FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 9 No. 9, September, 2025, pp 1 – 7
DOI: https://doi.org/10.33003/fjs-2025-0909-2798

# AN INTELLIGENT FARMLAND ADVISORY MODEL FOR BEST CROPPING AGRICULTURAL PRACTICES USING MACHINE LEARNING

**[1]Mohammed Umaru and [2]Gregory Wajiga**

[1]Department of Computer Science, Federal College of Education, Yola Adamawa State
[2]Department of Computer Science, Modibbo Adama University, Yola, Adamawa State

*Corresponding authors' email: umarsamu00@gmail.com Phone +2348033028625

## ABSTRACT

Agricultural practice is the major business activity of the well-being of people. Crop production is the important factor in agricultural practices, likewise soil pattern determines the suitability of crop to be cultivated. The cropping system in Nigeria faces challenges such as poor access to modern technology, climate change, inadequate digital infrastructure, and insufficient agricultural data for optimized crop management and productivity. This study aimed at developing an intelligent farmland advisory model for prediction of soil fertility for maize, guinea corn and millet crops using machine learning. The dataset used was 1550 data records. Dataset was pre-processed, cleaned, and divided into two; 80% for training while 20% for testing the model. Decision Tree technique and python Jupita notebook were used for classification and building the model. The model was evaluated using confusion matrix, precision, F1 score and recall. Precision and recall results were above 80% for both "Not fertile" and "Partially Fertile" in all the soil fertility tests for the three crops used. The performance of the model implies that it can be used to test soil fertility. The study recommended the implementation of the model to assist farmers in making informed decisions about crop cultivation based on soil suitability.

**Keywords**: Intelligence, Model, Prediction, Farmland, Cropping, Agriculture, Practices, Machine Learning

## INTRODUCTION

Traditionally, agricultural services have been dependent on human expertise and local knowledge, primarily provided by agricultural extension officers who visit farms and offer guidance based on their skills and experience. However, this method has proven to be inadequate and inefficient, as farmers often struggle to access these services when they need them most. This delay has had a negative impact on agricultural production, particularly in rural areas where farming is the primary occupation, leading to significant losses.

In recent years, many people have ventured into agriculture with minimal knowledge, skills, or experience, aiming to produce crops and livestock for domestic use. Unfortunately, due to their reliance on traditional farming methods, many have incurred significant losses, especially within their first year, causing them to abandon farming altogether.

The demand for food crops continues to rise, while their supply is steadily decreasing. Factors such as climate change, soil degradation, and a growing population have increased the need for higher-quality foods, making it difficult to balance food demand and supply (Chakraborty, 2022). Agriculture is facing numerous challenges, including shifting weather patterns, pests, diseases, a lack of reliable extension services, and market fluctuations. These challenges are particularly acute for farmers in remote or underserved areas who lack timely and accurate agricultural information. Traditional advisory services are often limited in scope and effectiveness, hindering farmers from adopting best practices to meet the growing demand for food. The United Nations Food and Agriculture Organization (FAO) projects that food demand will increase by 70% by 2050 (Arsene, 2021; Oshodi, 2023; FAO, 2009).

Vadlamudi (2019) notes that breakthroughs in artificial intelligence and climate-smart agriculture have significantly enhanced crop yields. In response to these developments, an intelligent farmland advisory model has been designed to help farmers improve their practices and increase crop yields.

Automated farming, also known as precision farming or smart farming, involves the use of various technological advancements to automate farming processes. These technologies aim to optimize food production, improve crop quality, and reduce the need for labor and time-intensive activities (Jha et al., 2019). Durai et al. (2022) suggests employing artificial intelligence (AI) agents, sensors, blockchain, the Internet of Things (IoT), expert systems, and machine learning algorithms to automate specific tasks in agriculture. Examples of automated agricultural tools include robotic systems for sowing and weeding, harvest automation devices, automatic irrigation systems, autonomous tractors, and drones for spraying fertilizers (Krishnan & Swarna, 2020).

