

لجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي
جامعة وهران 1 أحمد بن بلة
معهد العلوم والتقنيات التطبيقية

Ministry of Higher Education and Scientific Research
University of Oran 1 Ahmed Ben Bella
Institute of Applied Sciences and Techniques ISTA

Project submitted to the Institute of Applied Science and Technology With a
view to obtaining the diploma of **MASTER in Instrumentation and
Metrology**

Topic

***An IoT Low-Cost Fertilizer for Crop
Recommendation using Machine Learning***

Presented by:

Mrs. Sekfal Zahret el Afrah

Mr. Djemai Imad Eddine Ismail

Defended on 27 June 2024 before the jury composed of:

Mrs.TAHRI Ghrissi

MCB

Chair

Mr. DAHANE Amine

MCA

Supervisor

Mrs.HAMDANI Nesrine

MCA

Examiner

University year: 2023/2024

Abstract

Agriculture plays a crucial role in sustaining the world's growing population, yet farmers face challenges in optimizing crop production and resource management. This thesis introduces an innovative IoT device empowered by machine learning (ML) to monitor soil nutrients and offer precise crop recommendations. The device integrates sensors such as, an NPK sensor and MAX485 TTL to gather real-time data on soil composition, humidity, temperature, rainfall, pH, and nutrient levels. This data is then transmitted to a server via protocol. ML algorithms analyse the collected data to generate tailored recommendations, including optimal crop choices, suitable fertilizer types, and application rates based on crop needs and soil conditions. Field experiments validate the system's efficacy compared to traditional methods, demonstrating its ability to boost crop productivity, optimize resource allocation, and promote sustainable agricultural practices for enhanced food security.

Keywords

-  Machine Learning
-  Internet of Things (IoT)
-  Agriculture
-  Crop Recommendation
-  Soil Nutrients
-  Sensor

Acknowledgements

We want to thank Almighty God for granting us health and the will to begin and finish this dissertation. We also want to thank all the individuals who contributed to the success of our internship and assisted us during the writing of this dissertation.

First and foremost, we would like to thank our thesis advisor, A. DAHANE, for his patience, availability, and especially his insightful advice, which enriched our reflections. He was very supportive throughout the realization of this dissertation, providing inspiration, assistance, and the time he graciously dedicated to me. Without him, this dissertation would never have come to fruition.

We also express our gratitude to the entire teaching staff of Oran 1 University and the professional speakers responsible for our training, for providing the theoretical part of it.

We dedicate this humble work to our parents, who have always supported and encouraged us.

Thank you all.

❖ List of figures

Figure 1 . IoT application in agriculture	14
Figure 5. JXCT Soil NPK sensor	37
Figure 6 . the assembly carried out.....	38
Figure 7. The MAX485 TTL	40
Figure 8. Interfacing NPK Soil Sensor with Arduino	42
Figure 9. The assembly carried out	42
Figure 10. Soil NPK Tester.....	43
Figure 11 . Element of NPK Sensor	44
Figure 12. Display of Digital Soil Instant NPK Nutrient Tester	45
Figure 13 . Arduino card architecture.....	47
Figure 14 . Arduino ide.....	48
Figure 15. jupyter	49
Figure 16 . Soil nutriments monitoring.....	50
Figure 17 . Methodology for the current study	51
Figure 18 . Correlation Metrics on the crop recommendation dataset	52
Figure 19 . Training machine learning model.....	53
Figure 20. Diagram of gradient boosting machine learning method.....	54
Figure 21. Diagram of Decision Tree	56
Figure 22. Random Forest classifier.....	57
Figure 23. The sigmoid function	59
Figure 24. Visualizing the Knn algorithm behaviour	60
Figure 25. The normal or the Gaussian distribution function	61
Figure 26. Comparison of ML Techniques under study.....	64
Figure 27. Sampling procedure	64
Figure 28. Taking samples	75
Figure 29. [1 : Sensor N (Barly Soil), 2 : Sensor P (soil with no crops), 3 : Sensor K (Olive soil) , 4 : (Olive soil with soil sensor), 5 : Soil sensor]	76

❖ List of Tables

Table 2. IoT software used in agriculture.	17
Table 3. IoT hardware used in agriculture.	18
Table 4 . Different security concerns and attacks on smart agriculture.	18
Table 5 . Summary of the most recent studies.	23
Table 6. the specifications for Soil NPK Sensor	39
Table 7 . The specifications for Digital Soil Fertility Available Instant NPK Nutrient Tester Meter.....	44
Table 8. First 10 records of the Dataset	51
Table 9 . Statistical summary table	52
Table 10. Classification report table	58

❖ Table of contents

Abstract	2
Acknowledgements	3
General Introduction	11
1.1 Introduction	13
2.1 Internet of things in precision agriculture	14
2.1.1 Definition.....	14
3.1 Application of IoT in Agriculture.....	14
3.1.1 Irrigation Management System.....	15
3.1.2 Pest and Disease Control	15
3.1.3 Animal Movement Monitoring	15
3.1.4 Water Quality Monitoring	15
3.1.5 Greenhouse Condition Monitoring	15
3.1.6 Precision Farming Using UAV	15
4.1 Advantages of IoT in Agriculture.....	16
4.1.1 Increased Efficiency	16
4.1.2 Agility	16
4.1.3 Improved Product Quality	16
5.1 Smart agriculture applications, software and hardware	17
Conclusion	20
1.1 Introduction	22
1.1.1 Research goals	22
1.1.2 Data used	22
1.1.3 Sensors.....	22
1.1.4 Project cost.....	22
1.1.5 Machine learning use:	22
1.1.6 Targeted anomalies and limitations:	22
2. Table 1	23
3.1 Related Work.....	28
Article 1: A Self-Powered, Real-Time, Lora WAN IoT-based Soil Health Monitoring System.....	28
Article 2: Machine Learning Algorithm for Soil Analysis and Classification of Micronutrients in IoT-Enabled Automated Farms	29
Article 3: IoT Framework for Measurement and Precision Agriculture Predicting the Crop Using Machine Learning Algorithms	29

Article 4: IoT-Enabled Soil Nutrient Analysis and Crop Recommendation Model for Precision Agriculture	30
Article 5: Crop Yield Maximization Using an IoT-Based Smart Decision	30
Article 6: A Cloud Enabled Crop Recommendation Platform for Machine Learning-Driven Precision Farming	31
Article 7: Crop Recommendation System to Maximize Crop Yield using Machine Learning Technique	31
Article 8: Internet of Things (IoT) assisted Context Aware Fertilizer recommendation	32
Article 9: Crop Recommendation System using Machine Learning	32
Article 10: An innovative IoT based system for precision farming.....	32
Conclusion	33
1.1 Introduction	35
1. Nitrogen (N)	36
2. Phosphorus (P).....	36
3. Potassium (K)	36
Part one: Materials and software used	37
2 .1 what is an NPK sensor and how it works.....	37
2.1.1 Required Materials.....	37
2.1.2 Soil NPK sensor Pinout.....	38
3.1 Soil NPK sensor Specs.....	39
4.1 NPK sensor Application	40
4.1.1. Environmental Monitoring:	40
4.1.2. Horticulture	40
4.1.3. Soil Testing Laboratories.....	40
4.1.4. Education	40
5.1 MAX485 TTL to RS-485 Module Converter	40
5.1.1 Specs MAX485 TTL to RS-485 Module	41
5.1.2 Pinout MAX485 TTL to RS-485 Module	41
Data side	41
Output side	41
6.1 Interfacing NPK Soil Sensor with Arduino	42
7.1 Digital Instant Soil NPK Tester.....	43
7.1.1. Operation	43
7.1.2. Ease of Use	43
7.1.3. Portability	43
7.1.4. Accuracy.....	43

7.1.5. Applications	43
8.1 Product Description	44
9.1 The use of Digital Soil Instant NPK Nutrient Tester.....	45
10.1 NPK sensor Application.....	46
10.1.1. Field Agriculture	46
10.1.2. Soil Quality Monitoring.....	46
10.1.3. Analysis of Farmland	46
10.1.4. Field Research.....	46
10.1.5. Agricultural Education	46
11.1 Informatics cards	47
11.2 Arduino IDE	48
11.3. Jupyter	49
Part two: Implementation and results	51
12.1 Methodology	51
1. Data Collection	51
2. Exploratory Data Analysis (EDA)	52
3. Splitting the Data into Training and Testing Sets.....	52
4. Training the Models	53
5. .Model Evaluations and Selection.....	53
The Gradient Boosting Classifier	53
•Accuracy.....	54
•Precision	54
• Recall.....	54
• F1-score	55
Decision Tree	55
Random Forest	57
Logistic Regression	59
K Nearest Neighbours	60
Gaussian Naïve Bayes	61
Support Vector Machine	62
Result and discussion.....	63
6. Predictions and Evaluation.....	64
I – IDENTIFICATION	65
II – DELIMITATION.....	65
III – SUPERFICIE TOTALE.....	65

IV -Service	65
V - TECHNICAL SUPERVISION	65
VI - LABOR FORCE	65
VII - INTERVENTION ZONE.....	66
VIII - VOCATION	66
IX - PROGRAM.....	66
X - OBSERVATIONS.....	66
1.2 Main Missions of the Structure	66
1.3 Current Production Level.....	68
1.4 Description of the Studied Trial Plot	68
1.5 Plant Material	68
Nutrients necessary for plant growth.....	71
2.1. Secondary Elements.....	71
2.1.1 Calcium (Ca)	71
2.1.2 Magnesium (Mg).....	71
2.1.3 Sulfur (S)	71
2.2 Macronutrients: N - P - K	71
2.2.1 Nitrogen (N).....	71
2.2.2 Phosphorus (P)	72
2.2.3 Potassium (K).....	72
3- Signs of deficiencies in basic elements: N.P.K.....	72
3.1 Soil.....	73
1. Physical	73
2. Chemical.....	73
3. Biological	73
4.1 Soil Analysis	74
Characterization analysis.....	74
Monitoring analysis	74
Diagnostic analysis	74
4.2 Soil Analysis Steps	74
Pre-laboratory	74
Post-laboratory.....	74
5. Interest of Irrigation for Olive Trees	75
5.1 Rapid Entry into Production for Young Plantations	75
5.2. Increase in Production in Mature Orchards	75

Impact of Irrigation on Growth and Alternation	75
6. Sampling procedure	75
7. The economic study of our project	79
7.1. Cost Analysis of our Project.....	79
7.2. Comparison with Digital Soil NPK Tester	80
Cost Difference	80
7.3. Economic Efficiency.....	80
7.3.1 Cost Efficiency	80
7.3.2 Performance.....	80
7.3.3 Cost-Benefit Analysis	80
- Our Project.....	80
Significantly lower cost.	80
Commercial Tester	80
High initial investment.	80
Potential for additional features or higher precision.	80
8. Additional Considerations.....	81
Scalability	81
Maintenance and Support	81
Accuracy and Precision	81
Results.....	81
Conclusion	82
Perspectives	82
General conclusion	84
References	85

General Introduction

In recent years, the convergence of Internet of Things (IoT) technology and machine learning has opened up new avenues for innovation across various industries, with agriculture being no exception. This synergy has led to the development of intelligent systems capable of revolutionizing traditional agricultural practices by offering real-time monitoring, predictive analytics, and data-driven decision-making. One such ground-breaking application is the integration of IoT and machine learning in the design and implementation of systems for soil nutrients monitoring and crop recommendation.

The overarching goal of these systems is to address the challenges faced by farmers in managing soil health and optimizing crop yields. Traditional methods of soil analysis and crop management often lack precision and require significant time and resources. Additionally, factors such as changing environmental conditions, soil variability, and crop diversity further complicate decision-making processes. However, by harnessing the power of IoT sensors to gather continuous data on soil parameters and leveraging machine learning algorithms to analyze this data, these systems offer unprecedented insights and recommendations tailored to specific agricultural contexts. This project is broken down into four key chapters, each playing a critical role in the creation and deployment of our system.

In the first chapter, we will lay the foundations of our project by exploring the Internet of Things (IoT). We will define the fundamental principles of the IoT and its impact on our lives. In addition, we will discuss the preliminary work that paved the way for our project, including the goals and opportunities that were presented.

The second chapter will focus on a comparative study. The goal of this study is to thoroughly review ten current studies that investigate the connection between IoT and ML systems and agriculture. The main goal is to examine these papers in order to determine their goals, approaches, data, and technologies and tools utilized in order to put their answers into practice.

The third chapter will focus on the key technical components of our project. We will detail the sensors, the specific software, and computer maps we have chosen to operate our monitoring system and for data processing. This chapter will lay the essential hardware and software foundation for the implementation of our solution. and we will enter the realization phase of our project. We will reveal details of the implementation of our monitoring system, including the machine learning algorithms used, data processing methods, and the concrete results of our work. This chapter will demonstrate how our innovative solution materializes and its potential to revolutionize the way we manage .

This general introduction sets the stage for exploring the innovative approach of integrating IoT technology and machine learning in agricultural practices, with a particular focus on soil nutrients monitoring and crop recommendation. It outlines the significance of this integration in addressing key challenges faced by farmers and emphasizes the potential for transformative impact on agricultural productivity, sustainability, and environmental stewardship. Throughout this study, we delve into the design and realization of such systems, examining their components, functionalities, and implications for the future of agriculture

Chapter 1

IOT in Precision Agriculture

1.1 Introduction

The total world population has experienced significant growth, rising from 1.65 billion in 1990 to 7.4 billion in 2020, as indicated by a statistical analysis. Projections suggest that this demographic growth could reach 11.2 billion by the end of the 21st century if the current rate persists. Meanwhile, arable land per person has decreased, going from 0.5 hectares in 1967 to 0.2 hectares in 2020. Faced with this rapid global population growth, it becomes imperative to reconsider traditional agricultural approaches and policies to improve productivity and meet the fundamental needs of the world population.

Climate change also poses a significant threat, particularly to the agricultural sector, due to water scarcity, a crucial resource for crop irrigation. Inadequate management could have negative repercussions on crop availability, food supply, and the global economy.

Agriculture serves as the cornerstone of human life by providing essential food, grains, and other raw materials. Currently, it utilizes 85% of the world's freshwater resources, a proportion that is expected to persist due to population growth and increasing food demand. Technology plays a significant role in enhancing agricultural production while reducing dependence on labour. The emergence of the Internet of Things (IoT) offers a promising solution for smart agriculture; however, realizing this vision remains a challenge as long as connectivity does not effectively reach rural areas. (1)

Kevin Ashton first proposed the idea of the Internet of Things (IoT) in 1999. The goal of IoT is to link any object, anywhere, at any time. The rise of cloud computing, the proliferation of small and mobile devices, and the developments in data analysis have made it especially pertinent in recent times.

Another area where technology is changing conventional farming methods is precision agriculture. In order to maximize the use of resources like water, fertilizer, and pesticides, GPS, IoT, and data analysis are used. Farmers can make targeted treatments by using sensor-equipped drones and satellites to gather comprehensive crop health information. By doing the, farming operations have a less negative impact on the environment and increase crop yields and overall farm profitability.

Moreover, the analysis of large datasets produced by sensors and other Internet of Things (IoT) devices is greatly aided by artificial intelligence (AI) and machine learning (ML). Algorithms using AI and ML may recognize patterns, forecast crop production, find pests or illnesses, and suggest the best ways to farm. This data-driven methodology improves decision-making and helps farmers adjust to shifting circumstances.

This chapter outlines the possible applications of IoT and wireless sensors in agriculture, along with the anticipated difficulties in integrating these technologies with conventional farming methods. A thorough analysis is conducted on IoT devices and the communication methods connected to wireless sensors utilized in agricultural applications.

2.1 Internet of things in precision agriculture

2.1.1 Definition

Internet of things common definition is defining as: Internet of things (IOT) is a network of physical objects. The internet is not only a network of computers, but it has evolved into a network of device of all type and sizes , vehicles, smart phones, home appliances, toys, cameras, medical instruments and industrial systems, animals, people, buildings, all connected ,all communicating & sharing information based on stipulated protocols in order to achieve smart reorganizations, positioning, tracing, safe & control amp; even personal real time online monitoring , online upgrade, process control & administration. Internet of Things (IoT) is a concept and a paradigm that considers pervasive presence in the environment of a variety of things/objects that through wireless and wired connections and unique addressing schemes are able to interact with each other and cooperate with other things/objects to create new applications/services and reach common goals. In this context, the research and development challenges to create a smart world are enormous. A world where the real, digital and the virtual are converging to create smart environments that make energy, transport, cities and many other areas more intelligent. (2)

3.1 Application of IoT in Agriculture

IoT finds a variety of applications in the field of agriculture, they can be broadly characterized into six categories, which are briefly discussed in the following section and presented in Figure1 below (3)

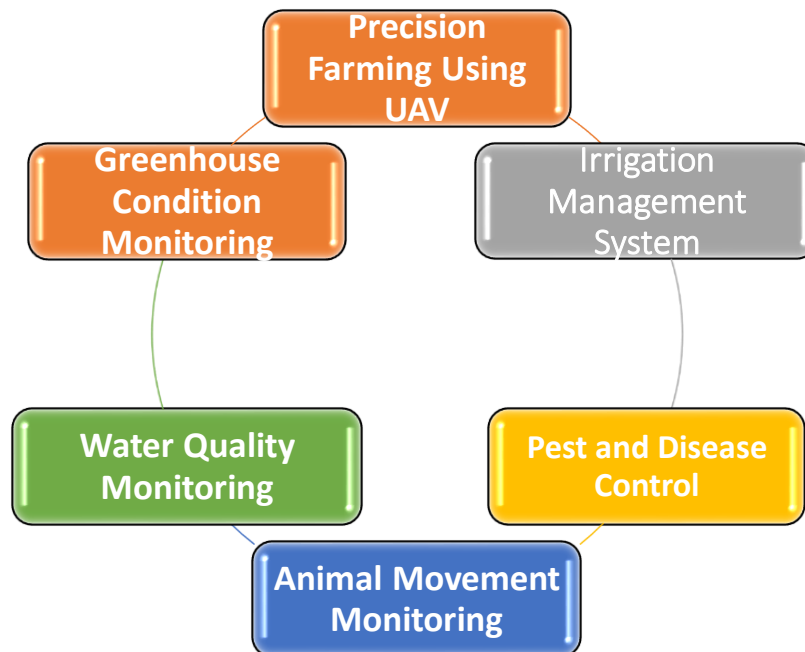


Figure 1 . IoT application in agriculture

3.1.1 Irrigation Management System

Sustainable use of resources is an important challenge nowadays. In the field of farming, water plays an important role in the production of crops. An IoT-based smart irrigation system is being developed which helps in improving the irrigation management system. These system help in optimizing the use of water in the fields. Smart irrigation systems mainly focus on real-time monitoring, controlling, and weather forecasting.

