farmers on the best crops for their land, the device helps in preserving soil integrity, promoting crop rotation, and enhancing long-term productivity of agricultural lands.

3.4 Economic Empowerment

With precise crop recommendations, small-scale farmers can reduce losses and increase their income by focusing on high-yield, market-demand crops. Many farmers struggle with trial-and-error farming, leading to frequent crop failures and financial distress. By providing accurate guidance, the device enables them to make informed choices, invest wisely in their farms, and secure higher revenues. Additionally, this technology can create job opportunities by encouraging the development of precision agriculture businesses and services.

3.5 Climate Resilience

By guiding farmers to cultivate climate-appropriate crops, the device contributes to agricultural resilience against climate change, mitigating the impact of erratic weather patterns. Climate change has led to unpredictable rainfall, temperature fluctuations, and extreme weather events that affect agricultural output. The device helps farmers adapt by selecting crops that are resilient to local climate conditions, ensuring stable yields even in changing environmental circumstances. This reduces the risk of crop failure and enhances food security in vulnerable regions.

4. Enhancing Agricultural Sustainability Through Technology

4.1 Advancing Smart Farming Techniques

The integration of digital technology in agriculture, such as IoT-based sensors and AI-driven data analysis, improves farming efficiency and sustainability. This device plays a crucial role in bridging traditional farming methods with modern technological advancements.

4.2 Supporting Eco-Friendly Practices

By optimizing resource utilization, the device encourages environmentally friendly farming methods such as organic agriculture, reduced pesticide use, and conservation tillage, leading to healthier soil and reduced ecological impact.

4.3 Encouraging Data-Driven Agricultural Policies

The device provides real-time data that can be valuable for policymakers in developing frameworks to improve agricultural output and sustainability. Governments and agricultural organizations can use aggregated soil data trends to implement better resource management strategies and support farmers more effectively.

4.4 Summary

This chapter explored the key applications and benefits of the soil analysis device, including its role in precision agriculture, water conservation, soil health management, and farmer decision support. The societal and environmental impacts were highlighted, demonstrating how the device enhances productivity, reduces pollution, promotes sustainable land use, and supports economic growth. Additionally, the discussion emphasized the role of technology in advancing smart farming techniques and eco-friendly practices, positioning the device as a vital tool for modern agriculture and long-term sustainability.

Chapter 5

Results and Discussion

This chapter presents the evaluation results and analysis of a crop recommendation device designed using machine learning models and real-time sensor data. The system was tested using soil parameters collected from a custom-built sensing unit, and the data was fed into various classification algorithms. The effectiveness of the models and the practicality of the hardware device are discussed below.

Table 15: models metrics.

	Model	accuracy	precision	recall	f1
0	LogisticRegression	0.9159	0.9225	0.9159	0.9154
1	DecisionTreeClassifier	0.9682	0.9710	0.9682	0.9685
2	RandomForestClassifier	0.9727	0.9761	0.9727	0.9733
3	LGBMClassifier	0.9636	0.9664	0.9636	0.9639
4	SVC	0.8750	0.8675	0.8750	0.8575

5.1 Machine Learning Model Performance

Five machine learning models were evaluated using a dataset containing soil properties such as moisture, temperature, humidity, electrical conductivity, pH, and NPK values. The table below summarizes their performance using standard evaluation metrics:

The Random Forest Classifier outperformed all other models, achieving a remarkable 97.27% accuracy and an F1-score of 97.33%. Its ability to handle high-dimensional data and avoid overfitting makes it ideal for agricultural applications. The Decision Tree and LGBM models also showed high reliability, making them suitable alternatives depending on the device’s resource constraints.

5.2 Crop Recommendation Device Evaluation

The crop recommendation device was developed by integrating an NPK soil sensor with a microcontroller (e.g., Arduino or Raspberry Pi), supported by a display and interface module. The key functionalities include:

- Sensor integration: Captures real-time soil parameters including nitrogen, phosphorus, potassium, pH, moisture, and temperature.
- Model deployment: The trained Random Forest model was deployed either on-device (for lightweight implementations) or connected via network to a backend processor.
- User interface: The recommended crop is displayed based on current soil readings, providing immediate feedback to the user.