Agriculture is on the brink of a new revolution driven by data and connectivity. Technologies such as artificial intelligence, analytics, and connected sensors have the potential to significantly enhance yields, optimize water usage, and improve input efficiency. These innovations are also vital for building sustainability and resilience in both crop cultivation and animal husbandry (Goedde, Katz, Menard & Revellat, 2020). The scope of digital farming technologies is vast, incorporating automation, sensors, and robotics in agricultural production systems, making farming more efficient and sustainable.

## MATERIALS AND METHODS

This research work focused on improving land selection for farming crops based on the chemical properties of the soil. This work utilizes supervised learning decision tree (DT) to build model for predicting the suitability of the soil for farming. The different crops grow differently, and suitability depend on the soil nutrient. Good and adequate crop yields depend on the soil compositions, and properties.

The model takes data as input which is entered, and based on the chemical properties inputted as data, model take a decision and classify soil as fertile, partially fertile, or not fertile. Fig. 1 shows the model development circle.

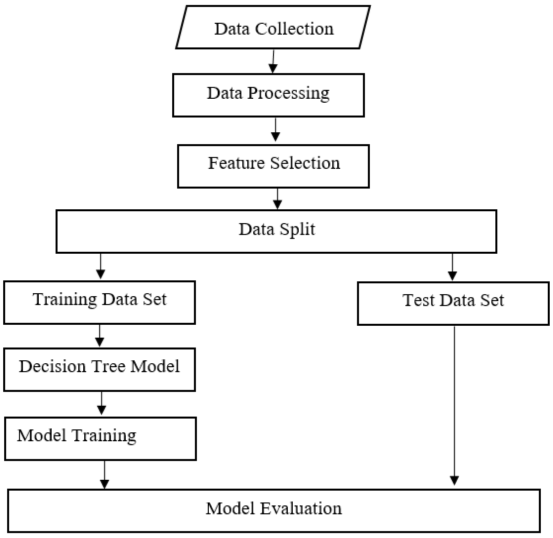


Figure 1: posed model life cycle

Many systems provide generic, block-specific advisories, overlooking the unique needs at the individual farm level, especially in African nations like Nigeria. Current systems also rely on one-way communication channels or offer limited interactivity, enabling global knowledge sharing but restricting the practical application and utilization of models and systems developed within the region.

Abdulbasit, Adewumi and Victoria (2023) conducted a research study on the topic Crop Yield Prediction in Nigeria Using Machine Learning Techniques: (A Case Study of Southern Part of Nigeria). The study highlighted that big data and machine learning are the major key tool for digitalizing the agriculture sector and other industries to predict farm produce. It also identified that the inability of farmers to accurately predict yield is a great problem using previous farming experience. The study adopted three (3) machine learning approaches, including a decision tree classifier, random forest, and support vector machine, to model data from different zones and make predictions. The techniques adopted were tested using root mean square error to ensure the right prediction algorithm is adopted and the right values are obtained. Results show prediction from the South East is the best in terms of yields and accuracy when tested and evaluated, with 138.9 %.

Pant et al. (2021) created a trained model to identify patterns among data and produce predictions for four important crops, including potatoes, rice, wheat, and maize. He employed machine learning techniques. The input fields for the dataset provided by FAOSTAT (food and agriculture organization of the United Nations) were as follows: item collected, country, item year beginning in 1990 through 2016, and yield value for the year 2021. The dataset was pre-processed, and then 70% was allocated to training and 30% to testing. Support vector machines, decision trees, random forests, and gradient-boosting regressors were used to build the model. When compared to other algorithms, the decision tree regressor had the best accuracy of 96%.

The data collection process is a fundamental component of this research. To gather the necessary data for model development and validation, a multi-pronged approach was employed. The first and foremost step in most ML predictive projects is picking a suitable dataset (Malti, & Apoorva, 2015). The study used a historical dataset extracted from www.hub.arcgis.com for building and testing the model. The dataset consists of 1550 rows and seven columns.

Data encoding in machine learning involves converting categorical features into a numerical format that can be used by machine learning algorithms. To achieve this, features were encoded to “Yes” or “No”.