3.1.2 Pest and Disease Control

Pesticides and fertilizers used in a controlled way help to increase crop quality and quantity. They also help in minimizing the farming costs. To achieve this, constant monitoring of the crops and disease identification is required, which is a tough job in the traditionally farming methods. IoT can provide a solution to this problem. By proper monitoring,

3.1.3 Animal Movement Monitoring

Many research articles show that the movement of animals in the field is constantly destroying agriculture fields. To avoid this traditionally puppets were used in the agriculture fields to mimic the presence of humans in the fields. But, this method does not guarantee success. IoT-based smart monitoring can be an efficient way to solve this problem. Using IoT the movement of animals into the field can be monitored and controlled.

3.1.4 Water Quality Monitoring

Water quality is another issue that can affect crop health and reduce their harvest. The factors which affect the quality of water are temperature, pH, conductivity, dissolved oxygen, etc. IoT-based water quality monitoring system is slowly coming up as a solution to tackle this problem. These systems can be used to remotely monitor the physical and chemical parameters of the water and control them if necessary.

3.1.5 Greenhouse Condition Monitoring

Other than water and pests, crop health is also related to temperature, humidity, pH, etc. Greenhouse gases have an adverse effect on the climatic condition and thus, directly impact the crops' health. Therefore, it is important that greenhouse gases are constantly monitored. Different works have presented ideas to remotely monitor greenhouse gases.

3.1.6 Precision Farming Using UAV

In this modern area, unmanned aerial vehicles (UAV) and drones are also finding a place in farming. UAVs and drones can help in precision farming, which contributes, to enhancing the production of crops. UAVs and drones provide support for advanced monitoring of weather conditions, fields, real-time data analysis, etc.

4.1 Advantages of IoT in Agriculture

The benefits of IoT in agriculture are summarized in the following section:

4.1.1 Increased Efficiency: Farmers can monitor their crops and conditions in real-time through IoT-enabled agriculture. They can quickly obtain information, anticipate challenges before they arise, and make informed decisions on how to prevent them. IoT solutions for agriculture also offer automation, such as demand-based irrigation, fertilization, and robotic harvesting.

4.1.2 Agility: Farmers can respond immediately to any significant changes in weather conditions, humidity, air quality, or the health of each crop or soil in the field through real-time monitoring and forecasting systems.

4.1.3 Improved Product Quality: Farmers can better understand the complex relationships between the environment and the quality of their crops using soil and crop sensors drone monitoring, and farm mapping. They can replicate optimal conditions and enhance the nutritional content of products using connected systems.

4.1.4 Cleaner Processes: Precision agriculture systems based on IoT make farming more environmentally friendly by significantly reducing the consumption of pesticides and fertilizers. This strategy generates a cleaner and more organic product compared to traditional farming practices.

4.1.5 Resource Optimization: IoT enables more efficient use of agricultural resources such as water, fertilizers, and pesticides by providing real-time data on specific crop needs.

5.1 Smart agriculture applications, software and hardware

Table 1. IoT software used in agriculture.

Software	Description
Android things	It is a Google-developed embedded operating system for Internet of Things applications built on Android. It optimizes resource use and simplifies the operating system for the company's gadgets as well as other Google certified hardware kits. The OS includes an automated upgrade framework to address the challenges inherent in different vendors being liable for upgrading their own handset operating systems.
Cooja	It includes a cross-layer WSN emulator based on Java. It enables simulation at several stages, from physical layer to application layer, as well as hardware emulation for a set of sensor nodes.
Contiki	It is a free and open source OS designed for interconnected, memory-constrained applications, with an emphasis on low-power IoT gadgets. It is used in a variety of applications, including streetlights, sound control for smart communities, environmental monitoring, and warning systems.
Dip Trace	It is an electronic design automation (EDA)/computer-aided design (CAD) software platform for constructing schematics and printed integrated circuit boards. It features a multilingual GUI. It is comprised of four modules: a circuit design editor, a PCB layout editor with an integrated shape-based auto-router, feature editor, and a design editor.
Proteus 8 Simulator	The Proteus Design Suite is a closed-source software application suite mainly used for electronic systems engineering. Electronic system engineers mostly use the program to produce blueprints and digital sketches for making integrated circuits.

Table 2.IoT hardware used in agriculture.

Hardware	Description
ESP8266	The ESp8266 is a Wi-Fi module that helps in establishing wireless connection between different components of IoT-based smart agriculture.
RTC module	This module helps in setting up a real time clock. It enables the module to keep track of precise timings of any operation to be performed by the designed IoT system
DHT11	The DHT11 is a widely used sensor for measuring humidity and temperature. The DHT11 temperature and humidity sensor is a widely used component. The sensor is equipped with a specialized NTC for temperature measurement and an 8-bit micro-controller for serial data output. Additionally, the sensor is factory tuned, making communication with other chipsets easy. The sensor has a precision of $\pm 1^{\circ}\text{C}$ and $\pm 1\%$ when measuring temperature between 0°C and 50°C & humidity between 20% and 90%. Therefore, if you need measurements within this range, this hardware might be the best choice
GSM module	GSM module It can allow SIM from any GSM network provider and can operate as a smartphone with its own specific contact number. The RS-232 protocol can be used to link the controller quickly
PIC Micro-controller	It is used in defence, monitoring devices and domestic appliances. Often included is an EEPROM that is being used to permanently store the content. This can be used to store the data collected by drones while analysing the farmland.

Table 3 . Different security concerns and attacks on smart agriculture.

Mode	Attack
Perception	Random Sensor incidents, Autonomous System Hijacking, Autonomous System Disruption, Irregular measurement, Sensor Weakening, Node Capture, Optical Deformation, Fake Node, Sleep deprivation
Data attack	Insider data leakage, Cloud data leakage, False data injection attack, Misinformation attack
Network	DoS/DDoS, Data Transit Attacks, Routing attacks, Signal disruptions, Radio Frequency (FR) jamming attack, Malware injection attack, Botnet attack, Side Channel attack
Edge	Forged control for actuators, Gateway-cloud request forgery, Forged measure injection, Booting, Unauthorized access, Man-in-the-middle, Signature wrapping, Flooding
Supply chain attack	Third-party attack, Supply chain software update attack, data fabrication attack, Interdiction attack
Application	Phishing, Malicious Scripts, DoS/DDoS
Other	Compliance and Regulation violation, Cyber Terrorism, Cloud computing attack

Conclusion

In this study, the combination of IoT technology, data collection sensors, and machine learning algorithms has demonstrated promising results in precision agriculture through an innovative IoT-compatible ML device for soil nutrition. Nutrient monitoring via the Internet of Things (IoT) provides an innovative approach to nutrition monitoring and recommendations in tropical regions. This technology allows for more precise and personalized nutrient management, offering potential benefits for health and well-being in tropical areas. However, considerations related to data privacy and accessibility must be taken into account to maximize the benefits of this solution.

Chapter 2

State of the art

1.1 Introduction

Modern agriculture faces increasingly complex challenges, including the need to produce more food for a growing global population while grappling with constraints such as climate change, dwindling natural resources, and pressure on arable land. In this context, the adoption of cutting-edge technologies such as the Internet of Things (IoT) and machine learning (ML) offers promising opportunities to transform traditional farming practices and enhance productivity, efficiency, and sustainability.

The goal of this study is to thoroughly review ten current studies that investigate the connection between IoT and ML systems and agriculture. The main goal is to examine these papers in order to determine their goals, approaches, data, and technologies and tools utilized in order to put their answers into practice. We shall pay particular attention to the following elements:

1.1.1 Research goals: We will look at the particular objectives of each project, including encouraging more sustainable farming methods, enhancing crop quality, managing resources efficiently, and optimizing agricultural output

1.1.2 Data used: We shall list the many kinds of data such as weather reports, soil profiles, sensor readings, etc. that are gathered and applied to each project.

1.1.3 Sensors: We will examine the sensors and Internet of Things (IoT) devices that are used to gather data, emphasizing their precision, dependability, and commercial accessibility.

1.1.4 Project cost: We will evaluate whether initiatives are feasible for farmers with low resources and whether they are created within a tight budget.

1.1.5 Machine learning use: We will ascertain whether projects use deep learning or machine learning methods for data analysis and insight generation.

1.1.6 Targeted anomalies and limitations: We shall list the particular anomalies such as crop diseases, nutrient deficits, etc. That each study has identified, along with the present constraints on their methods.

By looking at these elements, we intend to offer a thorough summary of new developments and prospects in the field of smart agriculture, along with recommendations for future study and encouragement of the use of cutting-edge technical solutions in the farming industry

2. Table 1

Table 4 . Summary of the most recent studies.

Study	Key findings	Data	Sensors	Low cost	ML	DL	Anomaly detection	Limitations
[1]	Continuous monitoring system of soil health The system aims to provide farmers with real-time actionable data to make informed decisions for enhancing agricultural productivity and the long-term health of the soil. (7)	Real time data	Water temperature, pH, turbidity, dissolved oxygen, conductivity, Water level, Teros 12,G MP25 1	No	Yes	No	Detecting unusual variations in sensor measurements, such as: Sudden spikes in temperature or humidity Unexplained drops in pH or electrical conductivity Decrease in nutrient content	No data analysis technique was conducted Limitations of current technologies High initial cost. LoRaWAN network coverage.
[2]	Soil analysis and classification of micronutrients to optimize resource utilization and improve overall crop yield and quality. (8)	Private data from “Tamil Nadu”	No sensors used	Yes	Yes	No	Data transmission issues: Errors in data transmission may indicate communication issues within the IoT system. Sudden or unexpected variations in micronutrient levels: Abrupt changes in micronutrient levels can indicate issues such as pollution, improper farming practices,	The dataset consists exclusively of the Tamil Nadu region The composition of soil can naturally vary from one location to another.

[3]	Developing an accurate model for high yield crops and advancing precision agriculture. (9)	Kaggle data (9)	No sensor s used	Yes	Yes	No	Imbalance of soil pH (too acidic or alkaline) Nutrient deficiency (iron, zinc, etc.) or toxicity due to excess	Data Quality: Accuracy and reliability of sensors: Data collected by IoT sensors may be inaccurate or erroneous.
[4]	Soil nutrients analysis and crop recommendation Optimizing the efficiency of agricultural practices by seamlessly integrating IoT data for accurate soil nutrient analysis and developing intelligent crop recommendation models, aiming to improve yields, reduce costs, and promote sustainable agriculture. (10)	Real time soil nutrient s data	LCD color sensor , moisture, temperature, Level indicator, global positioning system sensor	Yes	Yes	No	Out of range readings(sudden spikes or drops), sensor fusion, environmental factors Nutrient deficiency or excess. Detection of unusual variations in soil pH, indicating conditions that could affect nutrient absorption by plants.	The accuracy of the results depends on the reliability of IoT sensors. Measurement errors or hardware failures can lead to inaccurate data. Sensors often require regular calibration to maintain their accuracy, which can pose a logistical challenge for farmers.

[5]	Enhance the reliability and resilience of precision agriculture systems by developing innovative strategies for early anomaly detection, environmental stress management, and adaptation to unexpected crop responses, aiming to optimize agricultural yields and promote sustainable agriculture. (11)	Real time data	Integrated soil sensor , DHT11 MQ135	Yes	Yes	No	Sensors deviation or malfunction, environmental stress. Unexpected crop responses	The system is limited for few crops only.
-----	---	----------------	--------------------------------------	-----	-----	----	---	---

[6]	Cloud enabled crop recommendation platform Integrating a real-time weather-tracking feature into the platform, enabling farmers to receive tailored recommendations based on immediate climate changes. (12)	Kaggle Dataset (12)	No sensors used	Yes	Yes	No	Monitoring outliers in sensor data (extreme soil pH or humidity values). Detecting rapid weather changes (sudden temperature variations).	Not applicable for real time data Predictive models can be disrupted by unforeseen climate changes. "The costs associated with implementation can pose economic barriers for small-scale farmers."
-----	---	---------------------	-----------------	-----	-----	----	--	--

[7]	Develop and implement a crop recommendation system based on machine learning techniques to maximize agricultural yields. (13)	Politest Laboratories soil testing lab (India)	No sensors used	Yes	Yes	No	Sudden Variations in Nutrient Levels: Abrupt changes in soil nutrient levels can indicate issues with soil fertility or nutrient supply. Failures or deviations of IoT sensors: Measurement errors or malfunctions of sensors can lead to inaccurate data.	The models may not be sufficiently accurate due to complex variations in agricultural conditions. Some farmers, especially in rural areas, might have limited access to the necessary technologies.
[8]	Develop and implement an Internet of Things (IoT)-enabled and context-aware fertilizer recommendation system to optimize the efficiency of fertilizer application in agriculture. (14)	the department of agronomy (IUB) Pakistan. GitHub dataset (14)	NPK sensor	No	Yes	No	Sudden changes in the chemical composition of the soil that could indicate unusual conditions or measurement errors. An excessive or insufficient use of fertilizers compared to the recommendations.	Limited datasets, The deep neural network was not applicable.

[9]	Crop recommendation system Optimizing crop yield through Machine Learning (ML) and the Internet of Things (IoT) in the context of smart agriculture. (15)	Kaggle dataset (15)	No sensors used		Yes	No	Errors in machine learning models can lead to incorrect recommendations. Des rendements qui ne correspondent pas aux prévisions du modèle peuvent indiquer des problèmes ou des erreurs.	Limited dataset The initial installation of IoT sensors and the implementation of machine learning models can represent a significant financial investment for farmers. Agricultural conditions can vary significantly, making it challenging to create universal models suitable for all situations.
[10]	Continuously measures variable at the agro field Designing and implementing an innovative system based on the Internet of (16)Things (IoT) for precision agriculture.	Real time data	Yes 4 Soil moisture Soil temperature DHT11	Yes	Yes	No	Sensors deviation or malfunction, environmental stress. Unexpected crop responses	It must be connected with wires.

3.1 Related Work

In recent years, the integration of Internet of Things (IoT) technology with agriculture has gained significant momentum. This integration aims to enhance crop management, optimize resource use, and ultimately increase agricultural productivity. Specifically, the use of IoT for low-cost fertilizer recommendations using machine learning algorithms is an emerging area of research that promises to offer innovative solutions for precision agriculture.

In this chapter, we conduct a comprehensive comparative study of ten research articles that align closely with the theme of our project: "An IoT Low-Cost Fertilizer for Crop Recommendation using Machine Learning." Each article was carefully selected based on its relevance to IoT applications in agriculture, the use of machine learning for crop management, and the focus on cost-effective solutions. The purpose of this comparative study is to identify the strengths, weaknesses, and unique contributions of each work, thereby providing a solid foundation for our research.

The articles reviewed in this chapter cover a range of topics including IoT-based soil monitoring systems, machine-learning models for predicting crop nutrient requirements, cost-effective sensor networks, and case studies on the practical implementation of these technologies in different agricultural settings. By summarizing these studies, we aim to highlight the current state-of-the-art in the field, identify gaps in the existing literature, and position our work within the broader context of ongoing research

✓ **Article 1: A Self-Powered, Real-Time, Lora WAN IoT-based Soil Health Monitoring System**

Deep learning-based methods have shown promising results for precision agriculture according to the research analyzed in this technical document.

S.J. Ramson ,T. Blesslin Sheeba ,Kalaiselvi Bakthavatchalam and Murali Krishna Senapaty introduced a self-powered LoRaWAN IoT system for real-time soil health monitoring. The system employs sensors to measure parameters like soil temperature, pH and moisture. IoT devices transmit the data using LoRaWAN technology, enabling long-range communication with low energy use. This provides farmers with continuous insights into soil conditions.

They developed a machine learning algorithm for soil analysis and micronutrient classification in IoT-enabled farms. By leveraging IoT sensors, the algorithm provides real-time insights into soil nutrient levels. It aims to optimize resource use and improve crop yields through precision agriculture approaches and proposed an IoT framework combining sensors and machine learning algorithms to predict crop outcomes with high accuracy, aiding farmer decisions.

The researchers introduced an IoT-enabled system for real-time soil nutrient analysis and tailored crop recommendations, achieving accuracies above 98% in some cases.

In summary, this research demonstrates the potential of these methods to sustainably optimize agriculture through data-driven precision farming approaches. Further work can help address limitations and fully realize the benefits of these technologies.

✓ **Article 2: Machine Learning Algorithm for Soil Analysis and Classification of Micronutrients in IoT-Enabled Automated Farms**

In this article the authors analyze various techniques using the Internet of Things (IoT) and machine learning for soil analysis and classification of micronutrients in automated farms in the state of Tamil Nadu, India.

