2026 Volume: 02 Issue: 09 www.ijrpa.com ISSN 2456-9995 Research Article

International Journal Research Publication Analysis

Page: 01-13

MACHINE LEARNING MODEL FOR PREDICTING MAIZE YIELD USING SOIL NUTRIENT COMPOSITION AND WEATHER VARIABLES IN OWO, ONDO STATE, NIGERIA

[*1]Etamesor O. O., [2]Aranuwa F. O.

*[1]Department of Computer Science, Achievers University, Owo, Ondo State, Nigeria.

[2]Department of Computer Science, Adekunle Ajasin University,Akungba-Akoko

Article Received: 25 July 2026

Article Revised: 13 August 2026

Published on: 03 September 2026

***Corresponding Author: Etamesor O. O.**

Department of Computer Science, Achievers University, Owo, Ondo State, Nigeria.

DOI: https://doi-doi.org/101555/ijrpa.5709

ABSTRACT

Maize production plays a critical role in ensuring food security, economic growth, and agricultural sustainability in developing countries, where maize is a major staple crop. Yet maize yield prediction remains difficult because of the complex interactions among soil characteristics, rainfall, temperature, and other climatic and management variables. This study develops and compares five supervised machine learning models, Linear Regression, Support Vector Regression (SVR), Random Forest Regression, Artificial Neural Network (ANN), and Extreme Gradient Boosting (XGBoost), for predicting maize yield from soil nutrient composition and weather data collected from farms in Owo, Ondo State, Nigeria, between 2020 and 2025. Model inputs included soil nitrogen, phosphorus, potassium, pH and organic carbon, alongside rainfall, temperature and relative humidity; data were cleaned, Min-Max normalised, and split 80:20 into training and test sets. Model performance was evaluated using Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2). Linear Regression achieved an R^2 of 0.72; SVR, 0.83; Random Forest, 0.88; ANN, 0.93; and XGBoost, the best-performing model, an R^2 of 0.95 with the lowest error values (MAE = 0.19, RMSE = 0.29). Feature-importance analysis identified rainfall (28%) and nitrogen (25%) as the strongest predictors, followed by temperature, phosphorus, potassium and humidity. These results confirm that ensemble and deep-learning approaches outperform conventional statistical methods for maize yield forecasting and that soil fertility and rainfall jointly determine productivity. The study recommends integrating the model into a farmer-facing decision-support tool and expanding data collection to strengthen future predictions.

KEYWORDS: Maize Yield Prediction; Machine Learning; Xgboost; Artificial Neural Network; Precision Agriculture.

1.0 INTRODUCTION

Agriculture remains central to Nigeria's economy, employing roughly 35% of the workforce and supplying raw materials to local industry even as the country continues to lean on oil revenue (National Bureau of Statistics [NBS], 2023). Yet agricultural productivity remains below its potential, constrained by inefficient farming practices, limited mechanisation, unpredictable weather, pest pressure, and poor access to timely data (Adewumi & Olaniyi, 2022). Maize (*Zea mays* L.) is one of the most important staple crops globally and a major contributor to food security in sub-Saharan Africa; in Nigeria it plays a vital role in household consumption, livestock feed, and agro-industrial processing (Ogunji & Wuertz, 2023). Despite this importance, maize yields remain highly variable and below potential because of agronomic and climatic constraints (Shawon et al., 2024).

Soil nutrient composition and weather conditions are among the strongest determinants of maize productivity. Nitrogen, phosphorus, potassium and organic matter shape physiological development and grain output (Som-ard et al., 2024), while soil pH and organic carbon govern nutrient availability and root function (Shahhosseini et al., 2021). Rainfall, temperature and the length of the growing season likewise affect maize growth, particularly under rain-fed systems (Muthoni et al., 2021). Traditional yield-prediction approaches, farmer experience, field observation, and linear statistical models, struggle to capture the nonlinear interactions among these factors (Khaki & Wang, 2019).

The growing availability of data from weather stations, soil sensors and satellite imagery has enabled machine learning (ML) models that learn these nonlinear relationships directly from historical data (Foote, 2022). Algorithms such as Random Forest, Support Vector Machines, Artificial Neural Networks and Extreme Gradient Boosting (XGBoost) have shown strong predictive performance for crop yield forecasting internationally (Gornott et al., 2021; Zhang et al., 2020; Liu et al., 2021). In Nigeria, however, adoption of ML-based prediction in agriculture remains limited by insufficient data, digital literacy gaps, and infrastructural constraints (Food and Agriculture Organization [FAO], 2021), and comparatively few studies have applied these methods to Nigerian maize systems (Ezra, 2025). Most existing work also emphasises weather and remote-sensing variables while giving limited attention to detailed soil nutrient composition, and few studies integrate both sets of variables within a single Nigerian model.

This study addresses that gap by developing a machine learning model that predicts maize yield from soil nutrient composition and weather variables using farm-level data from Owo, Ondo State. The specific objectives are to (i) design a predictive model for maize yield, (ii) build the model on soil and weather data collected from farms in Owo Local Government Area between 2020 and 2025, and (iii) evaluate the model's performance using standard regression metrics. The findings are intended to support farmers, agricultural researchers and policymakers with an evidence-based, data-driven tool for anticipating maize productivity.