Each column represents whether a particular category is present or not for a given observation. The “Yes” encoding represents soil is fertile and “No” encoding represents Not fertile. Target class is encoded fertile to represent the present of all the nutrients, not fertile to represent the absent of all or most of the nutrients, and partially fertile represent nutrient partly present but enough to give good yield. Sample of features and their encoding is as shown table 1.

**Table 1: Feature Encoded**

| N | P | Ca | Mg | K | Cu | Mn | Fe | Target |
|---|---|---|---|---|---|---|---|---|
| Yes | No | Yes | No | No | No | No | Yes | Partially fertile |
| No | No | No | Yes | No | No | No | No | Partially fertile |
| No | No | Yes | No | No | No | Yes | No | Partially fertile |
| No | No | No | No | No | No | No | No | Not fertile |
| No | No | No | No | No | No | No | No | Partially fertile |
| No | No | No | No | No | No | Yes | No | Partially fertile |
| No | No | No | No | No | No | Yes | No | Partially fertile |
| No | No | No | No | No | No | Yes | No | Partially fertile |
| Yes | No | Yes | No | No | No | No | Yes | Partially fertile |
| No | No | No | No | No | No | Yes | No | Partially fertile |
| No | No | No | No | No | No | No | No | Not fertile |
| No | No | No | No | No | No | No | No | Not fertile |

| N | P | Ca | Mg | K | Cu | Mn | Fe | Target |
|---|---|---|---|---|---|---|---|---|
| No | No | No | No | No | No | Yes | No | Partially fertile |
| No | No | No | No | No | No | No | No | Not fertile |
| No | No | No | No | No | No | Yes | No | Partially fertile |
| Yes | No | Yes | No | No | No | No | Yes | Partially fertile |
| No | No | No | No | No | No | No | No | Not fertile |
| No | No | No | No | No | No | No | No | Not fertile |
| No | No | No | No | No | No | Yes | No | Partially fertile |
| No | No | No | No | No | No | Yes | No | Not fertile |
| No | Yes | No | No | No | No | No | No | Partially fertile |
| No | No | No | No | No | No | No | No | Not fertile |
| No | No | No | No | Yes | No | Yes | No | Partially fertile |
| No | No | No | No | No | No | No | No | Not fertile |
| No | No | No | No | No | No | No | No | Not fertile |
| No | Yes | No | No | No | No | Yes | No | Not fertile |
| No | No | No | No | No | No | Yes | No | Partially fertile |
| No | No | No | No | No | No | No | No | Not fertile |
| No | No | No | No | No | No | No | No | Not fertile |
| No | No | No | No | No | No | Yes | No | Partially fertile |
| No | No | No | Yes | No | No | Yes | Yes | Partially fertile |
| No | No | No | Yes | No | No | Yes | No | Partially fertile |
| No | No | No | Yes | No | No | No | No | Partially fertile |
| No | No | No | No | No | No | No | No | Not fertile |
| No | No | No | No | No | No | Yes | No | Partially fertile |

Data Splitting is a stage involves dividing the dataset into multiple subsets to train, validate, and test the machine learning model. The goal is to evaluate the model's performance on unseen data and avoid overfitting, where the model performs well on the training data but poorly on new data.

Training Set (80%): A total of 1240 randomized samples were used for training the machine learning model. This subset helps the model learn patterns and relationships within the data.

Testing Set (20%): A total of 310 randomized samples were set aside to test the model's final performance. This provides an unbiased estimate of the model’s ability to make predictions on unseen data.

The decision tree was selected for model building as it is a versatile machine learning algorithm widely used for both classification and regression tasks. The decision tree splits data into subsets based on input features, with each split leading to a tree-like structure of nodes and branches.