The authors T. Blesslin Sheeba, L. D. Vijay Anand, Gunaselvi Manohar, Saravana Selvan C. Bazil Wilfred, K. Muthukumar, S. Padmavathy, P. Ramesh Kumar, and Belete Tessema Asfaw developed a machine learning algorithm for soil analysis and micronutrient classification. By leveraging IoT sensors, the algorithm provides real-time insights into soil nutrient levels with a classification accuracy above 98%. It aims to optimize resource use and improve agricultural yields through precision techniques.

The study analyzed 1000 soil samples from 4 mulberry cultivating districts: Salem, Tirupur, Coimbatore and Erode. Chemical analyses revealed that soils generally contain high levels of potassium (35% of samples), nitrogen (80%) and sulfur (75%), but are sufficient or poor in magnesium, boron and zinc.

The authors propose improving soil management by developing diagnostic tools, private LoRaWAN networks, and applying statistical and machine learning models to IoT data. This would enable soil fertility mapping, anomaly detection and personalized recommendations for farmers.

Overall this study demonstrates the potential of IoT and machine learning technologies to sustainably optimize soil quality and agricultural productivity through precision techniques. Further research could overcome limitations and fully realize the promising methods.

✓ **Article 3: IoT Framework for Measurement and Precision Agriculture Predicting the Crop Using Machine Learning Algorithms**

This technical paper describes a methodology for predicting agricultural crops using an IoT framework and machine learning algorithms.

The authors propose a system with sensors connected to the IoT to collect data on soil and climate parameters. This data is stored in a cloud database and analyzed using supervised learning algorithms such as multilayer perceptron, JRip and decision tree.

A public dataset of 2200 instances describing 7 parameters for 22 crops was used. The algorithms were implemented with WEKA and achieved a classification accuracy above 98%.

The methodology demonstrates the potential of IoT and machine learning for precision agriculture. The authors suggest this type of system could help farmers optimize yields by recommending the most suitable crops.

Some possible improvements would be to include more parameters (soil type, diseases...) and test other deep learning algorithms. Field validation would also be needed to assess the reliability of the system under real-world conditions.

Nonetheless, this study presents a promising approach for the development of connected and precision agriculture using IoT and machine learning technologies.

✓ **Article 4: IoT-Enabled Soil Nutrient Analysis and Crop Recommendation Model for Precision Agriculture**

This technical paper describes an IoT-enabled soil nutrient analysis and crop recommendation (IoTSNA-CR) model for precision agriculture.

The model uses IoT sensors to collect real-time data on soil parameters like moisture, temperature, nutrients and pH value. This data is stored in cloud databases and analyzed using a novel MSVM-DAG-FFO algorithm.

The study collected data from 1000 soil samples across 4 mulberry cultivating districts in Tamil Nadu, India using sensors. Chemical analyses found soils generally contain high levels of potassium, nitrogen and sulfur but are sufficient/poor in other nutrients.

The MSVM classifier is optimized using a directed acyclic graph and fruit fly optimization algorithm (FFO) to select the best kernel function. Experiments on a public dataset achieved 97.3% accuracy, outperforming SVM, SVM-kernel and decision tree models.

The IoTSNA-CR model allows periodic soil data collection and analysis via sensors and cloud storage. It helps farmers maintain soil health and select optimal crops, reducing mineral use and increasing yields.

While demonstrating the potential of IoT and machine learning for precision agriculture, further work could include more parameters, crops and field validation of the model. Overall, this presents a promising approach for smart connected agriculture.

✓ **Article 5: Crop Yield Maximization Using an IoT-Based Smart Decision**

In this article the researchers introduced a smart decision system using IoT devices and machine learning algorithms for crop selection aimed at maximizing yields. The system considers parameters like soil nutrients, temperature, humidity, rainfall and more to accurately predict suitable crops.

- Real-time sensory data on soil nitrogen, phosphorus, potassium, pH, EC and temperature is collected using inexpensive sensors and sent to a cloud server for analysis. Five machine learning algorithms are ensembled, achieving 97-98% accuracy in crop prediction.

- Multiple linear regression is leveraged to predict future rainfall based on 15 years of historical data, obtaining a low root mean square error of 0.3%. This provides valuable insight for crop health.

- Experiments were conducted on loamy and clay soil samples to analyze nutrient levels. Initial tests showed the system correctly selecting wheat for a loamy field, resulting in good yields with optimized nutrient levels.

Then an Android application was developed to display selection results to farmers, aiding decisions through precision agriculture approaches. This addresses challenges of traditional methods facing low yields.

- The researchers introduced an IoT-enabled system combining real-time soil sensing and machine learning to predict crops with high accuracy above 98%, supporting farmer decision making.

- In summary, this research demonstrates the potential of such data-driven precision farming methods through IoT and algorithms to optimize and sustain agriculture by addressing limitations of conventional techniques.

✓ **Article 6: A Cloud Enabled Crop Recommendation Platform for Machine Learning-Driven Precision Farming**

- This article introduced a cloud-based Machine Learning (ML) platform for crop recommendation using real-time data from IoT sensors measuring soil nutrients, weather parameters and other factors.

- An IoT system was developed using sensors to monitor soil nitrogen, phosphorus, potassium, temperature, humidity, pH and rainfall levels on farms. They analyzed the data using ML algorithms like KNN, decision tree, random forest, XGBoost and SVM, achieving highest accuracy of 97.18% using random forest.

By leveraging the IoT sensors and ML algorithm, the system provided insights on soil conditions and suitable crops in real-time.

The researchers proposed an IoT and cloud-based framework combining the sensors and ML algorithms to predict crop outcomes with high accuracy, aiding farmer decisions.

A real-time crop recommendation platform was introduced leveraging the ML model for nutrient analysis and tailored recommendations.

- In summary, the study demonstrated how precision agriculture approaches optimized through data-driven ML methods can sustainably improve crop yields and resource use, further enhancements can help address limitations and realize full benefits by making the solutions more accessible to farmers worldwide.

✓ **Article 7: Crop Recommendation System to Maximize Crop Yield using Machine Learning Technique**

The researchers introduced an IoT-enabled machine learning-based crop recommendation system. An IoT soil monitoring system was developed employing sensors to measure soil nitrogen, phosphorus, potassium levels and other parameters from farms, leveraging these sensors to provide real-time insights into soil nutrient levels. The data was analyzed using machine learning algorithms SVM, ANN and Naive Bayes to predict suitable crops, developing an ensemble model through majority voting of the algorithms. This provided tailored crop recommendations to farmers, achieving over 98% accuracy in some cases. They proposed an IoT and machine learning-combined framework leveraging the sensors and algorithms to predict crop outcomes with high accuracy, aiding farmer decisions and optimizing resource use through precision agriculture approaches. The system introduced a real-time system for soil analysis and recommendations. In summary, the study demonstrated how these data-driven precision farming approaches through IoT and machine learning can sustainably enhance agriculture production, with further enhancements able to address limitations and fully realize the benefits of the technologies introduced.

✓ **Article 8: Internet of Things (IoT) assisted Context Aware Fertilizer recommendation**

The researchers introduced an IoT-enabled machine learning system for context-aware fertilizer recommendation. The system utilized IoT sensors to measure soil nitrogen, phosphorus, potassium levels and other parameters from farms, providing real-time insights into soil nutrient levels through an IoT-assisted soil fertility mapping system. The data was analyzed using machine learning algorithms such as logistic regression, SVM, Gaussian Naive Bayes, and KNN to predict suitable fertilizers. An ensemble model was developed through majority voting of the algorithms, which provided tailored fertilizer recommendations based on soil fertility levels, crop type, and soil type with over 98% accuracy in some cases, as demonstrated on real crop fields. The study demonstrated how data-driven precision agriculture approaches leveraging IoT and machine learning technologies can sustainably optimize agriculture by addressing limitations of conventional techniques. Further enhancements to the system could help address remaining challenges and fully realize the benefits of the introduced technologies for wider adoption.

✓ **Article 9: Crop Recommendation System using Machine Learning**

The document discusses developing a crop recommendation system that uses machine learning algorithms like Decision Tree, Support Vector Machine, Logistic Regression, and Gaussian Naive Bayes to analyze a dataset containing parameters like Nitrogen, Phosphorous, Potassium, soil pH, humidity, temperature, and rainfall for 2,200 data points on 22 crops specifically in India, in order to predict future crop production and recommend the best crops to farmers based on soil nutrients, soil pH, humidity, and rainfall. It outlines collecting datasets, pre-processing data, extracting features, applying the ML algorithms, and recommending crops, compares the algorithms' accuracies, discusses related works that have used ML for crop yield prediction and recommendation, and concludes that the system aiming to increase farmers' profits presented in the paper could help boost India's agricultural productivity and incomes, while noting the models using Decision Tree and Gaussian NB achieved the best accuracy of around 99% and identifying potential future work areas like improving the dataset and developing a user-friendly interface.

✓ **Article 10: An innovative IoT based system for precision farming**

This article presents a machine learning based crop recommendation system that analyzes different machine learning models such as decision trees, support vector machines, logistic regression, and naive bayes to predict the future production of 22 crops in India using a dataset containing parameters like nitrogen, phosphorus, potassium, soil pH, humidity, temperature, and precipitation for 2,200 data points. It outlines the methodology including data collection, pre-processing, feature extraction, applying the ML algorithms, and recommending crops, discusses related works using ML for crop yield prediction and recommendation based on soil characteristics, compares the algorithms' accuracies finding decision trees and naive bayes achieved around 99% accuracy, identifies potential future work such as improving the dataset and developing a user-friendly interface, aims to increase farmers' profits by recommending optimal crops based on environmental conditions, and concludes that the paper presents an ML model for crop recommendation that could help boost India's agricultural productivity and incomes

Through this comparative study, we have gained a deeper understanding of the current advancements and challenges in using IoT and machine learning for crop management. The insights drawn from these articles will inform the design and implementation of our own IoT-based fertilizer recommendation system. Our approach aims to build upon the strengths of existing work while addressing identified gaps, particularly in terms of cost-effectiveness and scalability.

Conclusion

In conclusion, this study's comprehensive review of 10 recent studies emphasizes how crucial it is for agriculture to use machine learning (ML) and Internet of Things (IoT) technology. These studies unequivocally show how important a role these technologies may play in streamlining resource management, encouraging more sustainable farming, and optimizing agricultural practices.

We have been able to determine the advantages and disadvantages of the present smart agriculture systems by looking at the goals, data utilized, sensors used, project costs, machine learning application, targeted anomalies, and obstacles faced. This data will be useful in directing future research and the creation of more efficient technical solutions that are more in line with the actual needs of farmers.

Additionally, the suggested ways to enhance current initiatives provide practical ways to address issues that have been found and optimize the benefits that these technologies may provide to agriculture. For farmers, no matter where they live or what resources they have at their disposal, to universally adopt these solutions, they must be easily obtainable, reasonably priced, and appropriate.

In the end, this research highlights the tremendous potential of IoT and machine learning to address present and future difficulties in global agriculture and strengthens our awareness of the growing significance of the convergence between agriculture and digital technology. We can build a future where agriculture is more productive, sustainable, and resilient to the economic and environmental problems it faces by funding research and development of creative solutions.

In the subsequent chapters, we will detail the methodology and experimental setup of our project, demonstrating how our solution integrates the best practices identified in this comparative study and introduces novel contributions to the field of precision agriculture.

Chapter 3

Conception and Development

1.1 Introduction

The design and execution of the suggested system are the main topics of Chapter 3 of this project. This chapter explores the manifestation of these notions into a workable and useful solution after outlining the prerequisites and specifications in the other chapters. The development process's design and execution phases are critical because they establish the viability and efficacy of the suggested solution.

We will cover each stage of the design process in this chapter, from identifying functional requirements to choosing the right technology. Additionally, we will outline the intended system design, emphasizing important parts and how they work together. We will also look at the technology decisions taken to put the system into place, such as those involving sensors, cloud services, machine learning algorithms, and IoT platforms.

Next, we shall describe the process of implementation, outlining the actions done to put the system together and assemble its components. In addition, we will give specific instances of how the system works and its potential to track soil nutrients, suggest crops, and give farmers insightful information.

Lastly, we will assess the system's effectiveness by contrasting the outcomes with the original goals and looking at its accuracy, dependability, and user-friendliness. Through this assessment, we will be able to make judgments about how well the system was designed and implemented as well as whether or not it could be used to solve the issues in the agricultural sector.

The term "soil NPK" describes the three major macronutrients potassium (K), phosphorus (P), and nitrogen (N) that are necessary for plant growth. These nutrients are essential for several stages of plant development, including:

1. Nitrogen (N): Nitrogen is necessary for the growth of leaves and is a key component in the production of chlorophyll, the green pigment that is involved in photosynthesis. It contributes to the synthesis of proteins, enzymes, and amino acids all essential for the metabolism and general growth of plants.

2. Phosphorus (P): Phosphorus is necessary for the growth of roots, flowers, and seeds. It is a part of phospholipids, nucleic acids, and ATP (adenosine triphosphate), all of which are necessary for plant membrane construction, cell division, and energy transfer.

3. Potassium (K): Potassium plays a role in several physiological functions in plants, such as nutrient transport, water management, and photosynthesis. Strong stems and general plant vigor are encouraged, disease and stress tolerance is improved, and enzymes involved in a variety of metabolic processes are activated.

Apart from NPK, plants also need additional vital components, which are commonly called micronutrients, but in smaller amounts. These elements have distinct functions in the growth and development of plants, and they include calcium, magnesium, sulfur, iron, manganese, zinc, copper, boron, molybdenum, chlorine, and nickel. It is imperative to maintain a sufficient supply of all these nutrients in order to sustain robust and fruitful crops.

Optimizing crop output, quality, and plant health depends on keeping the right amounts of nitrogen, phosphorus, and potassium in the soil. Farmers and other agricultural professionals can measure the amounts of these nutrients in the soil with accuracy thanks to NPK sensors, which are useful tools. Farmers may ensure that plants receive the proper amount of nutrients for optimal growth by making educated decisions about fertilization with the use of real-time data. Usually, these sensors identify physical properties of the soil, such as electrical conductivity, which are subsequently utilized to modify fertilizer application techniques, increasing nutrient usage efficiency and lowering environmental impact. Farmers may create precise nutrient management plans suited to their unique crop and field requirements by combining NPK sensor data with other agronomic data. This strategy reduces nutrient intake and maximizes resource use. Losses, and support environmentally friendly farming methods that are crucial for the long-term fertility and health of the soil.

Part one: Materials and software used

2.1 what is an NPK sensor and how it works

An NPK sensor is a device that measures the levels of nitrogen (N), phosphorus (P), and potassium (K) in soil. These three nutrients are the main ingredients of fertilizers and help plants grow faster. (19)



Figure 2. JXCT Soil NPK sensor

The sensor works by passing an electric current through the sample, and the resulting electrical resistance is used to calculate the concentrations of each element. The sensor is able to calculate the concentration of each element in the sample.

The JXCT Soil NPK sensor operates on a voltage range of 9-24V and has low power consumption. With a high-resolution capacity of up to 1 mg/kg (mg/l), the sensor can accurately measure the concentration of nitrogen, phosphorus, and potassium in the soil, allowing farmers and agricultural professionals to make informed decisions about fertilization and nutrient management.

2.1.1 Required Materials

To Interface the NPK soil sensor with Arduino, you will need:

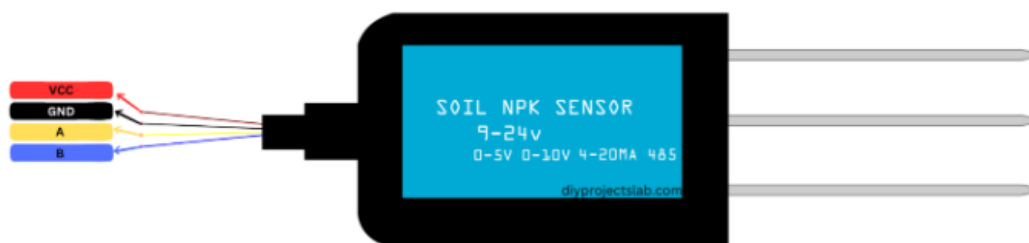
- ✓ Arduino Uno board
- ✓ NPK soil sensor kit
- ✓ Jumper wires
- ✓ USB cable



Figure 3 .the assembly carried out

2.1.2 Soil NPK sensor Pinout

The figure below displays the pinout for the sensor



-  **VCC pin:** is used to connect to a voltage supply ranging from 5V to 30V.
-  **A pin:** is a differential signal that should be connected to the A pin of the MAX485 Modbus Module.
-  **B pin:** is another differential signal that should be connected to the B pin of the MAX485 Modbus Module.
-  **Ground pin (GND):** should be connected to the ground.

Typically, a soil NPK (nitrogen, phosphorus, and potassium) sensor has several probes that are placed into the soil to collect exact data on the concentrations of these vital nutrients for plant growth. Farmers are able to make real-time adjustments to their fertilization operations because it continuously checks these levels. This sensor is built to endure a wide range of environmental conditions, such as intense heat and heavy rain, and it guarantees dependable performance even in the most difficult settings. Furthermore, the gathered data may be simply visualized with specialized software programs, making interpretation and analysis easier for well-informed soil nutrient management decision-making.

3.1 Soil NPK sensor Specs

Here are the specifications for Soil NPK Sensor: (19)

Operating voltage	DC 9-30V
Power consumption	less than 20Ma
Measurement range	0-9.99 mg/L
Baud Rate	2400/4800/9600
Measurement accuracy	±10% or ±0.1 mg/L
Resolution	1mg/kg
Precision	±2%FS
Measurement time	2-3 seconds
Output signal	RS485 (Modbus protocol)
Probe material	stainless steel
Protection class	IP68
Probe length	70mm
Probe diameter	12mm
Cable length	3m
Operating temperature	0-50°C
Operating humidity	5-90% RH
Weight	150g

Table 5. the specifications for Soil NPK Sensor

 **Note:** Soil NPK sensor NPK sensor can measure nitrogen, phosphorus, and potassium with a resolution of up to 1 mg/kg (mg/l).