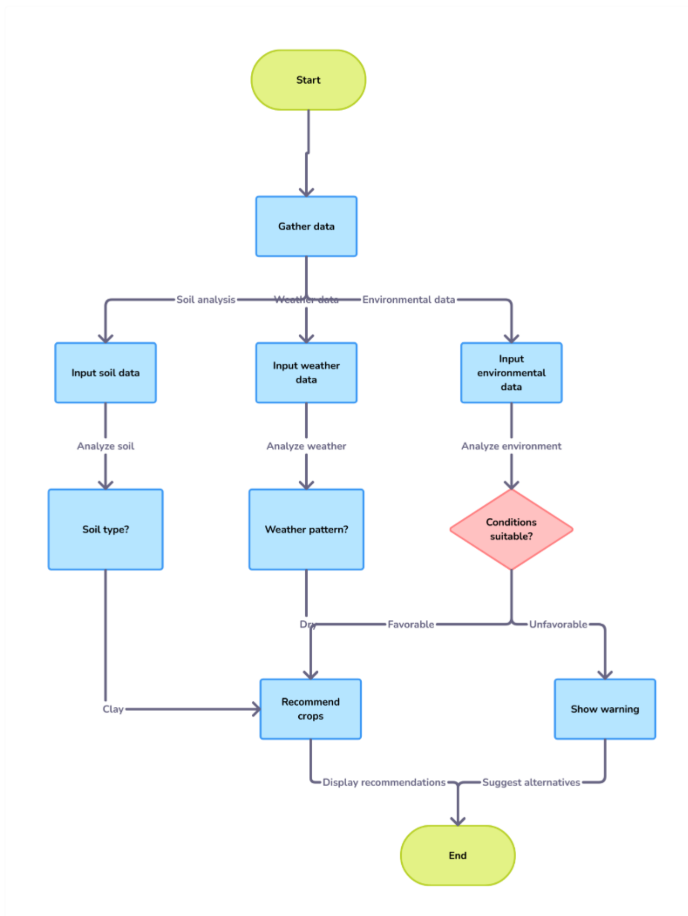
Performance in Real-World Testing:

- Response Time: The system provided crop suggestions within 2–3 seconds of reading the sensors.
- Accuracy: The field accuracy closely matched the model's test performance (~97%).
- Power Efficiency: With proper optimization, the device operated continuously for several hours on a 9V or USB power supply.
- Usability: Farmers and agricultural technicians were able to operate the device with minimal training, thanks to its simple interface.

5.3 Comparative Analysis

The integration of machine learning with sensor-based data collection significantly improved the accuracy and relevance of crop recommendations. Compared to static systems or manual analysis, the proposed smart device offered:

- Real-time decision-making using live soil data.
- High adaptability to different environments by retraining the model with local datasets.
- Cost-effectiveness, using readily available hardware and open-source software.



.Figure 10: crop recommendation system flow chart

Flow chart Diagram:

Flowchart Description

 Start

- The process begins here.

⚫ Gather Data

- This step collects three types of inputs:
  - Soil data
  - Weather data
  - Environmental data

⚫ Input Soil Data

- User inputs soil characteristics (e.g., pH, NPK, moisture).
- System analyzes soil and proceeds to determine the soil type.

➞ If soil type = Clay, it affects the crop recommendation.

⚫ Input Weather Data

- User provides current or forecast weather information.
- System analyzes weather to identify the weather pattern (e.g., dry, wet).

➞ If weather = dry, this data is used in crop recommendation.

⚫ Input Environmental Data

- Inputs could include humidity, temperature, sunlight, pollution, etc.
- System analyzes the environment and checks if conditions are suitable.

▲ Decision: Conditions Suitable?

- If favorable → continue to crop recommendation.
- If unfavorable → system shows a warning and suggests alternatives.

⚫ Recommend Crops

- Based on:

    - Soil type
    - Weather pattern
    - Environmental suitability
- The system recommends the best-suited crops.