Upta, Arora, Rawat, Jain and Dhami (2017) pointed out why decision tree be used:

i. Decision trees can be visualized and are simple to understand and interpret.
ii. They require very little data preparation whereas other techniques often require data normalization, the creation of dummy variables and removal of blank values.
iii. The cost of using the tree (for predicting data) is logarithmic in the number of data points used to train the tree.
iv. Decision trees can handle both categorical and numerical data whereas other techniques are specialized for only one type of variable.
v. Decision trees can handle multi-output problems.
vi. Uses a white box model i.e. the explanation for the condition can be explained easily by Boolean logic because there are mostly two outputs. For example yes or no.
vii. Decision trees can perform well even if assumptions are somewhat violated by the dataset from which the data is taken.

Python, with libraries like NumPy, Pandas, Matplotlib, Seaborn, and scikit-learn, was used to build the decision tree model. The `DecisionTreeClassifier()` function was employed to create and fit the model.

The decision tree was trained to predict soil fertility based on chemical properties, using features like Nitrogen (N), Phosphorus (P), Potassium (K), and other nutrients.

## RESULTS AND DISCUSSION

Accuracy, Precision, Recall, and confusion Matrix are the techniques and metric used to evaluate the model to assess its performance, and generalization ability. Table 2 describes the result of the model performance on soil for millet farming

**Table 2: Accuracy, precision, recall and f1_score results**

| Metric | Not fertile % | Partially fertile % | Accuracy % | |
|---|---|---|---|---|
| | | | Training set | Testing set |
| Precision | 87 | 91 | 89.7 | 90 |
| Recall | 87 | 91 | | |
| F1_score | 87 | 92 | | |

Table 2 displays the tabulated results for the precision, recall, and acquired when a model was tested for the dataset. From the table, it is observed different metric values of the model. The model achieved an accuracy of 89.7% on the training data and 90% on the test data. The accuracy is high for both the training and test datasets, indicating that the model is performing well and is likely not overfitting. Below is the confusion matrix for millet crop soil fertility

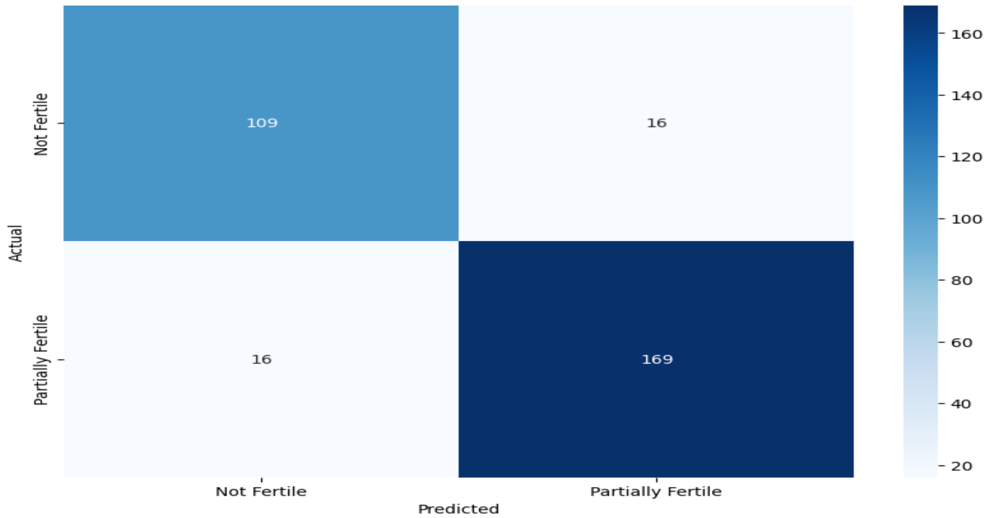


Figure 2: Confusion Matrix for Millet Crop Soil Fertility

109 stances are correctly predicted as "Not Fertile." Meaning the soil is not good for cultivating millet crops. 169 instances are correctly predicted as "Partially Fertile.". This implies that millet can partially be cultivated based on the soil properties. For the False Positives prediction, 16 instances were incorrectly predicted as "Not Fertile" when they are actually "Partially Fertile." Also, 16 instances are incorrectly predicted as "Partially Fertile" when they are actually "Not Fertile."
The result of the model features is shown in figure 3