4.1 NPK sensor Application

The Soil NPK sensor finds applications across various fields, including:

4.1.1. Environmental Monitoring: In environmental monitoring, NPK sensors are employed to assess soil health and nutrient dynamics in ecosystems. They aid in understanding nutrient cycling, soil degradation, and the impact of land use practices on environmental quality.

4.1.2. Horticulture: NPK sensors play a crucial role in horticulture by helping gardeners, landscapers, and greenhouse operators monitor soil fertility and nutrient levels. This information guides fertilizer application and irrigation practices to promote healthy plant growth and development.

4.1.3. Soil Testing Laboratories: Soil testing laboratories utilize NPK sensors as part of their soil analysis services. These sensors provide rapid and accurate measurements of nitrogen, phosphorus, and potassium levels in soil samples, aiding in soil fertility assessments and fertilizer recommendations.

4.1.4. Education: NPK sensors are valuable educational tools in agricultural and environmental science programs. They enable students to learn about soil fertility, nutrient management, and the importance of balanced nutrition for plant growth through hands-on experimentation and data analysis.

In summary, the versatility of Soil NPK sensors makes them indispensable in a wide range of applications, spanning agriculture, environmental monitoring, horticulture, soil testing laboratories, and educational settings.

5.1 MAX485 TTL to RS-485 Module Converter

The MAX485 TTL to RS-485 module is a reliable, low-cost interface for connecting TTL-level devices to RS-485 networks. It is based on the MAX485 IC, which is a low-power transceiver for RS-485 communication. The module features a 5V power supply and can operate at speeds up to 2.5 Mbps. (19)

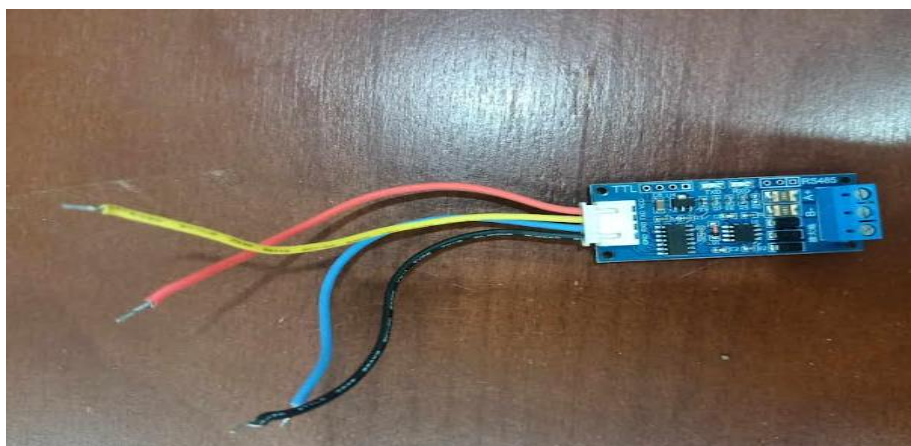


Figure 4. The MAX485 TTL

The RS-485 has the ability to communicate with multiple devices (up to 32) on the same bus/cable when used in a master and slave configuration.

We use a TTL (Transistor-Transistor Logic)-to-RS485 interface signal chip, MAX485, to convert the interface communication mode between the Arduino board and the Soil NPK Sensor. It receives the TTL (Transistor-Transistor Logic) signal from the Arduino and converts it to an RS-485 signal, which is then sent to the Soil NPK Sensor.

5.1.1 Specs MAX485 TTL to RS-485 Module

- Compatible with Arduino
- Supports half-duplex communication mode
- 5V operation
- Supports multiple data rates up to 2.5 Mbps
- Long-distance communication up to 1200 meters
- Compact size and easy to use
- Supports multiple interface standards such as EIA/TIA-485 and ISO 8482:1987(E)

5.1.2 Pinout MAX485 TTL to RS-485 Module



Data side:

- ✓ RO = TTL Data Output (connects to Arduino RX pin)
- ✓ RE = Receiver Enable. Generally joint to DE Pin.
- ✓ DE = Driver Enable. Generally joint to RE Pin.
- ✓ DI = TTL Data Input (connects to microcontroller TX pin)



Output side

- ✓ VCC = 5V
- ✓ B = RS-485 differential data signal B
- ✓ A = RS-485 differential data signal A
- ✓ GND = Ground

6.1 Interfacing NPK Soil Sensor with Arduino

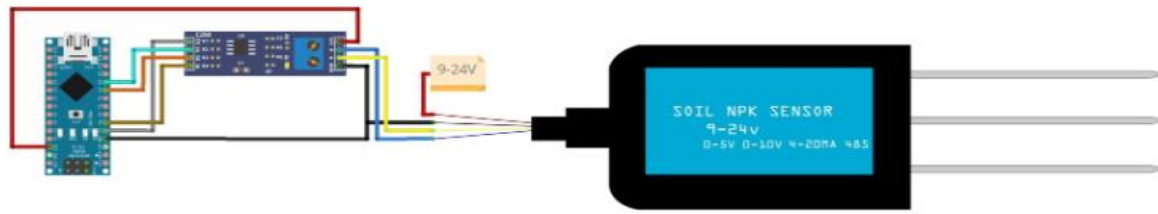


Figure 5. Interfacing NPK Soil Sensor with Arduino



Figure 6. The assembly carried out

To connect the NPK sensor to an Arduino, first, connect the R0 & DI pins from the Modbus to the D2 & D3 pins of the Arduino using Software Serial.

Then, connect the DE & RE pins of the Modbus to the D7 & D8 pins of the Arduino to enable DE & RE high. The NPK sensor has four wires; the brown wire is for VCC, which requires a 9V-24V power supply. Connect the black wire to the GND of the Arduino.

The blue wire (B pin) should be connected to the B pin of MAX485 and the yellow wire (A pin) should be connected to the A pin of MAX485.

Soil NPK Sensor	MAX485	Arduino
VCC (5V-30V)	-	External Power Supply
GND	GND	Ground
B	MAX485 B	
A	MAX485 A	
	Modbus RX	D2
	Modbus TX	D3
	RS485 DE	D7
	RS485 RE	D8

7.1 Digital Instant Soil NPK Tester



Figure 7. Soil NPK Tester

The Digital Soil Instant NPK Nutrient Tester is a valuable tool for farmers and agriculture professionals seeking to quickly and accurately assess essential nutrient levels in the soil. Here is some information about this device:

7.1.1. Operation: The soil NPK nutrient tester operates by measuring the levels of three key nutrients: nitrogen (N), phosphorus (P), and potassium (K) present in the soil. It typically employs rapid and precise detection technology to provide results within minutes.

7.1.2. Ease of Use: Designed to be user-friendly, the soil NPK tester often features a simple and intuitive interface, making it easy to use even for non-technical users. Some models may also be connected to mobile applications for deeper analysis and interpretation of results.

7.1.3. Portability: Most soil NPK nutrient testers are compact and portable, allowing users to easily carry them into the field for real-time testing. This portability provides great flexibility for farmers and agriculture professionals to monitor soil conditions across different sites.

7.1.4. Accuracy: Soil NPK nutrient testers are designed to deliver accurate and reliable measurements of nutrient levels in the soil. This precision enables users to make informed decisions regarding nutrient management and optimize fertilization practices for improved agricultural yields.

7.1.5. Applications: The soil NPK nutrient tester finds applications in various fields, including precision agriculture, crop management, agricultural research, environmental studies, and agricultural education.

In summary, the Digital Soil Instant NPK Nutrient Tester is a versatile and convenient tool that offers a quick and accurate method for assessing essential nutrient levels in the soil. It enables farmers and agriculture professionals to make informed decisions regarding nutrient management and improve crop yields.

8.1 Product Description

Digital Soil Fertility Available Instant NPK Nutrient Tester Meter (20)

Product Name	instant digital soil available NPK fertility nutriment
Measuring range	0 - 1999 mg/kg (mg/l)
precision	±2 % F. S.
Resolution	1 mg/kg (mg/l)
Show LCD screen with backlight	
Probe length	140 mm
Working temperature	5-45 41-113 ° ° C(F)/ATC
Functionality	Auto-switch-off in about 8 minutes
Calibration	Factory calibrated standard
The dimension	Ø50.5 x 250 mm
Weight	100g
Alimentation	DC 9V((9V280 / 6F22 rechargeable battery Ni-MH)

Table 6 . The specifications for Digital Soil Fertility Available Instant NPK Nutrient Tester Meter



Figure 8 . Element of NPK Sensor



Figure 9. Display of Digital Soil Instant NPK Nutrient Tester

9.1 The use of Digital Soil Instant NPK Nutrient Tester

Here is how to use the Soil NPK Nutrient Tester: (20)

1. Remove the protective cap from the metal probe.
2. Press the OFF/ON button. The tester will display 0.

Quick Test

- Water the soil with distilled water before testing. Ensure the water penetrates to a depth of at least 10 cm. ideally; the soil should have a moisture content of 70% to 80%.
- Insert the metal probe into the soil vertically and in a clockwise direction, to a depth of approximately 6-8 cm. Ensure the metal probe is in close contact with the soil.
- After about 6 seconds, the tester will display the measured value. Note that soil tightness can affect the measurement result. It is recommended to perform multiple tests at different locations and take the average as the result.

Precise Test

- Use a 200 ml beaker. Mix dried soil with distilled water by stirring to form a mud.
 - Insert the metal probe into the mud to a depth of about 6-8 cm. After about 6 seconds, the tester will display the measured value.
4. After each test, wipe the metal probe with a towel or tissue paper. Ensure it is clean and dry until the display returns to 0.

10.1 NPK sensor Application

The Soil NPK Nutrient Tester is designed for on-site use and offers a quick and easy way to measure the amounts of potassium, phosphorus, and nitrogen in the soil.

10.1.1. Field Agriculture: Before choosing how much fertilizer to apply, farmers can rapidly determine the nutrient levels of their soil by using the tester right there in their fields. This enables them to modify fertilization techniques in real-time to satisfy the distinct nutrient requirements of crops.

10.1.2. Soil Quality Monitoring: Using the tester, agronomists, and specialists in soil conservation can assess the quality of the soil in various agricultural settings. This makes it possible to precisely manage the soil by identifying regions where nutrient levels may be either excessive or inadequate.

10.1.3. Analysis of Farmland: The tester is a tool that soil analysis laboratories can use to quickly and accurately analyze farms. In order to make informed management decisions, this enables farmers to get real-time information on the nutrient levels in their soil.

10.1.4. Field Research: To get information on soil nutrients for field research investigations, researchers and agricultural scientists might utilize the tester. This advances our knowledge of soil nutrient cycles and advances sustainable farming methods.

10.1.5. Agricultural Education: To teach students the principles of crop fertilization and soil nutrient management, agricultural educators can integrate the tester into training curricula. This enables students to learn the practical skills of soil analysis as well as the significance of nutrients for plant growth.

In conclusion, using the Soil NPK Nutrient Tester in the field provides a quick and easy way to measure the amounts of potassium, phosphorus, and nitrogen in the soil; it has many uses in soil management, research, and agriculture.

11.1 Informatics cards

Arduino is an open-source electronic development board based on an easy-to-use hardware and software platform. This card is designed for electronics enthusiasts and professionals who want to create interactive projects using basic electronic components such as sensors, actuators and integrated circuits. The Arduino card is equipped with a programmable microcontroller, digital and analog inputs and outputs, as well as serial and USB communication ports for programming and data transfer. The Arduino card is available in different sizes and models to meet the specific needs of electronic projects and is widely used in home automation, robotics, interactive art and the Internet of Things (IOT) projects.

The types of the Arduino card:

- Arduino Uno.
- Arduino Nano.
- Arduino Due.
- Arduino Mega.

 **In our project, we will use Arduino Mega.**

The advantages of the Arduino Mega board are:

-  Its simplicity of use, its low cost.
-  Its large community of developers, which allow users to easily launch into electronics and quickly, develop functional prototypes.
-  It is also compatible with a wide range of sensors and electronic modules.

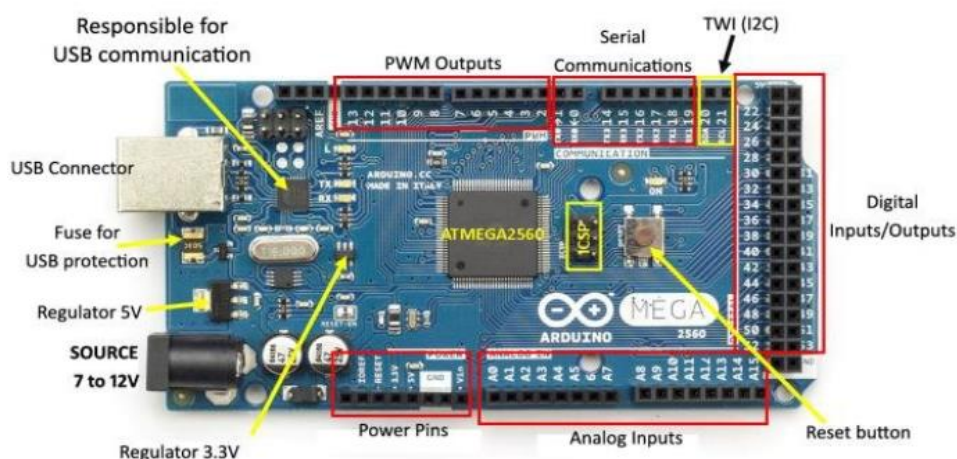


Figure 10 . Arduino card architecture

Implementation environment

11.2 Arduino IDE

Arduino IDE (Integrated Development Environment) is free software for programming Arduino boards. It provides a user-friendly interface for writing, compiling and uploading code to an Arduino board. It supports several programming languages, including C and C++, which are widely used in the Arduino community. IDE is available on various platforms including Windows, Mac OS and Linux. It also has a large library of built-in features to facilitate programming, as well as an active online community that provides code examples, additional libraries and community support. In addition, the IDE offers useful debugging tools to detect programming errors and to help solve problems encountered when developing Arduino projects. (22)



Figure 11 . Arduino ide.

The Arduino IDE interface consists of several elements:

-  Menu bar - which contains drop-down menus to access IDE functions such as File, Edit, Tools, Help, etc.
-  Toolbar - which contains icons for common functions such as New, Open, Save, Upload, etc.
-  Text box - which is where you write your Arduino code. There is also a message area at the bottom that displays information about errors, warnings and compilation and upload results.
-  Library window - which allows you to search and add additional libraries to use additional features with your Arduino board.
-  Map and serial port pull-down menu - which allows you to select the map and serial port you are working on.
-  Message area at the bottom - displays information about errors, warnings and compilation and upload results.

11.3. Jupyter

Jupyter emerged from the IPython project, originally conceived as an enhanced interactive interpreter for Python. In 2014, the project evolved into Jupyter to support multiple programming languages. It is compatible with different platforms including Windows, macOS, and Linux, making it accessible to a wide range of users.

We can be integrated with other popular data science tools and libraries such as NumPy, pandas, Matplotlib, SciPy, scikit-learn, TensorFlow, and PyTorch, making it a preferred choice for developing and analysing complex projects.

In addition to local installation, Jupyter can also be deployed in the cloud, allowing users to work on notebooks without having to install anything locally. Services like Google Colab, Microsoft Azure Notebooks, and AWS SageMaker offer Jupyter environments in the cloud.

Since Jupyter can be used to execute Python code and other languages, it is important to consider security and privacy issues, especially when running notebooks from external sources or sharing them publicly.

It has an active community of users and contributors who provide support, develop new features, and share knowledge through online forums, discussion groups, and conferences.



Figure 12. jupyter

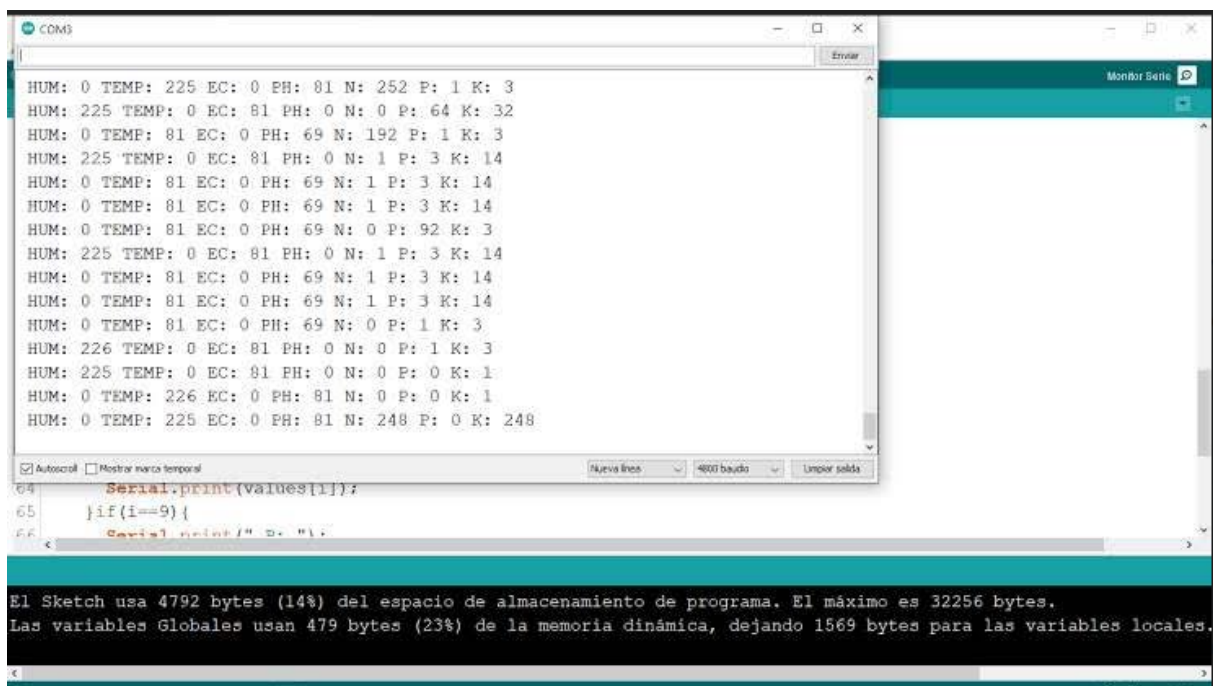
Data science libraries are pre-written toolsets that facilitate the work of analysts and researchers by providing ready-to-use functionalities for specific tasks. Here are some key libraries and their roles:

-  **NumPy:** For numerical computation and efficient data processing.
-  **Pandas:** For manipulation and analysis of tabular data.
-  **Matplotlib:** For creating static plots.
-  **Seaborn:** For creating attractive statistical visualizations.
-  **scikit-learn:** For developing machine learning models
-  **TensorFlow and PyTorch:** For deep learning model development.
-  **SciPy:** For advanced data science tasks like optimization and statistical analysis

Together, these libraries offer a comprehensive set of tools for all stages of the data analysis process, from initial pre-processing and exploration to modelling and visualization of results. They are widely used in the data science community to solve a variety of problems and make data-driven decisions.

