TOPIC: “Predicting Suitable Crops Based on Soil Properties Using Machine Learning: A Case Study of Major Crops in Nigeria”

CHAPTER 1

INTRODUCTION

* 1. BACKGROUND TO THE STUDY

Agriculture plays a crucial role in strengthening Nigeria’s economy. It contributes significantly to employment generation, poverty alleviation in rural areas, and the enhancement of food security. Agricultural activities not only provide income for farmers but also stimulate demand for related goods and services, creating a multiplier effect that drives growth across other sectors of the economy (Gomina et al., 2024)[4]. Furthermore, increasing crop production is essential for ensuring food availability and reducing hunger.

According to the Food and Agriculture Organization (FAO, 2025)[3], an estimated 30.6 million people across 26 states and the Federal Capital Territory (FCT) are projected to face acute food and nutrition insecurity—classified as Crisis (CH Phase 3) or worse—during the June to August 2025 lean season. When the production of major food crops fails to increase with the population, the demand for food cannot be met.

One way to increase agricultural production is crop selection. Crop selection deals with choosing which crop best suits the conditions, such as soil properties, weather, environmental factors, and temperature. Farmers understand the importance of crop selection which is why they employ their instinct and past experiences when picking these crops, unfortunately, they are not able to assess the effects of both natural and human events such as land degradation and climate change on soil properties which results in fluctuation of crop suitability on that land (Vedaste & Enan, 2025)[2]. Parameters such as nitrogen, phosphorus, potassium, and pH levels influence crop yield, yet many farmers lack access to this information or the tools to analyze it.

(Nde et al., 2024)[5] divided crop selection systems into two main types: rule-based and artificial intelligence (AI)-based systems. A branch of artificial intelligence called machine learning (ML) allows systems to automatically learn and improve from data without the need for explicit programming. Knowledge is extracted from the data to build a framework for making intelligent decisions and informed predictions (Sharma et al., 2020)[6]. By analyzing soil and environment data, ML models can predict the most suitable crops for a given land region.

This project aims to build a machine learning system that recommends crops from soil property data and environmental data, which will enhance crop selection accuracy, improve crop yield, and optimize fertilizer use -- supporting the efforts to tackle food insecurity in Nigeria.

* 1. STATEMENT OF THE PROBLEM

Crop prediction has traditionally relied on historical data and the expertise of experienced farmers. However, due to climate change and other environmental factors, such predictions have become increasingly inaccurate. Climate change poses both direct and indirect challenges to crop production, influencing yields and requiring strategic interventions from agricultural departments to enhance productivity (Ayoola et al., 2024).

Rainfall, temperature, and humidity have become difficult to predict; these are factors that influence crop production.

However, assessing each land and compiling the necessary data would be impossible, seeing as each farmland possesses unique soil characteristics that make manual soil analysis highly impractical. This stresses the need for a system that can predict suitable crops even with changing climate levels and unique soil properties.

In Nigeria, the adoption of technological tools such as artificial intelligence has not been successful. This has led to the waste of resources like fertilizers, financial losses, low crop yield, and poor crop quality. Utilizing tools that will address these issues by accurately predicting the crops that would grow best in specific conditions is essential. The goal of this study is to provide a machine learning-based system in Nigeria that offers precise analysis of soil and environmental data to recommend crops most compatible with the specific soil properties of a given farmland. This approach has the potential to increase efficiency and productivity, enhance the use of resources such as fertilizers, and support sustainable agricultural practices in Nigeria.

* 1. AIM AND OBJECTIVES

This study aims to develop a machine learning model that accurately predicts suitable crops based on soil properties (e.g., Nitrogen, Phosphorus, Potassium) and environmental factors (e.g., rainfall, temperature).

Objectives:

1. To gather, understand, and preprocess soil property data and crop labels from publicly available datasets.

2. To analyze the relationships between soil properties, environmental factors, and crop compatibility using visual tools such as plots, diagrams, and correlation matrices.

3. To train a machine learning model capable of predicting the most suitable crops for given soil and climate conditions.

4. To evaluate the accuracy and performance of the model using an external test dataset.

5. To simulate real-world testing by collecting soil condition input and generating crop recommendations using the model.

* 1. SCOPE OF THE PROJECT

This project focuses on developing a machine learning model that predicts and recommends suitable crops from given soil properties and some environmental factors.

The model will be trained using publicly available datasets with a focus on 5–7 major crops commonly grown in Nigeria, such as maize, rice, groundnut, cassava, and yam. From the dataset, the features will include nitrogen (N), phosphorus (P), potassium (K), pH level, and rainfall. If available, other environmental data such as temperature and humidity may be included in the training and testing of the model. An external dataset will be used to test the model to evaluate the model's applicability in real-world situations.

Additionally, a web-based interface will be developed to enhance usability and allow users to input soil and environmental data and receive crop recommendations from the trained model.

The project does not cover pest or disease prediction, crop yield quantity, or region-specific crop calendars. It is intended as a general-purpose crop recommendation tool based on soil compatibility, with potential for future improvement and expansion.