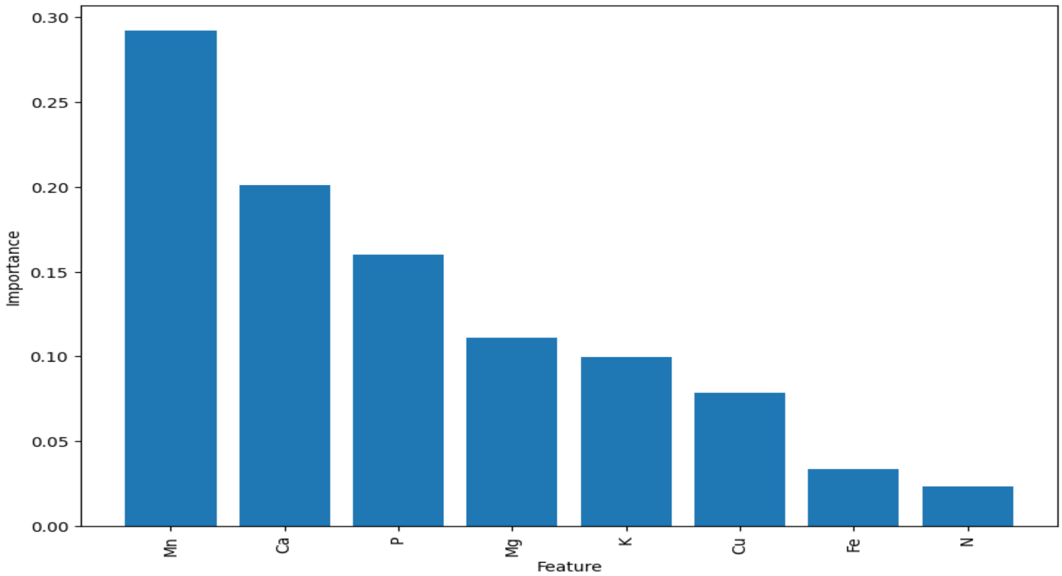


Figure 3: Feature performance for Millet crop

Manganese is the most important feature for the model. It has the highest importance score, suggesting that variations in manganese levels are highly influential in determining whether the sample is classified as "Not Fertile" or "Partially Fertile." Calcium is the second most important feature of 0.20. It plays a significant role, though not as critical as manganese, in the model's classification decisions. Phosphorus is also a key feature, contributing significantly to the model's predictions. Its importance is 0.15 slightly lower than calcium. Magnesium is the fourth most important feature with a value of 0.10, indicating that it still has a substantial impact on the model's decisions, but less than manganese, calcium, and phosphorus. 0.08 Potassium is moderately important. Its importance is lower than the top four features but still relevant. Copper has a moderate influence of 0.07 on the model's predictions, comparable to potassium. Iron and nitrogen have lower importance score of 0.3 and 0.2 respectively, indicating that they play a less significant role in the model's decision-making process compared to the other features. The following table 3 describe the result of the model performance on soil for Maiza and Guinea corn

**Table 3: Accuracy, precision, recall and f1_score results**

| Metric | Not fertile % | Partially fertile % | Accuracy % | |
|---|---|---|---|---|
| | | | Training set | Testing set |
| Precision | 83 | 92 | 86.11 | 89.01 |
| Recall | 84 | 91 | | |
| F1_score | 83 | 92 | | |

Table 3 displays the tabulated results for the precision, recall, and acquired when a model was tested for the dataset. From the table, it is observed different metric values of the model. The value tells how well the model performs in predictions in train and test datasets respectively. As shown the model accuracy on the training dataset was 86.11%, implying that the model did well on the training data set prediction. Similarly, the accuracy of the model predictions with test data as an input to the is 89.03%, which is slightly better in performance on test set for best soil for maize and guinea corn crops. The slightly higher accuracy on the test set compared to the training set suggests that the model performed well with new data and is not overfitting. The following figure describes the Model confusion matrix for maize and guinea corn crops soil fertility