To obtain the measurements from Arduino using the serial monitor, we follow these steps:

1. Connect our Arduino board to the computer using a USB cable.
2. Open the Arduino IDE on the computer and Select the correct board and port from the "Tools" menu in the Arduino IDE.
3. Upload the code to the Arduino board that includes the measurements we want to obtain and once the code is uploaded, click on the magnifying glass icon in the top right corner of the Arduino IDE to open the serial monitor.
4. Set the baud rate in the serial monitor to match the baud rate specified in your code (9600 in our case).
5. Then the serial monitor will display the measurements sent by the Arduino and each measurement will be printed on a new line.



The screenshot shows the Arduino IDE interface with the Serial Monitor open. The Serial Monitor displays a series of sensor readings for Humidity (HUM), Temperature (TEMP), Electrical Conductivity (EC), pH, Nitrogen (N), Phosphorus (P), and Potassium (K). The readings are printed on multiple lines, showing variations in values. Below the Serial Monitor, the Arduino IDE code editor is visible, showing a loop that prints the values of an array. At the bottom, a status bar indicates the memory usage of the sketch.

```
COM3
HUM: 0 TEMP: 225 EC: 0 PH: 81 N: 252 P: 1 K: 3
HUM: 225 TEMP: 0 EC: 81 PH: 0 N: 0 P: 64 K: 32
HUM: 0 TEMP: 81 EC: 0 PH: 69 N: 192 P: 1 K: 3
HUM: 225 TEMP: 0 EC: 81 PH: 0 N: 1 P: 3 K: 14
HUM: 0 TEMP: 81 EC: 0 PH: 69 N: 1 P: 3 K: 14
HUM: 0 TEMP: 81 EC: 0 PH: 69 N: 1 P: 3 K: 14
HUM: 0 TEMP: 81 EC: 0 PH: 69 N: 0 P: 92 K: 3
HUM: 225 TEMP: 0 EC: 81 PH: 0 N: 1 P: 3 K: 14
HUM: 0 TEMP: 81 EC: 0 PH: 69 N: 1 P: 3 K: 14
HUM: 0 TEMP: 81 EC: 0 PH: 69 N: 1 P: 3 K: 14
HUM: 0 TEMP: 81 EC: 0 PH: 69 N: 0 P: 1 K: 3
HUM: 226 TEMP: 0 EC: 81 PH: 0 N: 0 P: 1 K: 3
HUM: 225 TEMP: 0 EC: 81 PH: 0 N: 0 P: 0 K: 1
HUM: 0 TEMP: 226 EC: 0 PH: 81 N: 0 P: 0 K: 1
HUM: 0 TEMP: 225 EC: 0 PH: 81 N: 248 P: 0 K: 248

[Autoscroll] [Mostrar marca temporal] [Nuevas líneas] [4800 baudios] [Limpiar salida]

64 Serial.print(values[i]);
65 }if(i==9){
66   Serial.print("\n D: ");

El Sketch usa 4792 bytes (14%) del espacio de almacenamiento de programa. El máximo es 32256 bytes.
Las variables Globales usan 479 bytes (23%) de la memoria dinámica, dejando 1569 bytes para las variables locales.
```

Figure 13 .Soil nutriments monitoring

➤ Part two: Implementation and results

12.1 Methodology

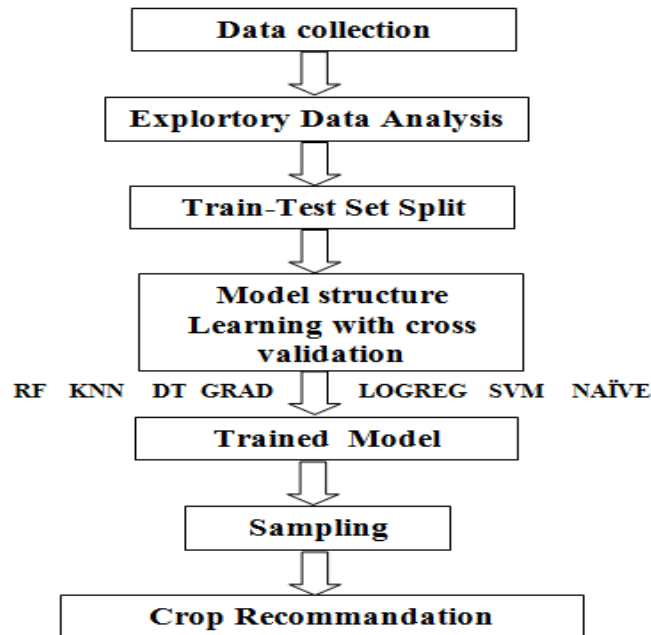


Figure 14 .Methodology for the current study

1. Data Collection

In our study we used a dataset that have been obtained from kaggle web site, it consists of 7 essential parameters such as Nitrogen (N), Phosphorous (P), Potassium (K), PH, Humidity,

Temperature and Rainfall. And it has 2200 instance and include 22 different crops such as rice, maize, chickpea, kidney beans, etc. And it was separated in Train and Test sets in which 80% of the dataset is taken as Train and 20% as a Test, The first 10 instances of our dataset are presented in table below.

	N	P	K	temperature	humidity	ph	rainfall	label
0	90	42	43	20.879744	82.002744	6.502985	202.935536	rice
1	85	58	41	21.770462	80.319644	7.038096	226.655537	rice
2	60	55	44	23.004459	82.320763	7.840207	263.964248	rice
3	74	35	40	26.491096	80.158363	6.980401	242.864034	rice
4	78	42	42	20.130175	81.604873	7.628473	262.717340	rice
5	69	37	42	23.058049	83.370118	7.073454	251.055000	rice
6	69	55	38	22.708838	82.639414	5.700806	271.324860	rice
7	94	53	40	20.277744	82.894086	5.718627	241.974195	rice
8	89	54	38	24.515881	83.535216	6.685346	230.446236	rice
9	68	58	38	23.223974	83.033227	6.336254	221.209196	rice

Table 7. First 10 records of the Dataset

2. Exploratory Data Analysis (EDA)

After loading the data, it is essential to perform exploratory data analysis to gain insights into the dataset. This step involves examining the data's structure, checking for missing values, identifying outliers, and visualizing the relationships between different features and the target variable (crop type).

	N	P	K	temperature	humidity	ph	rainfall
count	2200.000000	2200.000000	2200.000000	2200.000000	2200.000000	2200.000000	2200.000000
mean	50.551818	53.362727	48.149091	25.616244	71.481779	6.469480	103.463655
std	36.917334	32.985883	50.647931	5.063749	22.263812	0.773938	54.958389
min	0.000000	5.000000	5.000000	8.825675	14.258040	3.504752	20.211267
25%	21.000000	28.000000	20.000000	22.769375	60.261953	5.971693	64.551686
50%	37.000000	51.000000	32.000000	25.598693	80.473146	6.425045	94.867624
75%	84.250000	68.000000	49.000000	28.561654	89.948771	6.923643	124.267508
max	140.000000	145.000000	205.000000	43.675493	99.981876	9.935091	298.560117

Table 8 . Statistical summary table

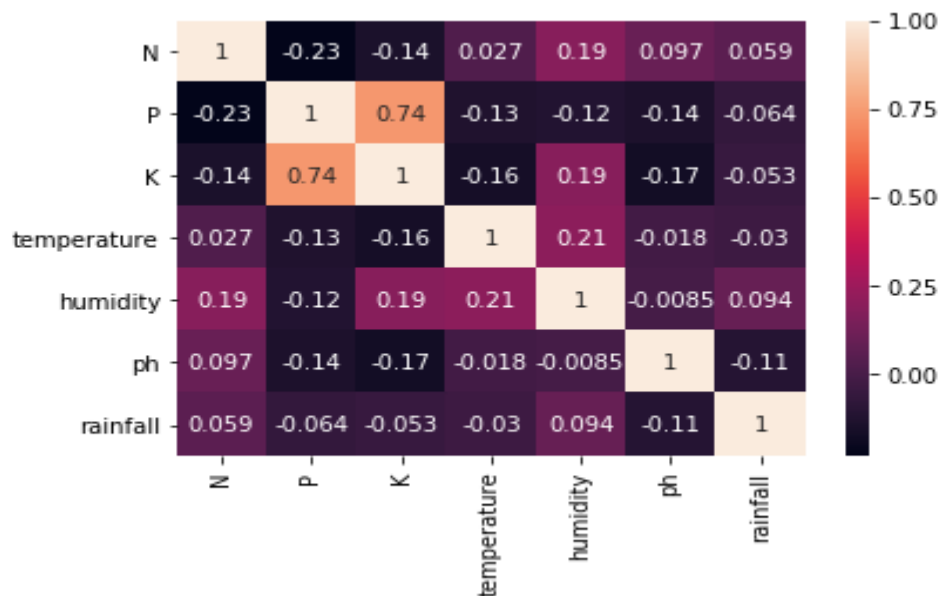


Figure 15 .Correlation Metrics on the crop recommendation dataset

3. Splitting the Data into Training and Testing Sets

To evaluate the performance of our models and ensure they generalize well to unseen data, we need to split the preprocessed dataset into training and testing sets. The training set is used to train the models, while the testing set is held out for evaluating the models' performance on unseen data.

4. Training the Models

In this step, we trained various machine learning models, such as Decision Trees, Random Forests, Logistic Regression, Support Vector Machine, K Nearest Neighbors, Gradient Boosting Classifier and Gaussian Naïve Bayes, on the training dataset. This involves feeding the training data to the individual classifiers and allowing them to learn the patterns and relationships between the features and the target variable.

Training Models

Decision tree

```
from sklearn.tree import DecisionTreeClassifier
DecisionTree = DecisionTreeClassifier(criterion="entropy",random_state=2,max_depth=5)
#trains the decision tree model on the training data
DecisionTree.fit(Xtrain,Ytrain)
#predict the target values using the trained decision tree model on the test data
predicted_values = DecisionTree.predict(Xtest)
#calculate the accuracy score
x = metrics.accuracy_score(Ytest, predicted_values)
acc.append(x)
model.append('Decision Tree')
print("DecisionTrees's Accuracy is: ", x*100)
print(classification_report(Ytest,predicted_values))
```

Figure 16 .Training machine learning model

5. .Model Evaluations and Selection

After training the models, we evaluated their performance on the testing set using appropriate metrics such as accuracy, precision, recall, and F1-score. This step helped us to identify the model that perform best on the given task and select the most suitable one.

In this section, we will provide a detailed explanation for each algorithm employed in this study:

➤ The Gradient Boosting Classifier

The Gradient Boosting Classifier is an ensemble machine-learning algorithm that combines multiple weak decision tree models to create a strong predictive model. In our study , it is used to solve a crop recommendation problem, where the model is trained on various features (N, P, K, temperature, humidity, ph, rainfall) to predict the type of crop (label).And here's how the Gradient Boosting Classifier works in this example:

First of all our data is loaded from a CSV file named "Crop_recommendation.csv" and split into features (N, P, K, temperature, humidity, ph, rainfall) and the target variable (label) .After loading the data from the CSV file and split it into training and testing sets using the train_test_split function from scikit-learn. Our Gradient Boosting Classifier is finally created, and the fit method is called to train the model on the training data (Xtrain, Ytrain).

During the training process, the Gradient Boosting Classifier starts with a single decision tree model and iteratively adds new decision tree models to the ensemble. Each new model is trained on the residual errors of the previous ones, with the goal of correcting the errors made by the previous models.

The learning process continues until a specified number of models are added or until a certain performance threshold is reached. After training, the predict method is used to make predictions on the test data (Xtest).

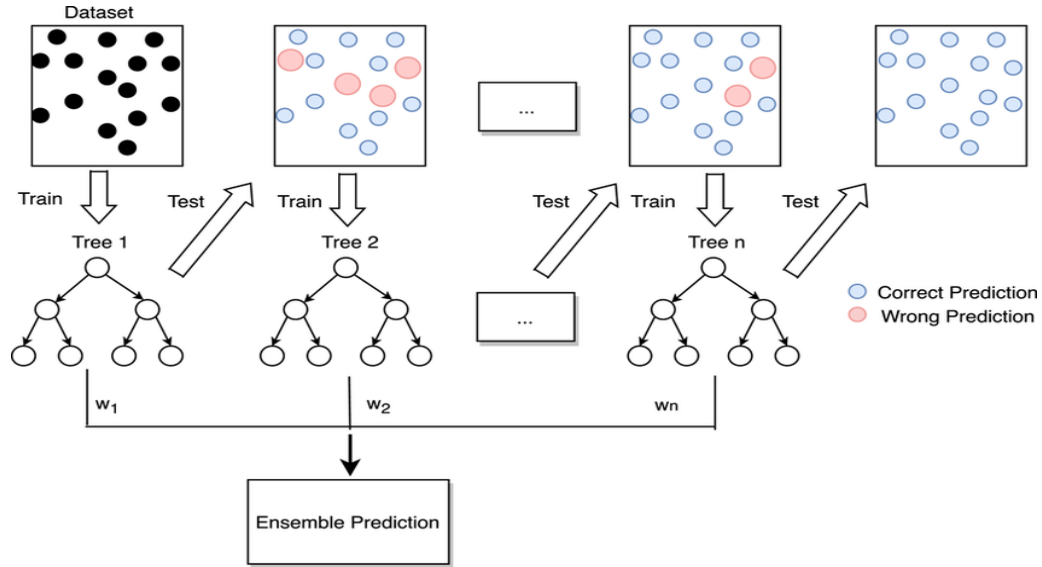


Figure 17. Diagram of gradient boosting machine learning method

The accuracy of the model is evaluated using the accuracy score metric from scikit-learn which compares the predicted values with the actual target values (Ytest).

In the provided output, the Gradient Boosting Classifier achieves an accuracy of 0.995454545455 (or 99.55%) on the test data. The classification report shows that the model has a precision, recall, and F1-score of 1.0 (or 100%) for most classes, indicating excellent performance in correctly classifying the different crop types

•**Accuracy:** The overall accuracy score on the test set indicates the proportion of correct predictions made by the model.

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

•**Precision:** This metric measures the proportion of true positives among the positive predictions made by the model for each class. And it's determined by dividing the number of true positive samples by the total number of positive samples predicted by the model:

$$\text{Precision} = \frac{(TP)}{(TP+FP)}$$

• **Recall:** this metric measure the proportion of actual positives that the model correctly identified for each class and it is calculated by dividing the number of correctly recommended or predicted crops or fertilizers by the total number of crops or fertilizers. It can be defined as:

$$\text{Recall} = \frac{(TP)}{(TP+FN)}$$

- **F1-score:** The F1-score is the harmonic mean of precision and recall, providing a balanced measure of the model's performance for each class.

$$\text{F1-score} = \frac{(2 * TP)}{(2 * TP + FP + FN)}$$

True Positive (TP): represents the number of abnormal samples correctly classified as exacerbation (correct detection)

False Positive (FP): represents the number of normal samples wrongly classified as attacks (incorrect detection)

True Negative (TN): represents the number of normal samples correctly classified (correct detection)

False Negative (FN): represents the number of abnormal samples wrongly classified as attacks (incorrect detection)

➤ Decision Tree

A decision tree is a machine-learning algorithm that resembles a flowchart-like structure. It represents decisions and their possible consequences in the form of a tree. Each internal node of the tree represents a feature or attribute, and each branch represents a decision rule based on that feature. The leaves of the tree represent the outcome or prediction.

Decision trees are used for classification and regression tasks, as they provide a clear and interpretable decision-making process. They are capable of handling both categorical and numerical data and are known for their ability to handle complex interactions between features. At each internal node of the decision tree, the algorithm evaluates the best feature to split the data based on a certain criterion. In this code, the criterion used is “entropy”, which measures the impurity of a node. The algorithm selects the feature that results in the greatest reduction in entropy when splitting the data.

The maximum depth parameter (max_depth) limits the depth of the decision tree, preventing over fitting. It controls the number of splits or levels that can be made in the tree

Once the decision tree is trained, it can be used to make predictions on new, unseen data by traversing the tree based on the feature values of the input data.

At each internal node, the decision tree checks the value of the corresponding feature and follows the appropriate branch until it reaches a leaf node, which represents the predicted class for that input.