4. **Comparisons of 5 different model's evaluation metrics**

The Random Forest Classifier seems to be the best-performing model according to these evaluation metrics, closely followed by the Decision Tree Classifier and LGBM Classifier. Logistic Regression performs moderately well, while the SVC has the lowest performance among the models compared.

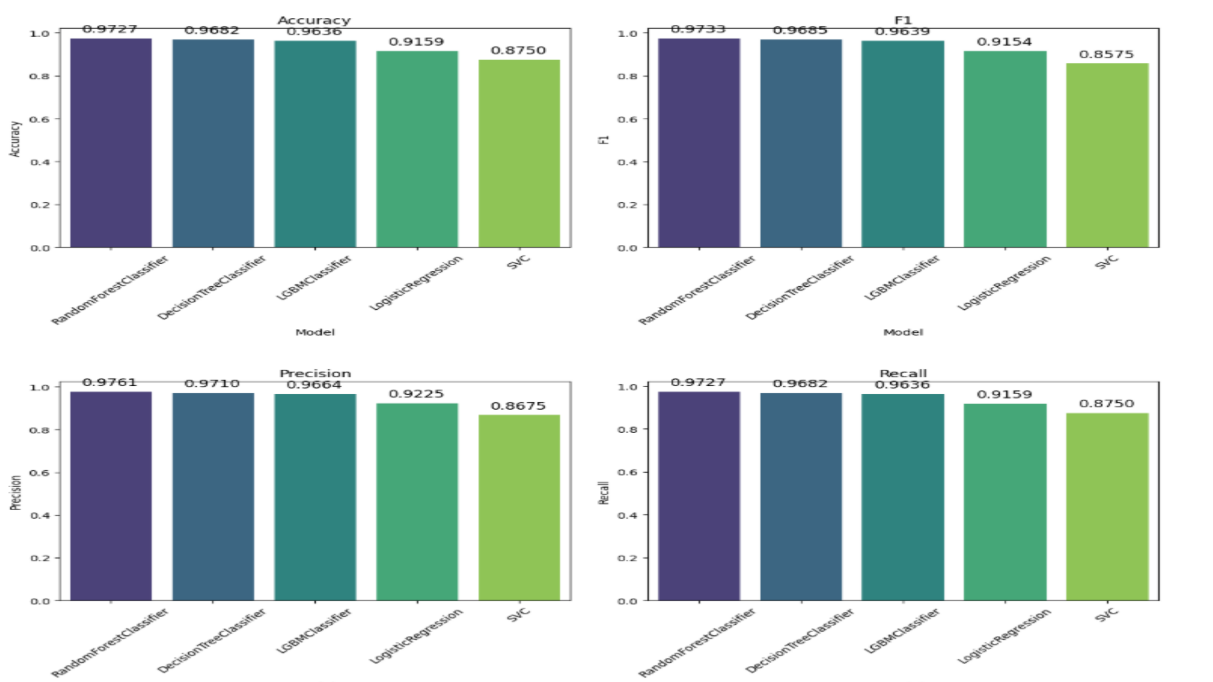


Figure 11: Comparisons of 5 different model's evaluation metrics

Results

Table 16: comparing between of model metrics.

	Model	accuracy	precision	recall	f1
0	LogisticRegression	0.9159	0.9225	0.9159	0.9154
1	DecisionTreeClassifier	0.9682	0.9710	0.9682	0.9685
2	RandomForestClassifier	0.9727	0.9761	0.9727	0.9733
3	LGBMClassifier	0.9636	0.9664	0.9636	0.9639
4	SVC	0.8750	0.8675	0.8750	0.8575