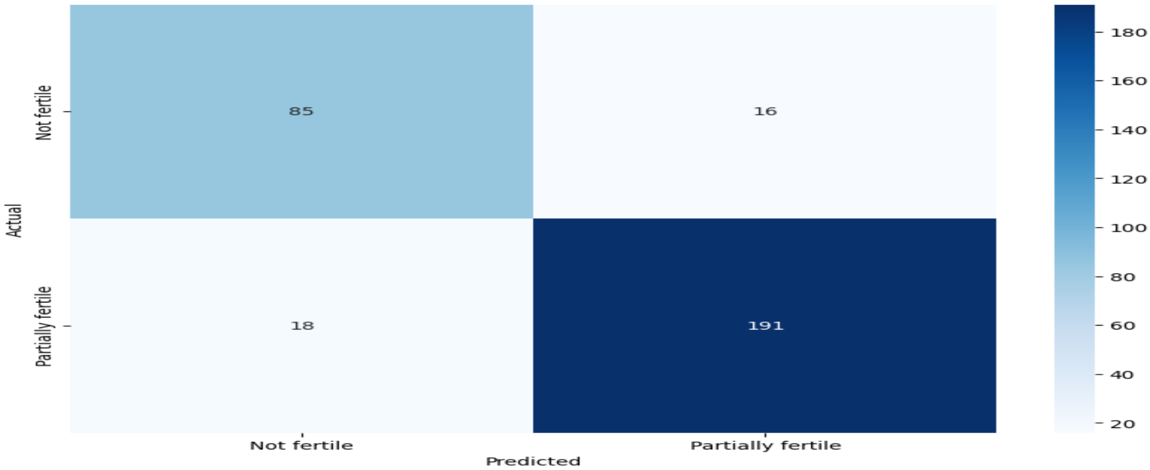


Figure 4: Confusion Matrix for Maize and Guinea corn Crops Soil Fertility

85 stances are correctly predicted as "Not Fertile." Meaning the soil is not good for cultivating maize and guinea corn crops. 191 instances are correctly predicted as "Partially Fertile.". This implies that maize and guinea corn can partially be cultivated based on the soil properties.
For the False Positives prediction, 16 instances were incorrectly predicted as "Not Fertile" when they are actually "Partially Fertile." Also, 18 instances are incorrectly predicted as "Partially Fertile" when they are actually "Not Fertile."
The model performs better at predicting the "Partially Fertile" class, as indicated by higher precision (92.2%) and recall (91.4%) compared to the "Not Fertile" class. The "Not Fertile" class has a slightly lower precision (82.5%) and recall (84.2%), but these values indicate that the model performance is good.