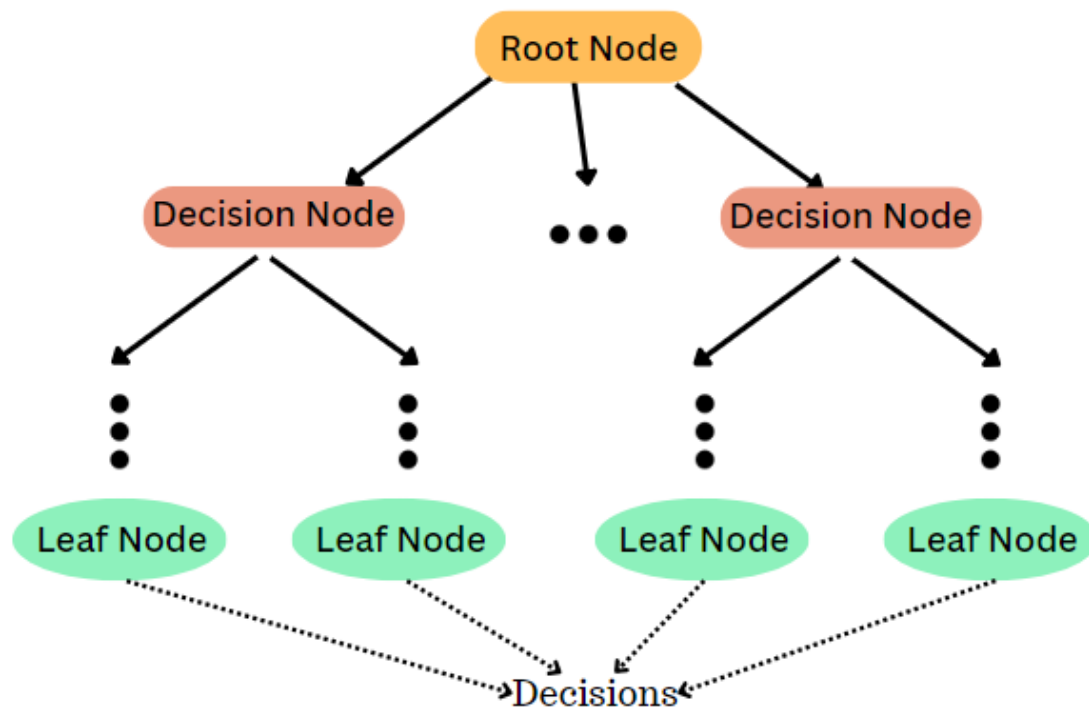


Figure 18. Diagram of Decision Tree

The model was trained and evaluated using the decision tree classifier from scikit-learn. The dataset was split into training and testing sets with a test size of 20% and a random state of 2. It was initialized with entropy as the criterion and a maximum depth of 5.

The model achieved an accuracy score of 90% on the test data. Additionally, a classification report was generated to provide insights into precision, recall, and F1-score for each crop class. The model achieved perfect precision and recall for several crops, while others exhibited lower scores.

Cross-validation was performed using 5-fold cross-validation, which yielded high accuracy scores ranging from 87% to 93%, indicating the model's generalization capability.

This model proved to be a reliable method for crop recommendation based on the dataset. It achieved a high accuracy score and demonstrated favorable precision and recall for most crops. However, continued refinement and exploration can enhance its performance for specific crops.

➤ Random Forest

Random Forest is an ensemble-learning algorithm that combines multiple decision trees to make predictions. Each decision tree is trained on a random subset of the data and features, and the final prediction is determined by aggregating the predictions of all the trees.

In our study, the Random Forest model is created using the Random Forest Classifier class from the `sklearn.ensemble` module. It is configured with 20 estimators, meaning that 20 decision trees will be created. The random state parameter is set to 0 for reproducibility.

The dataset is split into training and testing sets using the `train_test_split` function. The training set (80% of the data) is used to train the Random Forest model, while the testing set (20% of the data) is used to evaluate the model's performance.

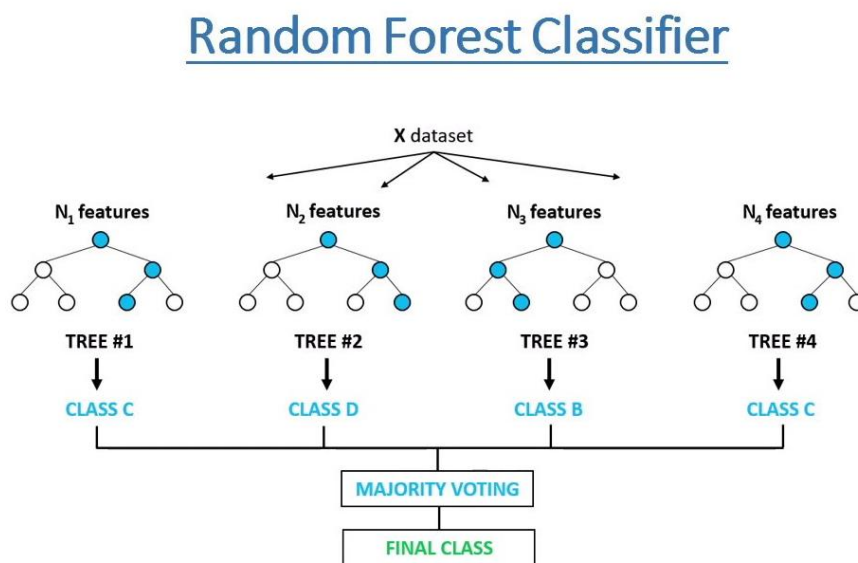


Figure 19. Random Forest classifier

Once the model is trained, predictions are made on the testing set using the `predict` function. The accuracy of the model is then calculated by comparing the predicted labels with the actual labels from the testing set.

Precision, recall, and F1-score are important metrics to evaluate the performance of a classification model.

The classification report function from `sklearn.metrics` is used to calculate precision, recall, and F1-score for each class (crop label) in the dataset. The results are displayed in the classification report as shown in the table below:

("RF's Accuracy is: ", 0.990909090909091)				
	precision	recall	f1-score	support
apple	1.00	1.00	1.00	13
banana	1.00	1.00	1.00	17
blackgram	0.94	1.00	0.97	16
chickpea	1.00	1.00	1.00	21
coconut	1.00	1.00	1.00	21
coffee	1.00	1.00	1.00	22
cotton	1.00	1.00	1.00	20
grapes	1.00	1.00	1.00	18
jute	0.90	1.00	0.95	28
kidneybeans	1.00	1.00	1.00	14
lentil	1.00	1.00	1.00	23
maize	1.00	1.00	1.00	21
mango	1.00	1.00	1.00	26
mothbeans	1.00	0.95	0.97	19
mungbean	1.00	1.00	1.00	24
muskmelon	1.00	1.00	1.00	23
orange	1.00	1.00	1.00	29
papaya	1.00	1.00	1.00	19
pigeonpeas	1.00	1.00	1.00	18
pomegranate	1.00	1.00	1.00	17
rice	1.00	0.81	0.90	16
watermelon	1.00	1.00	1.00	15
micro avg	0.99	0.99	0.99	440
macro avg	0.99	0.99	0.99	440
weighted avg	0.99	0.99	0.99	440

Table 9.Classification report table

By analyzing the classification report, we can observe the precision, recall, and F1-score for each crop label. Precision reflects how accurate the model is in predicting a specific crop label, while recall indicates the model's ability to correctly identify instances of a specific crop label. F1-score provides a balanced measure of precision and recall.

The Random Forest model achieved high accuracy, precision, and recall for most crop labels as it shown in the table (0.99) The accuracy score of approximately 99.1% (0.9909)indicates that the model correctly predicted the crop labels for the majority of the test instances.

The high precision and recall scores for most crop labels suggest that the Random Forest model effectively differentiated between different crops based on the given features. It successfully captured the patterns and relationships in the data to make accurate predictions.

The success of Random Forest in this code can be attributed to its ability to handle complex relationships between features and crop labels. By combining multiple decision trees, our model captured a wide range of patterns and made more robust predictions. Additionally, Random Forest is known for its resilience to noisy data, which contributes to its reliable performance.

Overall, the Random Forest algorithm demonstrated strong performance in this code, achieving high accuracy, precision, and recall in predicting crop labels based on the provided dataset. Its ability to handle complex relationships, handle noisy data, and leverage multiple decision trees makes it a powerful tool for our study.

➤ Logistic Regression

Logistic regression is a commonly used statistical model for binary classification problems. It aims to predict the probability of an event occurring based on input features.

The logistic regression algorithm works by fitting a sigmoid function to the training data, which maps the input features to a probability value between 0 and 1. The threshold of 0.5 is commonly used to classify instances into one of the two classes. If the predicted probability is greater than or equal to 0.5, the instance is classified as one class (ex: 'rice'), and if it is less than 0.5, it is classified as the other class (ex 'maize').

The sigmoid function, also known as the logistic function, plays a crucial role in mapping the predicted values to a probability between 0 and 1 and it's defined as:

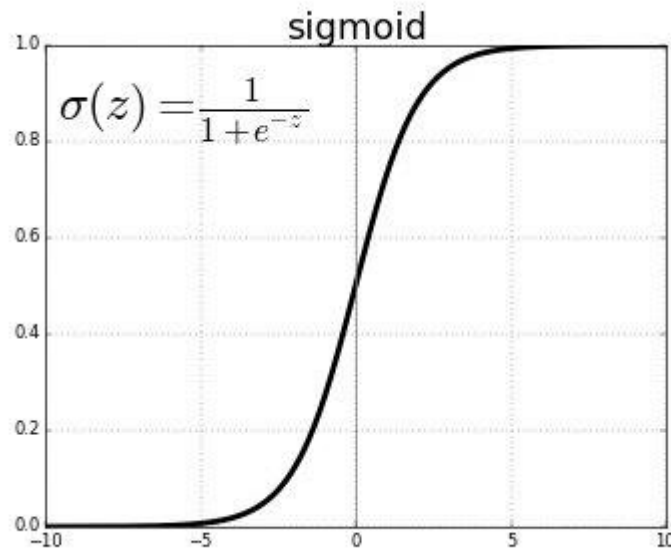


Figure 20.The sigmoid function

Where $\sigma(z)$ represents the output value or probability, and z is the input to the sigmoid function, which is a linear combination of the input features and their respective coefficients.

The sigmoid function has an S-shaped curve that starts from 0 at negative infinity and approaches 1 as z tends to positive infinity. It asymptotically approaches these values but never reaches them entirely. The function ensures that the predicted probabilities are bounded within the range of 0 and 1, which is essential for interpreting them as probabilities.

In logistic regression, the goal is to estimate the parameters (coefficients) that best fit the sigmoid function to the training data. This is done by using a process called maximum likelihood estimation. The model learns the optimal coefficients that maximize the likelihood of observing the given data.

The sigmoid function is a key component of logistic regression as it enables the model to make probabilistic predictions and define decision boundaries. By fitting the sigmoid function to the training data, logistic regression learns the optimal coefficients that maximize the likelihood of the observed data and allows us to predict the probabilities of new instances belonging to a specific class.

It is important to note that logistic regression assumes a linear relationship between the features and the log odds of the target variable. The model learns the optimal coefficients that maximize the likelihood of the observed data, which allows it to make predictions on new, unseen instances.

The accuracy of the logistic regression model in our study is approximately 94.3%. This means that the model correctly predicts the crop type for about 94.3% of the test data instances. The precision, recall, and F1-score are also calculated for each class in the target variable.

Looking at the classification report, we can observe that the model performs well in predicting most crop types with high precision, recall, and F1-scores close to 1 (100%) However, some crop types like black gram, jute, rice, and papaya have lower scores, indicating that the model has some difficulty in accurately classifying these crops.

In conclusion, the logistic regression model in this code demonstrates good overall performance in predicting crop types based on the given features. However, further analysis and improvements may be needed to address the challenges faced in accurately classifying certain crop types.

➤ K Nearest Neighbours

K-Nearest Neighbors (KNN) is a supervised machine-learning algorithm used for both classification and regression tasks. It works by finding the K nearest data points in the training set to a given test data point and making predictions based on the majority class or average value of the K nearest neighbors.

After loading the dataset and splitting it into training and testing sets. The KNN classifier is then created with a parameter `n_neighbors=9`, indicating that it will consider the 9 nearest neighbors when making predictions.

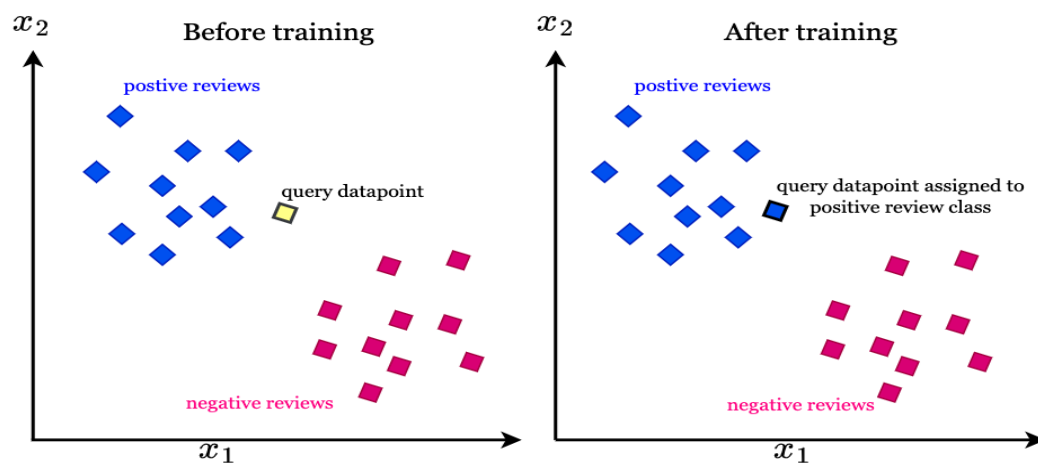


Figure 21. Visualizing the Knn algorithm behaviour

The classifier is trained on the training data using the `fit()` method, and predictions are made on the test data using the `predict()` method. The accuracy of the classifier is evaluated using the `accuracy score()` function, which compares the predicted crop types with the actual crop types in the test set.

In this example, the accuracy of the KNN classifier is printed as 0.9727, which means that it correctly predicted the crop type for approximately 97.27% of the test samples. This indicates that the model performs well in classifying the crops based on the given features.

Additionally, the classification report provides more detailed evaluation metrics such as precision, recall, and F1-score for each crop type. Precision measures the proportion of true positive predictions out of the total predicted positive instances. Recall measures the proportion of true positive predictions out of the actual positive instances. F1-score is the harmonic mean of precision and recall, providing a balanced measure of model performance.

Looking at the classification report, we can observe that most crop types have high precision, recall, and F1-scores close to 1. This indicates that the model is able to accurately classify the majority of crop types. However, some crop types like 'rice' have lower scores, indicating that the model struggles to correctly predict those classes.

The cross-validation scores further validate the performance of the KNN classifier. The scores range from approximately 0.968 to 0.979, which shows consistency in the model's performance across different folds of the dataset.

Overall, the KNN algorithm in this example demonstrates good accuracy and performance in predicting crop types based on the given features. However, it is important to note that the effectiveness of the algorithm may vary depending on the specific dataset and its characteristics.

➤ Gaussian Naïve Bayes

Naive Bayes is a probabilistic classifier that assumes independence between features. The Gaussian Naive Bayes variant assumes that the features follow a Gaussian distribution; it means that we assume the values of those features are normally distributed.

In statistics, the Gaussian distribution (also known as the normal distribution) is a continuous probability distribution that is symmetric and bell-shaped. It is characterized by two parameters: the mean (μ) and the standard deviation (σ).

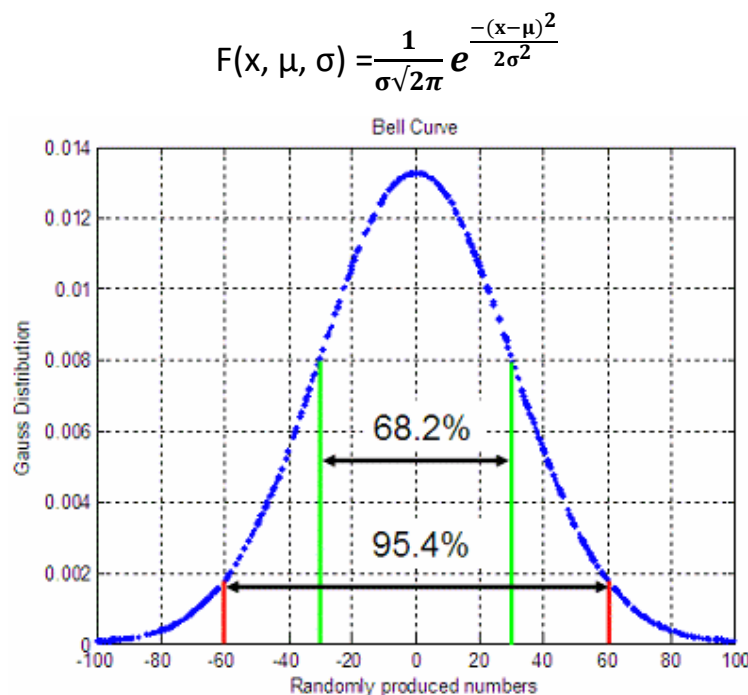


Figure 22. The normal or the Gaussian distribution function

A Gaussian distribution is often used to model real-world phenomena where the values tend to cluster around a central value and have a symmetric spread. A Gaussian distribution can approximate many natural phenomena, such as height, weight, and test scores.

In the context of the Gaussian Naive Bayes algorithm, we assume that each feature follows its own Gaussian distribution. This assumption simplifies the modeling process, as it allows us to estimate the parameters (mean and standard deviation) for each feature based on the training data. By assuming a Gaussian distribution for the features, the algorithm calculates the likelihood of observing a particular value of a feature given a specific class. It then combines these likelihoods with the prior probability of each class to calculate the posterior probability, which is used to make predictions.

The accuracy score of 0.9909 indicates that the model correctly predicted the crop label for 99.09% of the instances in the test data.