* 1. PROJECT RISK

|  |  |  |  |
| --- | --- | --- | --- |
| RISK | DESCRIPTION | MITIGATION | IMPACT |
| Lack of access to datasets specific to Nigerian soil and crop | Datasets from public or private Nigerian farms are not made public and agricultural agencies may not respond | Make use of publicly available datasets and use synthetic test data that simulate Nigerian soil and crop data. | High |
| Incomplete and noisy data | Incomplete and noisy data lead to errors in the models output. | Applying preprocessing techniques such as data cleaning, transformation, validation, and normalization is critical for improving the data quality. | Medium |
| Overfitting | The model becomes too closely tailored to the data used for training therefore its accuracy on unseen data is lower. | Deploy large and diverse data for training. | High |
| Limited real world testing | Lack of access to live data to test with | Medium | Manually create test data that repliacte soil profiles |
| Time constraints for developing both the machine model and integrating it into a website. | Training and testing models coupled with developing a website and ensuring documentation is correct all take time | High | Apply Agile methodology by breaking down tasks into flexible phases. |

Table 1.1 - Risk Assessment

* 1. SWOT ANALYSIS

|  |  |
| --- | --- |
| STRENGTHS | WEAKNESSES |
| - it is relevant in agriculture today and solves issues that stem from inaccurate crop yield prediction.  - Uses public dataset for training  - Improves yield which addresses food insecurity happening in Nigeria.  - Focus on major Nigerian crops which cuts training time down and is relevant to Nigerian farmers.  - It is scalable.  - User friendly interface to accommodate non-technical users. | - it does not include all the features that influence crop yield (e.g. pests)  - Nigerian specific soil datasets may not be available to the public.  - Limited real world testing due to not having access to farm data |
| OPPORTUNITIES | THREATS |
| - Can be extended to include more crops.  - if adopted, it will increase users trust for technological advances in agriculture.  - Future integration with IoT (Internet of Things) in agriculture. | - Rapid changes in climate will cause the model to be inaccurate therefore, needing constant updates  - Users may be slow to adopt if they don’t understand and trust technology  - Technical errors |

Table 1.2 - SWOT Analysis

* 1. SIGNIFICANCE OF THE PROJECT

The purpose of the machine learning-based system that predicts suitable crops from a set of soil and environmental data is to reduce the likelihood of selecting the wrong crops to cultivate on a piece of land. Incorrect selection of crops reveals a gap in knowledge related to the soil and weather, and how they affect the cultivation of those crops. When these issues are met with inadequate solutions, such as applying excessive amounts of fertilizers-- without identifying what properties are deficient -- it can lead to further degradation of the farmland. Since soil composition varies in terms of its properties, and without accurate analysis of the soil, choosing a crop will be incredibly difficult. Coupling the differences in soil and the irreversible effects of climate change, this system is capable of improving crop production.

It shows the correlation between the land chosen for cultivation and the crop to be produced on it. Also, it illustrates the importance of soil properties: Nitrogen, Potassium, Phosphorus, pH levels and how a strategic combination of these elements tailored to the crop enhances its quality.

Beyond the practical benefits, the farmers’ knowledge of their crops greatly improves, maximizing profits and production. Proper implementation of this project also has the potential to make the local farmers more receptive to technological advancements in Nigerian agriculture.

* 1. ORGANIZATION OF THE PROJECT

The structure of the project is as follows:

Chapter 1: Introduction

Overview of the project, including background, objectives, problem statement, and risk assessment.

Chapter 2: Literature Review

Reviews past research on crop prediction systems, machine learning in agriculture, and soil data analysis, highlighting gaps that this project aims to fill.

Chapter 3: Methodology

Describes the tools, datasets, and techniques used to develop and evaluate the machine learning model, including system architecture and workflow.

Chapter 4: Implementation and Testing

Details the model development, web interface, and performance testing using both training and external data.

Chapter 5: Conclusion

Summarizes the project, key findings, limitations, and suggestions for future improvement.

Reference:

1. *(PDF) 2024) Data-Driven Framework for Crop Categorization using Random Forest-Based Approach for Precision Farming Optimization*. Available from: <https://www.researchgate.net/publication/381281441_2024_Data-Driven_Framework_for_Crop_Categorization_using_Random_Forest-Based_Approach_for_Precision_Farming_Optimization> [accessed Jun 30 2025].
2. *(PDF) Applicability of Machine Learning and Internet of Things-Based for Crop Selection*. Available from: <https://www.researchgate.net/publication/392505168_Applicability_of_Machine_Learning_and_Internet_of_Things-Based_for_Crop_Selection> [accessed Jul 05 2025].
3. <https://www.fao.org/nigeria/news/detail-events/en/c/1735060/#:~:text=The%20findings%20show%20that%2030.6,June%E2%80%93August%202025%20lean%20season.>
4. *(PDF) Analyzing agricultural funding, poverty alleviation, and economic growth in Nigeria: A Focus on the Abuja Federal Ministry of Agriculture*. Available from: <https://www.researchgate.net/publication/383561341_Analyzing_agricultural_funding_poverty_alleviation_and_economic_growth_in_Nigeria_A_Focus_on_the_Abuja_Federal_Ministry_of_Agriculture> [accessed Jul 05 2025].

*5.(PDF) Crop Selection*. Available from: <https://www.researchgate.net/publication/384997215_Crop_Selection> [accessed Jul 05 2025].

*6.(PDF) Machine Learning Applications for Precision Agriculture: A Comprehensive Review*. Available from: <https://www.researchgate.net/publication/348141354_Machine_Learning_Applications_for_Precision_Agriculture_A_Comprehensive_Review> [accessed Jul 05 2025].

[P.S: I will edit references later also major crops in scope]