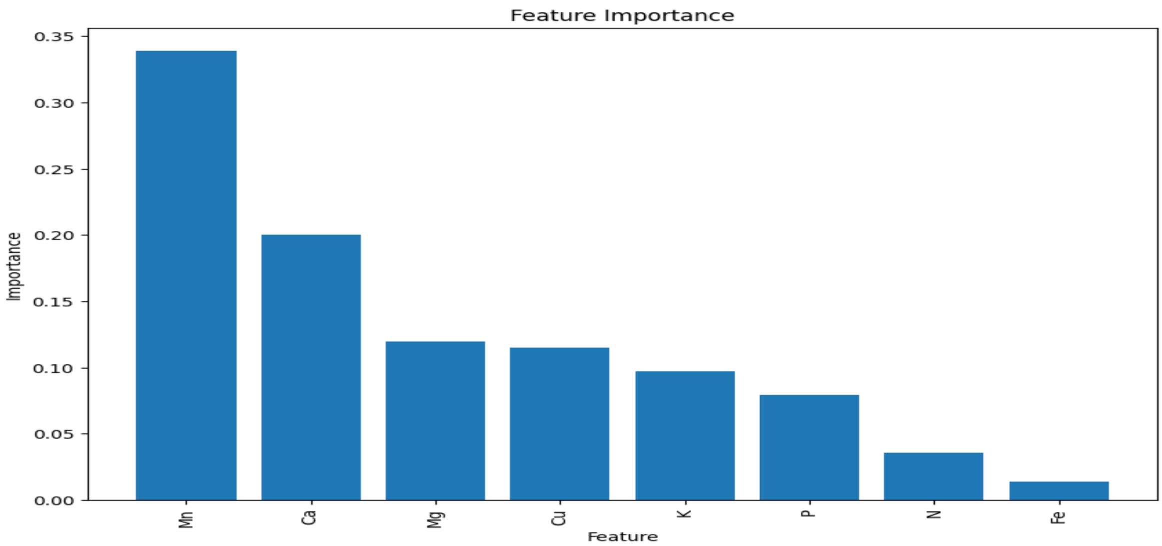


Figure 5: Feature importance for Maize and Guinea Corn Crops

Manganese is the most important feature for the model. It has the highest contribution towards predicting the fertility status of the soil for maize and guinea corn. This implies that variations in manganese levels are highly indicative of whether the soil is fertile or not.

0.20 of calcium is the second most important feature with a value of 0.20. Its relatively high importance indicates that calcium levels are also crucial in determining soil fertility. This indicates that adequate calcium is essential for plant growth and soil structure.

Magnesium is the third most important feature of 0.12. its significance indicates that magnesium levels in the soil play a notable role in fertility. Copper is another important feature as classified by the model, contributing significantly to the model's predictions with a value of 0.11. Copper is essential for various plant physiological processes, and its availability can affect soil fertility.

Potassium is essential for plant metabolism and growth. Its importance in the model indicates that potassium levels are a key factor in determining soil fertility. 0.08 Phosphorus is critical for energy transfer and genetic material in plants. Its importance reflects its role in soil fertility.

0.04 Nitrogen is crucial for plant growth and development. Although it has a lower importance compared to other nutrients, it still contributes to fertility predictions. 0.02 Iron, although the least important feature, still plays a role in the model. Iron is essential for chlorophyll synthesis and enzyme function in plants.

## CONCLUSION

The research focused on developing a model using machine learning to predict suitable farmland for the cultivation of specific crops. The study demonstrated that the Decision Tree outperformed several alternatives in terms of both classification accuracy and interpretability, making it a practical tool for guiding data-driven agricultural decisions. The model has the potential to support sustainable farming practices, optimizes land use and ultimately enhance food production.

The model was trained and tested on a dataset sourced from Adamawa State in Nigeria, achieving a high accuracy rate. The results indicated that the model could reliably predict soil suitability based on various nutrient parameters.

## REFERENCES


Abdulbasit, A., Adewumi, S., E. & Victoria, Y. (2023). Crop Yield Prediction in Nigeria Using Machine Learning Techniques: (A Case Study of Southern Part of Nigeria). A periodical of the Faculty of Natural and Applied Sciences, UMYU, Katsina.

Bichri1, H., Chergui, A., & Hain M. (2024). Investigating the Impact of Train/ Test Split Ratio on the Performance of Pre-Trained Models with Custom Datasets. International Journal of Advanced Computer Science and Applications, Vol. 15, No. 2, 2024

Blesslin, T. S., vijay, A. L. D., Gunaselvi, M., Saravana, S., Wilfred, C. B., Muthukumar, K., Padmavathy S., Ramesh P. K., & Belete T. A. (2022). Machine learning Algorithm for Soil Analysis and Classification of Micronutrients in IOT-Enabled Automated Farms. Journal of Nanomaterials. Volume 2022 | Article ID 5343965. Retrieved on Saturday, 16th march, 2024 from https://www.doi.org/10.1155/2022/5343965

Chakraborty, S. K., Chandel, N. S., Jat, D., Tiwari, M. K., Rajwade, Y. A., & Subeesh, A., (2022). Deep learning approaches and interventions for futuristic engineering in Agriculture. Neural Computing and Applications, 34 (23), 20539–20573

Eastwood, C. R., Chapman, D. F.& Paine, M. S. (2012). Networks of practice for construction of agricultural decision support systems: Case studies of precision dairy farms in Australia. Agricultural Systems 108:10-18.