The classification report provides more detailed evaluation metrics for each class. Precision, recall, and F1-score are calculated for each crop label. Precision represents the proportion of true positives among the instances predicted as a particular class. Recall represents the proportion of true positives that were correctly predicted as a particular class. F1-score is the harmonic mean of precision and recall, providing a balanced measure of performance.

Looking at the classification report, we can see that for most crop labels, the precision, recall, and F1-score are all perfect (1.00), indicating that the model performed well in predicting those classes. However, for the “jute” class, the precision is slightly lower (0.88), which means that there were some instances predicted as “jute” that were actually other crops. This indicates a potential area for improvement.

Support Vector Machine

SVM is a supervised machine-learning algorithm used for classification and regression tasks. In classification, SVM finds an optimal hyper plane that separates different classes by maximizing the margin between the classes. It can handle both linear and non-linear classification tasks. The algorithm identifies support vectors, which are data points closest to the decision boundary, and uses them to construct the optimal hyperplane. SVM can also handle multi-class classification by using techniques like one-vs.-one or one-vs.-rest.

In our study, the SVM model is trained using the “fit” method and then used to make predictions on new data points. Overall, in this example, the SVM model exhibits poor performance with low accuracy, precision, and recall.

- The accuracy of the SVM model is reported as 0.1068, which means it correctly predicted only 10.68% of the instances in the test data.
- The precision is generally high for most classes, indicating a low false positive rate.
- The recall is quite low for most classes, indicating a high false negative rate.
- The macro-average f1-score is 0.14, indicating poor overall performance.
- The scores range from approximately 0.268 to 0.290, which suggests inconsistent performance across different folds.

In summary, the SVM model in our study shows poor performance with low accuracy and relatively low recall for most classes. This suggests that the model struggles to correctly predict positive instances, leading to a high false negative rate.

➤ Result and discussion

After analyzing the results, it is evident that the Gradient Boosting Classifier algorithm outperforms the other algorithms in this study. It achieves the highest accuracy of 99, 55% and a perfect precision, recall, and F1-score. This indicates that the model is highly accurate and reliable in predicting crop recommendations. Table 6 and Figure 29 also present the comparative analysis for all machine-learning techniques used in our study.

The Random Forest algorithm also demonstrates strong performance, with high accuracy (99, 09%) and consistent precision, recall, and F1-scores. It is a very powerful ensemble method.

On the other hand, the Decision Tree algorithm, while achieving decent accuracy (90%), falls behind the Gradient Boosting Classifier and the Random Forest and in terms of precision, recall, and F1-score. This suggests that it may not capture the complexity of the dataset as effectively as the ensemble methods.

K-Nearest Neighbors (KNN) also performs well, with high accuracy (97, 27%) and consistent precision, recall, and F1-scores. Logistic Regression and Gaussian Naive Bayes algorithms exhibit good performance as well, with respectable accuracy and consistent macro average precision, recall, and F1-scores.

However, the Support Vector Machine (SVM) algorithm shows significantly lower performance across all metrics, indicating that it may not be suitable for this particular crop recommendation system.

Considering the overall performance, the Gradient Boosting Classifier emerges as the most suitable choice for our crop recommendation system. It consistently achieves high accuracy and demonstrates reliable performance across multiple evaluation metrics.

In conclusion, based on the provided metrics, the Gradient Boosting Classifier algorithm appears to be the most appropriate choice for our study.

Algorithm	Accuracy	Precision	Recall	F1-score
Random forest	0,9909	0,99	0,99	0,99
Decision tree	0,9	0,84	0,88	0,85
KNN	0,9727	0,97	0,97	0,97
Gradient Boosting Classifier	0,9955	0,99	1	1
Logistic regression	0,9432	0,95	0,94	0,94
SVM	0,1068	0,66	0,13	0,14
Gaussian naive Bayes	0,9909	0,99	0,99	0,99

Table 6. Experimental results for ML Techniques under study

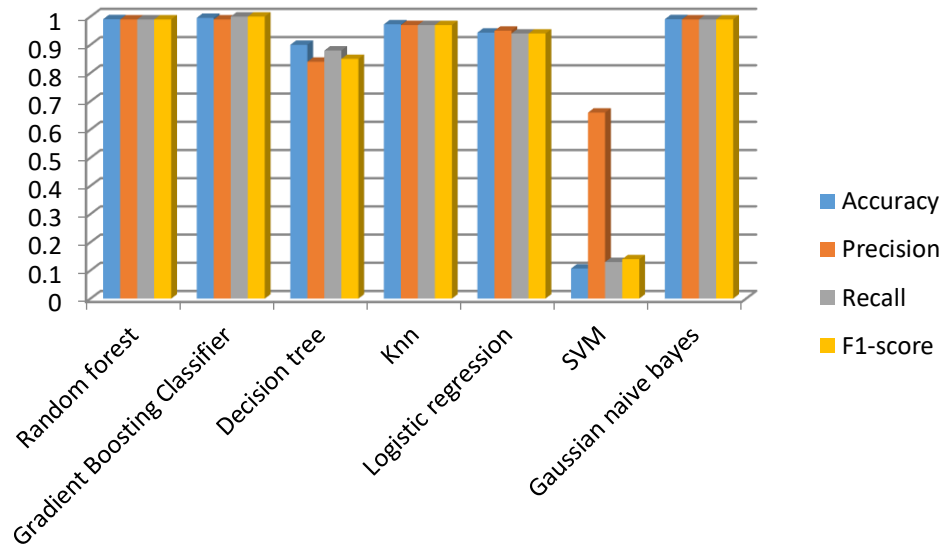


Figure 23. Comparison of ML Techniques under study

6. Predictions and Evaluation

In this final step, we selected the best-performing model (the Gradient Boosting Classifier) and retrain it on the entire dataset (combining the training and testing sets) then we use it to make predictions on new data instances. Therefore, to ensure accurate predictions, a diverse range of soil samples was collected from various locations as it shown in figures 30 .These samples were carefully selected to represent different soil types and agricultural regions. The collection process involved obtaining soil samples from different depths and ensuring proper labeling for identification purposes.

In addition, achieve this; we conducted a month internship at the technical Institute of Fruit Tree Arboriculture and Vine Demonstration Farm and Seedling Production of MOHAMMADIA. During this internship, we had the opportunity to collect various soil samples from their fields. These samples provided a practical basis to validate the effectiveness of our project. The data obtained from these samples were crucial in assessing the performance and reliability of the crop recommendation system in real-world conditions



Figure 24. Sampling procedure

Presentation of the Demonstration Station of the Technical Institute of Fruit Tree Arboriculture and Vine in Mohammadia

I – IDENTIFICATION

- Province: Mascara
- District: Mohammadia
- Address: Route d'Alger BP N° 52 Mohammadia 29400
- Origin of the lands: Formerly D.A.S. BAKHTI Mahi
- Date of establishment: Ministerial decree No. 228/SM dated 23/02/1989

II – DELIMITATION

- North: Mohammadia-Alger railway line
- South: National Road No. 4 (R.N. N° 4)
- East: Main drainage channel
- West: City of Mohammadia

III – SUPERFICIE TOTALE 35 ha, 47 ares 50 ca

- Terre cultivé : 27 has

IV – Service

- Experimental and Research Service
- Technical Support Service: Training, Outreach, and Technical Assistance
- Certified Basic Plant Material Production Service
- Study Service

V - TECHNICAL SUPERVISION

- 01 Director
- 02 Agricultural Engineers
- 01 Senior Technician

VI - LABOR FORCE

- 01 Assistant Mechanic
- 03 Casual Workers
- 02 Seasonal Workers
- 02 Guards

VII - INTERVENTION ZONE

- Oran
- Mostaganem
- Relizane
- Mascara
- Tiaret
- Saida
- Naama

VIII - VOCATION

The station initially focused on citrus farming.

Due to the drought in recent decades, it has shifted its activities to species that require less water in quantity and quality, such as rosaceae, olive trees, and fig trees.

IX - PROGRAM

1. Experimentation: Testing the behavior of olive, fig, and pear varieties.
2. Support for production:
 - Development of olive farming.
 - Development of techniques for managing perennial crops in rain-fed conditions.
 - Assistance to farms in implementing good agricultural practices.
- Assistance to table olive processors in improving manufacturing techniques.

X - OBSERVATIONS

The station is located near the city of Mohammadia. Part of its area is being encroached upon by the expansion of urban areas. A road serving as a bypass for the city has crossed the station widthwise, dividing it into two parts for over two years. This bypass was created with the agreement of the Department of Agriculture (D.S.A.) of Mascara and the Department of Forestry and Pasture (D.O.F.P.P.). This demonstration farm must be well protected with a solid fence to ensure its proper functioning and fulfill its mandates in its intervention zone.

1.2 Main Missions of the Structure

- Provide technical assistance for the implementation of fruit tree development programs.
- Contribute to training, skill development, and retraining of technical staff and trainers in areas related to fruit tree development.
- Transmit and enhance updated technical knowledge (agriculture, development agents).
- Establishment of demonstration and production orchards as part of technical assistance in fruit tree cultivation.
- Assess the behavior of new cultivars adapted to the intervention zone.

- Prospect and characterize the indigenous heritage found in the region and preserve it.
- Establish and produce plants in quality and quantity and support investors in this field.
- Introduce or develop technical innovations and disseminate them to fruit growers and vine growers.
- Participate with producers in its intervention area in fairs, exhibitions, and national and international competitions in its areas of expertise.

For this purpose, the Demonstration Farm has:

- A land potential consisting of orchards with a rich heritage
- A young technical staff
- proven experience in table olives, citrus farming, and processing
- The current composition of the orchards is dominated by olive and citrus varieties
- As part of experimentation activities, the demonstration farm has introduced varieties of indigenous and foreign fruit species for study in behavior orchards.
- These include Italian olive varieties
- Syrian olive varieties including:
 - Sigoise
 - Table olive collection
 - Citrus collection
 - Indigenous fig tree varieties
 - Foreign varieties of rosaceae: Nashi and pear.
- The study of varietal behavior allows for understanding the agronomic performance of each studied variety.

However, these strengths are counteracted by numerous weaknesses, including:

- Lack of a reliable water source.
- The station's land is within the Habra perimeter, which relies on dam releases depending on rainfall inputs, which are generally insufficient and irregular (these inadequacies may reflect the degradation of the conserved genetic heritage).
- Insufficient infrastructure for the development of the nursery for the production of authentic and quality plants.
- The station is highly coveted by industry and animals due to the lack of fencing to protect plantations and land.
- Insufficient agricultural labor force.

1.3 Current Production Level

The production of citrus fruits experienced a decline in the 1990s, dropping from 180,000 Qx in 2009 to 340,000 Qx in 2010. (Source: Department of Agriculture)

The varietal diversity of citrus fruits should provide fresh fruits for 8 months per year to cover national needs and generate exportable surpluses. However, unfortunately, this orchard is in a situation where it struggles to meet local consumption demands. The factors leading to this situation are numerous, including:

- Aging orchards
- Significant decrease in water resources
- Salinity
- Defective irrigation and drainage networks

1.4 Description of the Studied Trial Plot

This is a small olive orchard composed of a varietal mix of Italian and French origin. The plot covers an area of four ares. The trees are planted in a square planting pattern with a North-South exposure. The distances between the rows and within the rows are 7 meters, resulting in a density of 204 trees/ha. Other olive orchards (Syrian varieties), marabou, and a nursery surround the experimental orchard.

1.5 Plant Material

The plant varieties of olive trees planted at the study site. The varieties Nocellara Del Belice, Termite Di Bitteto, and Nocellara Messinese are of Italian origin, one variety is from France: Picholine, and the last one is from Algeria: Sigoise.

The properties of the studied olive varieties are represented in Table (8) below.

- **Nocellara Del Belice**

Origin: Italy

Vigor: Medium

Habit: Semi-spreading

Foliage density: Medium

Internode length: Medium

Intended use: Table olives

Growth scale: 6 months and 17 days



- **Termite Di Bitteto**

Origin: Italy

Vigor: Strong

Habit: Semi-spreading

Foliage density: Compact

Internode length: Medium

Intended use: Table olives

Growth scale: 7 months and 23 days



- **Picholine**

Origin: France

Vigor: Medium

Habit: Semi-spreading

Foliage density: Medium

Internode length: Medium

Intended use: Dual purpose

Growth scale: 7 months and 5 days



- **Nocellara Messinese**

Origin: Italy

Vigor: Strong

Habit: Semi-spreading

Foliage density: Compact

Internode length: Medium

Intended use: Dual purpose

Growth scale: 6 months and 10 days



- **Sigoise**

Origin: Algeria

Vigor: Strong

Habit: Erect

Foliage density: Medium

Internode length: Medium

Intended use: Dual purpose

Growth scale: 6 months



➤ **Nutrients necessary for plant growth**

For their development, the vast majority of plants require 16 essential nutrients from the air and soil around them. The following elements come from:

- From the air: Carbon (C) in the form of CO₂ (Carbon dioxide);
- From water: Hydrogen (H) and oxygen (O) in the form of water (H₂O);
- From the soil and mineral and organic fertilizers.
- Basic elements (macronutrients): Nitrogen (N), phosphorus (P), potassium (K).
- Secondary elements: Calcium (Ca), magnesium (Mg), sulfur (S), and trace elements: Iron (Fe), manganese (Mn), zinc (Zn), copper (Cu), boron (B), molybdenum (Mo), and chlorine (Cl).

Secondary elements and trace elements are usually found in sufficient quantities in the soil and should only be added if deficiencies are observed.

2.1. Secondary Elements

2.1.1 Calcium (Ca)

Calcium plays an extremely important role in the formation of plant tissues and facilitates better plant development. It enhances the resistance of plant tissues, promotes stronger stem structure, and supports normal root system growth.

2.1.2 Magnesium (Mg)

As a fundamental component of chlorophyll, magnesium enables the plant to undergo chemical transformations of nutrients and facilitates the assimilation and movement of phosphorus within the plant.

2.1.3 Sulfur (S)

Sulfur is essential for plant growth as it is involved in the composition of leaves and green parts of the plant. It is an important constituent of amino acids.

2.2 Macronutrients: N - P - K

2.2.1 Nitrogen (N)

Nitrogen is a major element for plant fertilization, taken up from the soil in either nitrate (NO₃⁻) or ammonium (NO₄⁺) form. It plays several roles in plant development. It drives vegetative growth and contributes to the vegetative development of all aerial parts of the plant, including leaves, stems, and seed formation, thereby enhancing yield.

CAUTION: Excess nitrogen has several negative effects on plants including: The development of foliage to the detriment of flowering and fruiting poor resistance to diseases; attracts aphids to young, very green shoots.

2.2.2 Phosphorus (P)

Phosphorus plays a role in strengthening plant resistance and contributes to root growth, fruiting, and seed setting.

2.2.3 Potassium (K)

Potassium promotes flowering and fruit development. It also enhances resistance to diseases and cold, reduces evapotranspiration, strengthens stem rigidity, and contributes to the formation of nutrient reserves (bulbs)

3- Signs of deficiencies in basic elements: N.P.K

↓ For Nitrogen (N)

- Yellowing leaves
- Reduced growth
- Smaller leaf size
- Shorter internodes
- Reduced fruit setting
- Insufficient protein content



↓ For phosphorus (P)

- Dark green, bronzed, or reddish-stained foliage
- Thin or malformed branches
- Sparse flowering
- Flower abortion
- Late fruit ripening



↓ **For potassium (K)**

- Brown necrosis at the tips, edges, and between the veins of the leaves
- Reduced distance between nodes on the stem
- Leaf edges sometimes curl upwards
- Plants susceptible to diseases
- Fruits are less sweet or flavorful
- Poor root vegetable preservation



3.1 Soil

Soil is the layer of earth worked by agricultural implements. It is characterized by its thickness (shallow or deep soil) and by the nature of the material; it is composed of (clayey, silty, sandy, etc.).

Soil fulfils three roles:

- 1. Physical:** supporting plant roots;
- 2. Chemical:** providing mineral nutrition to plants;
- 3. Biological:** acting as a living environment.

Soil is primarily composed of rocks, sand, silt, clay, limestone, and humus. Its physical properties (moisture, temperature, aeration) and fertility depend on its structure and texture (particle composition).

Soil is not an inert medium where a nutrient solution circulates. Soil colloids (clay, humus) regulate the nutrient content. This regulatory mechanism affects the content of cations (calcium, potassium, etc.) and the content of H^+ ions in the soil solution. Through soil analysis, the most suitable mode of exploitation can be determined.

4.1 Soil Analysis

Soil analysis is a valuable informational tool to aid decision-making. The method of soil sampling and the parameters of its analysis are determined based on three objectives: characterization, monitoring, and diagnosis.

-  **Characterization analysis:** This is initially performed to determine the general characteristics of a soil. It helps determine its physical, chemical, and biological fertility. It is one of the most comprehensive analyses.
-  **Monitoring analysis:** This is conducted periodically (every four years) at specific points in the plot. It determines the availability level of soil fertility elements that have been selected (such as monitoring the phosphorus fertility level of soil) to carry out maintenance or adjustments.
-  **Diagnostic analysis:** This is conducted when there is a difference in vegetation behavior within the same plot. The aim is to find an explanation by first examining crop profiles, soil analysis, and vegetation.

4.2 Soil Analysis Steps

There are two steps:

-  **Pre-laboratory:** This involves characterizing a plot using a plot sheet for the collection of one or more samples. Sampling is done at homogeneous units previously defined, with criteria such as soil color, clod size, abundance of stones, and depth.
-  **Post-laboratory:** This involves quality control of the information contained in the analysis report and its assessment (appreciations and recommendations) regarding specific elements such as particle size, total limestone, organic matter and biological activity, permeability, acidity, exchange capacity, assimilable phosphorus and nitrogen, etc.