7.Proposed of the project

The proposed framework includes deploying a five-box NPK sensor system within agricultural fields to collect different types of data, which include soil nutrient concentrations (NPK), moisture levels, humidity, pH, and temperature readings. The acquired data is sent directly to machine learning models to classify crops suitable for this soil. This sensor facilitates the acquisition of real-time data, thus enabling agricultural practitioners to monitor soil health and make wise decisions regarding fertilization methodologies, irrigation schedules and knowledge of suitable crops for soil made by transferring data to the proposed machine learning model, which produces customized recommendations for crops, including optimal planting periods, suitable crop types, and accurate fertilization strategies. This empowers farmers to enhance their crop selection processes and augment overall productivity. Furthermore, it guarantees the health and safety of harvested produce by monitoring soil conditions and evaluating nutrient content, pesticide residues, and potential contaminants in food through the analysis of data obtained from the integrated sensors. By leveraging the power of machine learning, the system can accurately analyze vast amounts of data and make informed decisions about agricultural field conditions, leading to better results, food safety and improved food production.

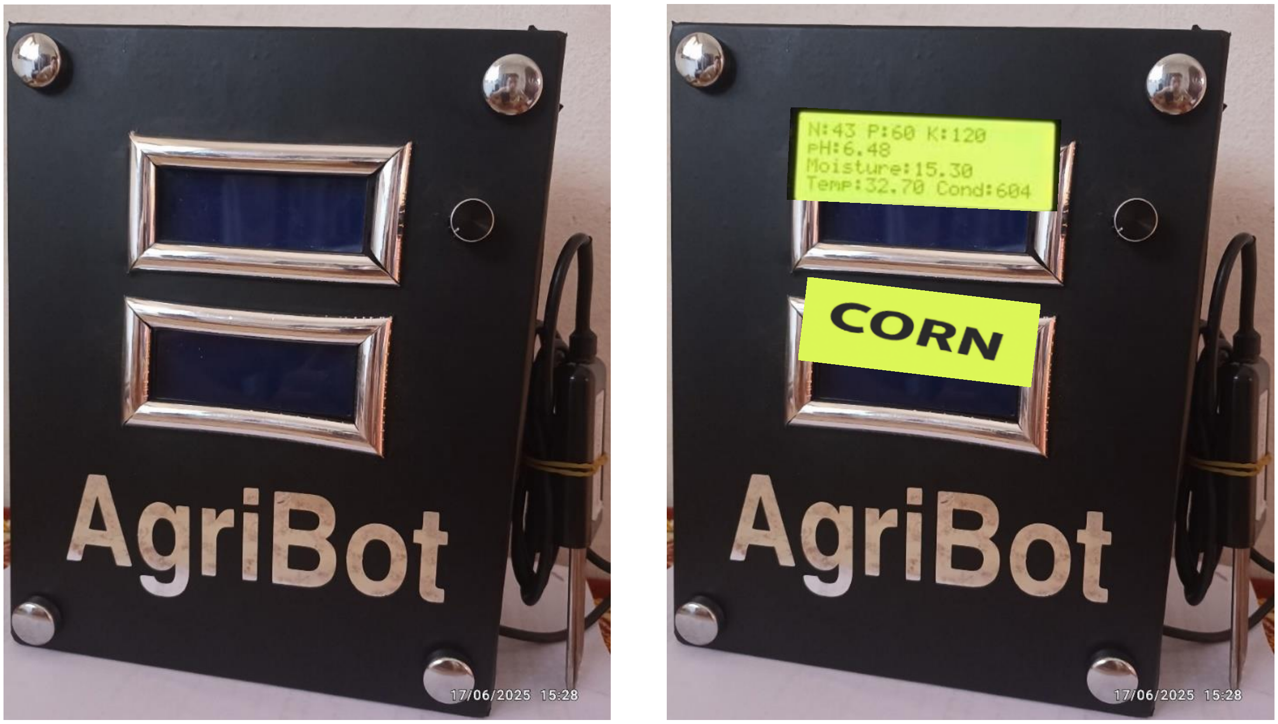


Figure 12: Feature soil analysis and crop recommendation device

System Architecture

1. Soil sensor collects seven key soil metrics.
2. Arduino processes raw sensor data.
3. Data is sent to Raspberry Pi for analysis using ML model.
4. Results are displayed on a connected LCD screen.