Elbasi, E., Zaki, C., Topcu, A.E., Abdelbaki, W., Zreikat, A.I., Cina, E., Shdefat, A. & Saker, L. (2024). Crop Prediction Model Using Machine Learning Algorithms. Appl. Sci. 2023, 13, 9288. Retrieved on 16th March, 2024 from https://doi.org/10.3390/app13169288

Khaki, S. and Wang, L., (2019). Crop Yield Prediction Using Deep Neural Networks. Frontiers in Plant Science, 10.

Kuehne, G., Llewellyn, R., Pannell, D. J., Wilkinson, R., Dolling, P., Ouzman, J. & Ewing, M. (2017). Predicting farmer uptake of new agricultural practices: A tool for research, extension and policy. Agricultural Systems 156:115-125.

Mishra, S., Paygude, P., Chaudhary, S., & Idate, S. (2018). "Use of data mining in crop yield prediction." In: 2018 2nd International Conference on Inventive Systems and Control (ICISC), pp. 796-802. IEEE.

Motia, S. & Reddy, S. R. N. (2021). Exploration of machine learning methods for prediction and assessment of soil properties for agricultural soil management: a quantitative evaluation. Journal of Physics: Conference Series. 1950 (2021) 012037. Retrieved on Sunday 17th March, 2024 from https://doi.org/10.1088/1742-6596/1950/1/012037

Neethirajan, S. (2017). Recent advances in wearable sensors for animal health management. Sensing and Bio-Sensing Research 12:15-29.

Nettle, R., Crawford, A. & Brightlng P. (2018). How private-sector farm advisors change their practices: Journal of Rural Studies 58:20-27.

Nigam, A., Garg, S., Agrawal, A., & Agrawal, P., (2019). "Crop yield prediction using machine learning algorithms." In: 2019 Fifth International Conference on Image Information Processing (ICIIP), pp. 125-130. IEEE.

Pannell, D. J., Marshall, G. R., Barr, N., Curtis, A., Vanclay, F. & Wilkinson, R. (2006). Understanding and promoting adoption of conservation practices by rural landholders. Australian Journal of Experimental Agriculture 46:1407-1424.

Pant, J., Pant, R. P., Singh, M. K., Singh, D. P., & Pant, H., (2021). "Analysis of agricultural crop yield prediction using statistical techniques of machine learning." Materials Today: Proceedings, 46, 10922-10926.

Panwar, E., Kukunuri, A. N. J., Singh, D., Sharma, A. K., & Kumar, H., 2022. An Efficient Machine Learning Enabled Non-Destructive Technique for Remote Monitoring of Sugarcane Crop Health. IEEE Access, https://doi.org/10,75956-75970.

Prager, K., Labarthe, P., Caggiano, M. & Lorenzo-Arribas, A. (2016). How does commercialization impact on the provision of farm advisory services? Evidence from Belgium, Italy, Ireland and the UK. Land Use Policy 52:329-344.

Prasad, N. R., Patel, N. R.,& Danodia, A., (2021). "Crop yield prediction in cotton for regional level using random forest approach." Spatial Information Research, 29(2), 195-206.

Prasad, N. R., Patel, N. R.,& Danodia, A., (2021). "Crop yield prediction in cotton for regional level using random forest approach." Spatial Information Research, 29(2), 195-206

Sajja, G. S., Jha, S. S., Mhamdi, H., Naved, M., Ray, S. & Phasinam, K. (2021). "An investigation on crop yield prediction using machine learning." In: 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA) (pp. 916-921). IEEE.

Upta, B., Arora, A., Rawat, A., Jain, A., & Dhami, J. (2017). Analysis of Various Decision Tree Algorithms for Classification in Data Mining. International Journal of Computer Applications (0975 – 8887) Volume 163 – No 8, April 2017

Vadlamudi, S. (2019). "How Artificial Intelligence Improves Agricultural Productivity and Sustainability: A Global Thematic Analysis." Asia Pacific Journal of Energy and Environment, 6(2), 91-100.