5. Interest of Irrigation for Olive Trees

5.1 Rapid Entry into Production for Young Plantations

- Irrigation positively impacts the growth of aboveground parts (wood and leaves), but has no effect on root development. Therefore, the reduced growth of aboveground parts in trees grown under dry conditions may be attributed to the olive tree's adaptation to dry conditions.

5.2. Increase in Production in Mature Orchards

Impact of Irrigation on Growth and Alternation

During drought, olive trees reduce their photosynthetic activity, which is detrimental to shoot growth. Low water availability slows down shoot and bud growth.

Irrigation helps restore optimal growth, thus encouraging harvest in the following year. It has been demonstrated that olive trees can achieve vegetative growth during spring and early summer as long as soil moisture reserves are sufficient. Additionally, the role of water inputs in late August on the formation of autumn fruiting branches should be appreciated.

-Pruning and fertilization are certainly the most important operations in reducing alternation. However, irrigation also plays a significant role by allowing for a regular olive production, as there is a good correlation between the growth of the current year and the yield of the following year.

6. Sampling procedure



Figure 25. Taking samples

To obtain the necessary data for our crop recommendation system, we employed advanced sensors such as NPK sensor capable of measuring key parameters such as nitrogen (N), phosphorus (P), potassium (K) levels in the soil and pH, humidity, temperature, and rainfall installed in the study area. These sensors provided us with precise and real-time measurements, allowing us to gather comprehensive data for analysis.

Once the sensor readings were obtained as it shown in figure bellow the data was collected and compiled into our code. After training our crop recommendation system, we evaluated its performance by testing it on a separate set of soil samples. These samples were not included in the training dataset, ensuring an unbiased evaluation. The system utilized the measured parameters of these samples to predict the most suitable crop to be planted.

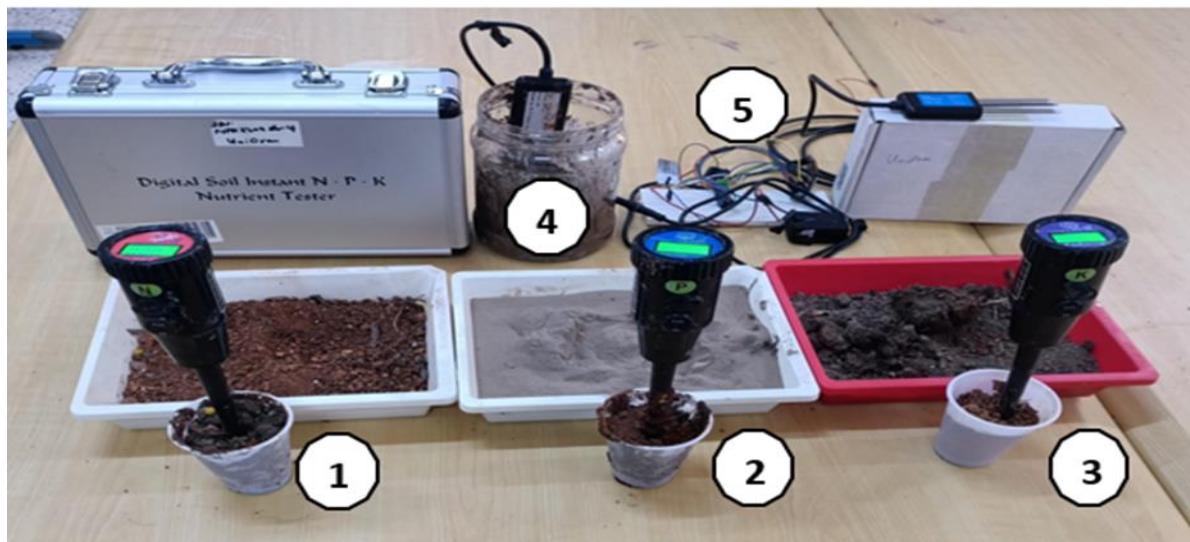


Figure 26. [1 : Sensor N (Barly Soil), 2 : Sensor P(soil with no crops), 3 : Sensor K(Olive soil) , 4 : (Olive soil with soil sensor) , 5 : Soil sensor]



Figure 32. NPK monitoring

The results obtained from our crop recommendation system demonstrated promising accuracy and reliability. The predictions made by our model aligned perfectly with the actual crops already cultivated in those soil samples. This indicates that our machine-learning algorithm and code effectively analyzed the soil parameters and provided very suitable crop recommendations as it shown in the figure below:

Making a prediction

```
N = 20
P = 17
K = 15
temperature = 22.3
humidity = 92.12
ph = 6.95
rainfall = 110.33

data = np.array([[N,P, K, temperature, humidity,ph, rainfall]])
prediction = grad.predict(data)
print(prediction)

['orange']
```

The successful implementation of our crop recommendation system highlights its potential to revolutionize agricultural practices. By leveraging machine learning algorithms and advanced sensor technology, farmers can make informed decisions regarding crop selection, leading to improved yields and resource utilization. One of the key advantages of our system is its ability to consider multiple parameters simultaneously. By analyzing the combined effect of nitrogen, phosphorus, potassium, pH, humidity, temperature, and rainfall. Our system provides a holistic approach to crop recommendation. This comprehensive analysis enables farmers to optimize crop selection based on the specific requirements of their soil and environmental conditions. Furthermore, our system can be easily scaled and adapted to different regions and soil types. By collecting soil samples and measuring the relevant parameters, farmers can train the system to provide tailored recommendations for their specific agricultural contexts. This flexibility ensures that farmers worldwide, regardless of their location or soil characteristics, can utilize our system.

7. The economic study of our project

7.1. Cost Analysis of our Project

Here is a breakdown of the costs for each components in our project (21)

RS485 NPK Sensor 	18.16 \$
MAX485 TTL to RS485 Converter Module (OT5003) 	2.39 \$
Arduino Mega 	21.64 \$
USB Cable 	2.12 \$
Jumper Cable 	1.57 \$
Total Cost	45.88\$

7.2. Comparison with Digital Soil NPK Tester



The commercially available Digital Soil NPK Tester costs 1278.84 \$ (21)

Cost Difference: 1278.84 \$ - 45.88 \$ = 1232.96 \$

Our project costs only about 3.6% of the price of the commercial Digital Soil NPK Testers, making it a highly cost-effective solution.

7.3. Economic Efficiency

7.3.1 Cost Efficiency

- Our project costs 45.88 \$, which is approximately 96% or 27 times cheaper than the commercial Digital Soil NPK Tester (1278.84 \$).

7.3.2 Performance

- Both systems perform the same primary function of measuring soil N, P, K levels.

7.3.3 Cost-Benefit Analysis

- Our Project

- ✓ Significantly lower cost.
- ✓ DIY (do it yourself system means building and repairing it on your own) and customizable.
- ✓ Accessible for small-scale farmers, researchers, and educational purposes.
- ✓ Commercial Tester
- ✓ High initial investment.
- ✓ Potential for additional features or higher precision.

- ✓ Better suited for large-scale agricultural enterprises or professional use where long-term reliability is critical.

8. Additional Considerations

Scalability: our project can be scaled or modified according to specific needs, such as adding wireless communication or automated data logging.

Maintenance and Support: DIY solutions might require more hands-on maintenance and troubleshooting compared to commercial products with dedicated support.

Accuracy and Precision: While our project works perfectly, it might be beneficial to compare the accuracy and precision of measurements against the commercial tester to validate the performance thoroughly.

Results: In summary, our project demonstrates a highly cost-effective solution for soil nutrient measurement. By utilizing affordable components and successfully integrating them, we have created a system that achieves the same core functionality as a significantly more expensive commercial product. This makes our project a viable option for budget-conscious users or those seeking customizable and scalable solutions. Our low-cost project not only works effectively but also provides an accessible alternative to expensive commercial options, highlighting the potential for innovation and cost savings in agricultural technology.

Conclusion

In conclusion, our project successfully developed a crop recommendation system using machine-learning algorithms. By collecting soil samples, utilizing sensors to measure key parameters, and training a model on the collected dataset, we were able to predict the most suitable crops for a given set of soil conditions. The alignment of our predictions with the actual crops grown in the soil samples validates the effectiveness of our system. This project holds great potential for optimizing crop selection and enhancing agricultural productivity.

Perspectives

Our IoT Low-Cost Fertilizer for Crop Recommendation using Machine Learning,"project introduces a system that leverages IoT sensors and machine learning algorithms to optimize crop production and resource management, the system provides tailored recommendations that help farmers make informed decisions about crop selection and fertilizer application. This approach not only aims to increase crop yields but also promotes sustainable agricultural practices.

Despite the promising results demonstrated through field experiments, there is always room for improvement. Enhancing the accuracy and variety of sensors, improving data transmission and storage, adopting advanced machine learning techniques, and focusing on user interface and scalability are essential steps to further develop this low-cost fertilizer recommendation system. By addressing these areas, the system can better meet the diverse needs of farmers and adapt to varying agricultural contexts, ultimately contributing to enhanced food security and sustainable farming practices.

And these are our perspectives to improve our project in the future:

1. Enhanced Sensor Accuracy and Variety:

***Upgrade Sensors:** Consider using more advanced sensors with higher accuracy and additional capabilities (for example sensors for detecting more soil nutrients or environmental conditions).

*** Expand Sensor Types :** Introduce new types of sensors for monitoring additional parameters such as soil pH, salinity, and pest presence.

2. Improved Data Transmission and Storage:

***Reliable Communication Protocols:** Implement more robust communication protocols (LoRaWAN, NB-IoT) to ensure reliable data transmission even in remote areas.

***Cloud Integration :** Enhance cloud computing integration for better data storage, processing, and accessibility. This can also facilitate the use of more sophisticated machine learning models.

3. Advanced Machine Learning Techniques:

***Algorithm Optimization:** Experiment with different machine learning models (ensemble methods, deep learning) to improve the accuracy of predictions and recommendations.

***Real-Time Processing:** Implement real-time data processing capabilities to provide immediate feedback and recommendations to farmers.

4. User Interface and Usability:

***Mobile Application:** Develop a user-friendly mobile application that provides farmers with easy access to real-time data, recommendations, and alerts.

5. Energy Efficiency:

***Low-Power Sensors:** Use low-power sensors and optimize the power consumption of the system to extend the battery life and reduce maintenance needs.

***Solar Power:** Consider integrating solar power solutions to make the system more sustainable and suitable for off-grid areas.

6. Collaboration and Feedback:

***Farmer Feedback:** Establish a feedback loop with farmers to gather insights on the system's performance and areas for improvement.

***Collaborative Research:** Collaborate with agricultural research institutions to continuously refine the algorithms and incorporate the latest scientific findings

The proposed IoT and machine learning-based system for crop recommendation and fertilizer application offers significant potential for improving agricultural productivity and sustainability. By enhancing sensor accuracy, improving data transmission, leveraging advanced machine learning techniques, and focusing on user experience and system scalability, the system can be further optimized to better meet the needs of farmers and adapt to various agricultural contexts. Continuous feedback and collaboration with stakeholders will be crucial in ensuring the system's ongoing development and success.

General conclusion

This project has successfully demonstrated the powerful integration of IoT technology and machine learning to revolutionize soil nutrient monitoring and crop recommendation systems for the agricultural sector. The developed IoT device, equipped with an array of sensors, has the capability to continuously gather real-time data on critical soil parameters such as nutrient levels, moisture, temperature, and humidity.

By leveraging advanced machine learning algorithms, the system is able to analyze this data and generate tailored recommendations for farmers, advising on optimal crop choices, appropriate fertilizer types, and precise application rates. This data-driven approach addresses the key challenges faced by farmers in managing soil health and optimizing crop yields, which are often hindered by the limitations of traditional soil analysis methods and the complexities of variable environmental conditions.

The implementation and field-testing of the proposed solution have demonstrated its effectiveness in improving agricultural productivity, resource utilization, and environmental sustainability. Farmers can now make informed decisions based on the system's recommendations, leading to enhanced crop yields, reduced waste of fertilizers and other inputs, and more sustainable farming practices.

The success of this project paves the way for wider adoption of IoT-ML systems in the agricultural domain, transforming traditional farming methods into precision-driven, data-informed approaches. As the world grapples with the challenges of food security and environmental preservation, this research offers a promising pathway to address these critical issues through the synergistic power of emerging technologies.

Overall, this thesis has made a significant contribution to the field of smart agriculture, showcasing the potential of IoT and machine learning to revolutionize the way we manage and optimize crop production, ultimately supporting the global effort towards sustainable and efficient food systems.

References

1. An IoT Low-Cost Smart Farming for Enhancing Irrigation Efficiency of Smallholders Farmers Amine Dahane^{1, 2,3} · Rabaie Benameur^{2,3} · Bouabdellah Kechar². [Online]
2. <https://www.researchgate.net/publication/330425585> Internet of Things IoT : Definition, Charac Wireless Personal Communications A Comprehensive Review on the
3. Application of Internet of Thing (IoT) in Smart Agriculture <https://doi.org/10.1007/s11277-021-08970-7>teristics, Architecture, Enabling Technologies, Application & Future Challenges
4. Making, 10. Internet-of-Things (IoT)-Based Smart Agriculture: Toward. [Online]
5. García, L., Parra, L., Jimenez, J. M., Lloret, J., & Lorenz, P. (2020). IoT-based. [Online]
6. A comprehensive review on the application of internet of thing (IoT) in smart agriculture. Srivastava, A., & Das, D. K. (2021). Wireless Personal Communications. <https://doi.org/10.1007/s11277-021-08970-7>. [Online]
7. S. J. Ramson, W. D. León-Salas, Z. Brecheisen, E. J. Foster, C. T. Johnston, D. G. Schulze, T. Filley, R. Rahimi, M. J. C. V. Soto, J. A. L. Bolivar, et al. A self-powered, real-time, lorawan iot-based soil health monitoring system. IEEE Internet of T. [Online]
8. T. Blesslin Sheeba, 1L. D. Vijay And, 2Gunaselvi Manohar, 3Saravana Selvan, 4C. Bazil Wilfred, 5K. Muthukumar, 6S. Padmavathy, 7P. Ramesh Kumar, 8and Belete Tessema Asfaw , Machine Learning Algorithm for Soil Analysis and Classification of Micronutrients in I. [Online]
9. Kalaiselvi Bakthavatchalam 1, Balaguru Karthik 1, Vijayan Thiruvengadam 1, Sriram Muthal 2, Deepa Jose 3, Ketan Kotecha 4, and Vijayakumar Varadarajan 5, IoT Framework for Measurement and Precision Agriculture: Predicting the Crop Using Machine Learning Algorithms.. <https://doi.org/10.3390/technologies10010013>
- Data set link: <https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset/discussion>
10. Murali Krishna Senapaty 1, Abhishek Ray 1 and Neelamadhab Padhy IoT-Enabled Soil Nutrient Analysis and Crop Recommendation Model for Precision Agriculture Computers 2023, 12(3), 61 and <https://doi.org/10.3390/computers12030061>. [Online]
11. Amna Ikram, 1Waqar Aslam, 2Roza Hikmat Hama Aziz, 3Fazal Noor, 4Ghulam Ali Mallah, 5Sunnia Ikram, 2Muhammad Saeed Ahmad, 1Ako Muhammad Abdullah 6and Insaf Ullah Crop Yield Maximization Using an IoT-Based Smart Decision , journal of sensors Published 17 May 20. [Online]
12. Navod Neranjan Thilakarathne, Muhammad Saifullah Abu Bakar, Emeroly Lariffion Abas and Hayati Yassin. A Cloud Enabled Crop Recommendation Platform for Machine Learning-Driven Precision Farming Sensors 2022, 22(16), 6299 and <https://doi.org/10.3390/s22166299>.

Data set link: <https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset/discussion>

13. Rohit Kumar Rajak, Ankit Pawar, Mitalee Pendke , Pooja Shinde, Suresh Rathod, Avinash Devare ,Crop Recommendation System to Maximize Crop Yield using Machine Learning Technique International Research Journal of Engineering and Technology (IRJET) Volum.

14. Arafat Ahmed Khan, Muhammed Faheem, Rab Nawaz Bashir, Chitapong Wechataisong, and Muhammed Zahid Abbas Internet of Things (IoT) Assisted Context Aware Fertilizer Recommendation IEEE internet of things journal
Doi:10.1109/ACCESS.2022.3228160. [Online]

Data set link : <https://github.com/ameyaagsilver/Fertiliser-Recommendation-System>

15. Maaz Patel , Anagha Rane
https://www.researchgate.net/publication/37005671_Crop_Recommendation_System
Doi:10.31224/2959

Data set link : <https://www.kaggle.com/datasets/atharvaingle/croprecommendation-dataset>

16. S. V. Gaikwad, A. D. Vibhute, K. V. Kale, and S. C. Mehrotra. An innovative iot based system for precision farming. Computers and Electronics in Agriculture, 187,(2021) 106291

17. <https://www.jetir.org/papers/JETIR2305686.pdf>. [Online]

18. https://www.mdpi.com/2227-7080/10/1/13/review_report. [Online]

19. <https://diyprojectslab.com/guide-for-npk-soil-sensor-with-arduino-tutorial/>. [Online]

20. ZD, [https://fr.aliexpress.com/item/1005005282971941.html?gatewayAdapt=glo2fr&testeur de sol NPK testeur instantané de sol testeur de nutriments azote phosphore Potassium](https://fr.aliexpress.com/item/1005005282971941.html?gatewayAdapt=glo2fr&testeur%20de%20sol%20NPK%20testeur%20instantan%C3%A9%20de%20sol%20testeur%20de%20nutriments%20azote%20phosphore%20Potassium). [Online]

21. https://fr.aliexpress.com/item/1005003169937270.html?src=google&src=google&albch=s hopping&acnt=548-301-0399&slnk=&plac=&mtctp=&albbt=Google_7_s

22. An innovative low-cost IoT-based asthma exacerbation prediction system using machine learning