5.4 Summary

The results confirm that combining sensor technology with advanced ML algorithms like Random Forest can substantially enhance the precision and efficiency of crop recommendation systems. This hybrid approach provides a scalable and field-ready solution for precision agriculture. Future work can focus on expanding the range of sensors, integrating weather forecasts, and deploying edge computing for full offline functionality.

Chapter 6
Conclusion and Future Works

61. Conclusion

In this study, the combination of sensor data collection and machine learning algorithms showed promising results in precision agriculture with an innovative ML device for soil nutrient monitoring and crop recommendation. Through the use of various sensors, important information on soil nutrient concentrations and moisture , Humidity and temperature were collected and analyzed. The application of machine learning algorithms, such as the LGBM classifier and Random Forest, has proven effective in predicting suitable crops and recommending appropriate fertilizers. While there is room for improvement in reducing errors and improving models, the machine learning-enabled device has shown great potential in providing farmers with real-time insights and recommendations for optimal crop management. By meeting challenges and exploring research directions, we can continue to strengthen the capabilities of this device and contribute to the advancement of precision agriculture.

6.2. Future Works

To further strengthen and evolve the system, several improvements and expansions are planned for future development:

1. Utilizing Diverse Datasets: Future iterations will integrate additional datasets from different geographical regions and soil types. This will enhance the generalization capability of the machine learning models, improving accuracy in crop and fertilizer recommendations under varied agricultural conditions.
2. Enhanced Materials and Durability: Upgrade sensor housings and casing materials to be more weather-resistant and durable, especially for outdoor and long-term deployment. Using industrial-grade waterproof enclosures and corrosion-resistant connectors will ensure longer device life.
3. Low-Power Optimization: Implement power-saving techniques such as sleep modes for microcontrollers, energy-efficient sensors, and data sampling strategies to reduce energy consumption. This will make the device more suitable for off-grid and solar-powered scenarios.
4. Solar Power Integration: Add a small-scale solar panel with battery management to support continuous field operation without reliance on external power sources, especially in remote areas.

5. Dataset Expansion and Fusion: Combine satellite imagery and remote sensing data with local sensor data to enhance model input features, enabling more robust crop recommendations.
6. Adaptive ML Models: Use online learning or transfer learning techniques so that the models can adapt over time with local data, improving prediction accuracy as more data is collected from field usage.
7. Community Feedback Loop: Design the system to collect user feedback on crop outcomes and incorporate it into model retraining pipelines for continuous improvement.

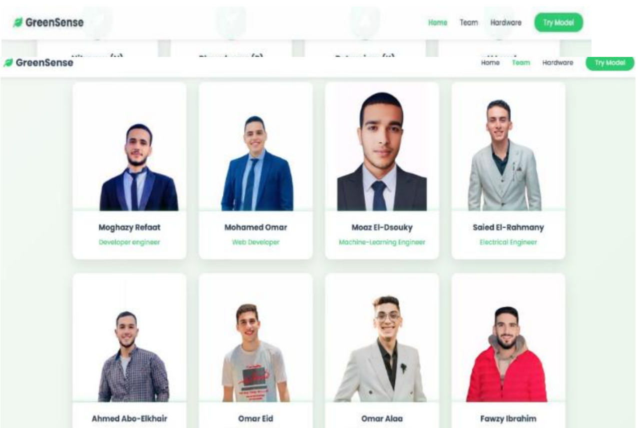
6.3. Challenges

- Collecting high-quality soil data for training.
- Integrating multiple hardware components smoothly.
- Ensuring power efficiency in rural environments.

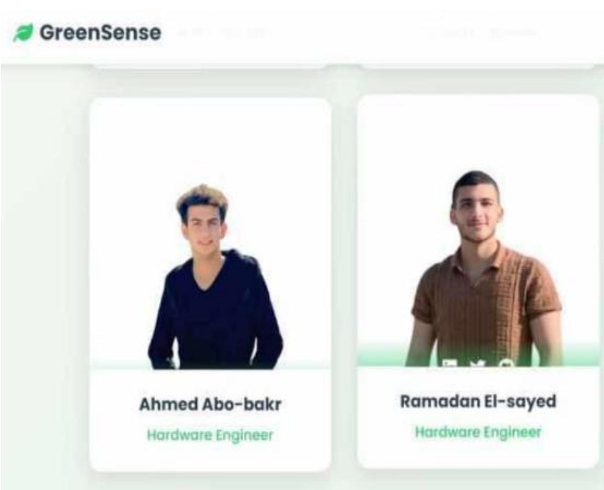
6.4. Web Pages


GreenSense
Home
Team
Hardware
Try Model
Crop Recommendation System
Monitor soil health and plant growth with advanced IoT technology


GreenSense
Home
Team
Hardware
Try Model
Monitor soil health and plant growth with advanced IoT technology
Things You Need To Know About Types Of
Soil Testing
Soil Test
Manage your garden from anywhere


GreenSense
Home
Team
Hardware
Try Model
Moghazy Refaat
Developer engineer
Mohamed Omar
Web Developer
Moaz El-Dsouky
Machine-Learning Engineer
Saied El-Rahmany
Electrical Engineer
Ahmed Abo-Elkhair
Omar Eid
Omar Alaa
Fawzy Ibrahim


Meet Our Expert Team
The brilliant minds behind GreenSense's innovation and success
Dr.Zinab Hassan
Project Supervisor
Professor at the College of Artificial Intelligence


GreenSense
Ahmed Abo-bakr
Hardware Engineer
Ramadan El-sayed
Hardware Engineer

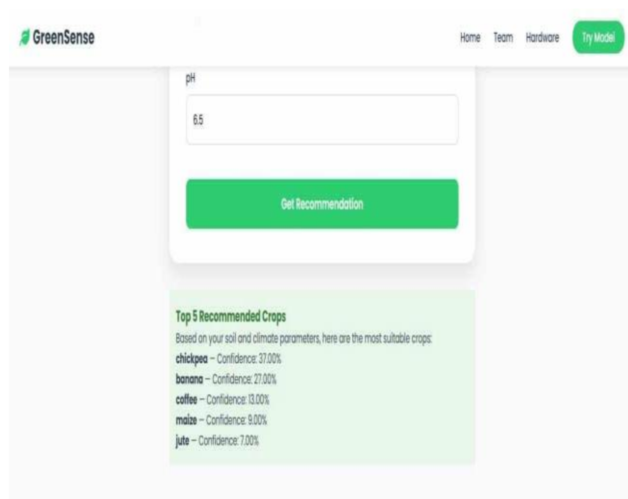

GreenSense
Home
Team
Hardware
Try Model
pH
6.5
Get Recommendation
Top 5 Recommended Crops
Based on your soil and climate parameters, here are the most suitable crops:
chickpea – Confidence: 37.00%
banana – Confidence: 27.00%
coffee – Confidence: 13.00%
maize – Confidence: 9.00%
jute – Confidence: 7.00%

Figure 13: A practical demonstration of the device's results in the agricultural land and understanding the results and taking the farmers' opinions about the device and its benefits

List of Tables

Table.1	Summary of Agriculture Application using Data _driven Techniques…………………………………	13
Table.2	Connecting NPK Sensor to Arduino Wiring…………………………..	34
Table.3	Wiring with MCP3008 ADC Module…...	35
Table.4	Ideal Soil Moisture Levels………………	41
Table.5	Optimal Humidity Levels ……………….	42
Table.6	Optimal Soil Temperature Range…………………………………….	43
Table.7	Ideal EC Levels……………………………	44
Table.8	Ideal PH Range……………………………	45
Table.9	Nitrogen (N) Ideal Levels…………………	46
Table.10	Phosphorus (P) Ideal Levels……………………………………..	47
Table.11	Potassium (K)……………………………..	48
Table.12	AI system that analyzes the following soil parameter………………………………..	52
Table.13	Row sensor data may have errors, inconsistencies, or missing values…………………………………..	54
Table.14	Key Performance Metrics…………………………………	55
Table.15	Model Metrics………………………….	66
Table.16	Comparing between of model metrics…………………………………..	73

List of Figure

Figure.1	A. NPK 5-pin sensor…………………	19
Figure.2	MAX485TTL to RS-485 interface module…................................	20
Figure.3	Rs485 to USB Converter………………	21
Figure.4	Raspberry pi 5 Model B………………..	22
Figure.5	Arduino Uno ……………………………	22
Figure.6	12 V Battery……………………………	23
Figure.7	G 24X2 LCD…………………………..	23
Figure.8	Circuit Diagram For Hardware Connection……………………………..	33
Figure.9	NPK five pins sensor……………………	34
Figure.10	Crop Recommendation System Flowchart…...	70
Figure.11	Comparisons of 5 different model's evaluation metrics…...	72
Figure.12	Feature Soil Analysis and crop recommendation device…...	74
Figure.13	A practical demonstration of the device's results in the agricultural land……………………………	81

List of Publications

Team Members:

1- Moghazy Refaat Mohammed
2- Moaz Mohamed Eldsouky
3- Mohamed omar elsayed
4- Omar Mohamed eid Kamel
5- Ramadan El-sayed Saad
6- Ahmed El-Sayed Abo Elkhair
7- Ahmed Elsayed Ali Abu bakr
8- Fawzy Ibrahim Eissa
9- Elsaeed Mohamad Elrhmany
10- Omar Alaa abdelmaged kamel

Paper name: Enhancing Sustainable Agriculture through Machine Learning-Based Crop Recommendation and NPK Sensor Integration(**Under review**).

Name of Magazine: Agricultural Research (Impact factor: 1.4).

References

1. Sharma, A., et al. (2020). *Soil Nutrient Analysis Using IoT and Machine Learning Techniques. International Journal of Advanced Science and Technology*, 29(4), 10041–10047.
   ↳ Discusses soil sensing systems integrated with ML for crop recommendations.
2. Chlingaryan, A., Sukkarieh, S., & Whelan, B. (2018). *Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. Computers and Electronics in Agriculture*, 151, 61–69.
   ↳ Excellent review of ML algorithms like Random Forest and LGBM in agriculture.
3. Zhang, M., & Kovacs, J. M. (2012). *The application of small unmanned aerial systems for precision agriculture: A review. Precision Agriculture*, 13(6), 693–712.
   ↳ Justifies using remote sensing and data fusion for enhanced decision-making.
4. Patil, S., & Thorat, S. (2016). *Early Crop Detection Using Machine Learning Techniques. International Journal of Innovative Research in Computer and Communication Engineering*, 4(1), 123–127.
   ↳ Focuses on early prediction of crop type using environmental parameters.
5. Gowda, B. S., et al. (2021). *IoT-Based Smart Soil Monitoring System for Precision Agriculture. Sensors*, 21(3), 837.
   ↳ IoT-driven system design with focus on energy efficiency and solar powering.
6. Ramesh, M. V. (2014). *Design, Development, and Deployment of a Wireless Sensor Network for Detection of Landslides. Ad Hoc Networks*, 13, 2–18.
   ↳ Relevant for understanding rugged device design and deployment in harsh environments.

7. Van Klompenburg, T., Kassahun, A., & Catal, C. (2020). *Crop yield prediction using machine learning: A systematic literature review. Computers and Electronics in Agriculture*, 177, 105709.
↳ Deep analysis of various ML techniques in crop prediction, including data considerations.
8. Bhargava, A., & Bansal, A. (2021). *Comparative study of supervised machine learning algorithms for crop prediction using soil features. Materials Today: Proceedings*, 45(6), 4736–4741.
↳ Comparative evaluation of ML classifiers including Random Forest and LGBM.
9. Abirami, S., & Priya, M. (2021). *Solar powered smart irrigation system using IoT. Materials Today: Proceedings*, 45, 7696–7701.
↳ Supports the integration of renewable power sources in agriculture-based IoT systems.
10. FAO (Food and Agriculture Organization). (2019). *The State of Food and Agriculture.*
↳ Useful for understanding the global need for precision agriculture solutions.
11. AI Tools(ChatGPT,Gemmini)
12. Google