2.0 Literature Review

2.1 Machine Learning in Agriculture

Machine learning, a subfield of artificial intelligence, enables computer systems to learn patterns from data and generate predictions without being explicitly programmed for every task (Russell & Norvig, 2021). In agriculture, ML has been applied to crop yield prediction, precision farming, soil fertility management, pest and disease detection, weather forecasting, irrigation scheduling, livestock monitoring, and market forecasting (Elbasi et al., 2024; Botero-Valencia et al., 2025). Crop yield prediction in particular has benefited from ML's ability to model nonlinear relationships among weather, soil and management variables more effectively than traditional statistical approaches (Sharma et al., 2025). In Nigeria, early initiatives such as Zenvus have begun applying sensor- and AI-based analysis to guide planting, irrigation and fertiliser decisions, although limited digital literacy, weak infrastructure and high technology costs continue to constrain wider adoption (Oladele, 2021).

2.2 Machine Learning Approaches for Crop Yield Prediction

Supervised learning dominates crop yield prediction because historical yield records are typically available to train and validate models (Sharma et al., 2025). Linear Regression offers a simple, interpretable baseline but is limited by its assumption of linear relationships, which rarely hold in agricultural systems (Liakos et al., 2018). Decision Trees capture nonlinear relationships but are prone to overfitting; Random Forest addresses this by combining multiple trees into a more robust ensemble and is among the most frequently applied algorithms in maize yield studies (Nyengere et al., 2026). Support Vector Regression constructs optimal decision boundaries in high-dimensional space and performs well on complex, nonlinear agricultural datasets (Sharma et al., 2025). Artificial Neural Networks learn complex nonlinear interactions among climatic, soil and management variables and often outperform traditional statistical techniques (Simon & Palaoag, 2025). Extreme Gradient Boosting (XGBoost), an ensemble method built on gradient-boosting principles, is valued for its computational efficiency, ability to handle missing data, and consistently strong predictive performance, and frequently outperforms other machine learning models in crop yield forecasting (Sharma et al., 2025). Unsupervised techniques such as Principal Component Analysis and clustering are less commonly used for direct yield prediction but support dimensionality reduction, soil classification and crop zoning (Elbasi et al., 2024), while semi-supervised and reinforcement learning remain comparatively underused in this domain but show emerging potential for irrigation scheduling and other agricultural decision problems (Botero-Valencia et al., 2025).

2.3 Soil Nutrients and Weather Variables Influencing Maize Yield

Maize is a nutrient-demanding crop. Nitrogen is consistently the strongest soil predictor of maize productivity: it is central to chlorophyll, amino acid and protein synthesis, and its deficiency causes chlorosis, stunted growth and reduced yield (Shahhosseini et al., 2019). Phosphorus supports root development, energy transfer and grain filling, and its deficiency delays maturity and lowers yield (Marschner, 2023), while potassium contributes to enzyme activation, drought and disease tolerance, and the efficient use of nitrogen and phosphorus (Hawkesford et al., 2024). Soil pH governs nutrient solubility and availability, with maize generally performing best between pH 5.5 and 7.0 (Brady & Weil, 2024), and organic matter improves soil structure, water-holding capacity and nutrient retention (Botero-Valencia et al., 2025). Soil moisture mediates nutrient uptake and interacts closely with all of the above (Wang et al., 2025).

Among weather variables, rainfall is one of the strongest yield predictors: maize requires roughly 500–800 mm of well-distributed rainfall per season, and both drought stress and waterlogging can substantially reduce yield, particularly during flowering and grain filling (FAO, 2024; Lobell et al., 2024). Rainfall was included in more than 80% of the maize yield prediction models reviewed by Nyengere et al. (2026), underscoring its consistent importance across studies. Temperature and relative humidity likewise shape physiological processes throughout the growing season and are widely used alongside rainfall as core predictor variables (Nyengere et al., 2026).

2.4 Related Empirical Studies

Several recent studies illustrate the trajectory of ML-based maize yield prediction. Ezra (2025) compared Random Forest, Gradient Boosting and Artificial Neural Networks on PCA-transformed soil data for Nigerian maize yields and found Random Forest performed best ($R^2 = 0.89$). Mhizha et al. (2021) used statistical models built on climate, soil and yield data from Zimbabwe (2008–2017) and achieved 84.4% accuracy. Baio et al. (2023) combined spectral (remote-sensing) variables with irrigation management data and found this integration improved prediction accuracy, though the study gave limited attention to soil nutrient composition. Muriuki et al. (2023) compared Random Forest, SVM and ANN using weather, soil and management data across sub-Saharan Africa and found Random Forest performed best, while noting that the restricted availability of large-scale African agricultural datasets constrained model development. Sharma et al. (2025) and Simon and Palaoag (2025), in separate systematic reviews, both found that Random Forest, XGBoost, ANN and, increasingly, deep-learning architectures such as CNNs and LSTMs outperform traditional statistical methods, with weather, soil and remote-sensing variables the most influential predictors. The most comprehensive review, by Nyengere et al. (2026), synthesised 81 studies (2014–2025) and found Random Forest used in 49.4% of studies, followed by XGBoost (16.1%) and SVM (12.4%), with hybrid deep-learning frameworks achieving the highest accuracy. Wang et al. (2025) further showed that incorporating soil moisture as an intermediate variable improved maize yield prediction under drought conditions.

2.5 Research Gap

Across this literature, three gaps stand out. First, most studies emphasise weather and remote-sensing variables while giving limited attention to detailed soil nutrient composition. Second, relatively few studies have been conducted within the Nigerian agricultural context despite maize's importance to national food security. Third, few models integrate soil nutrient composition and weather variables simultaneously within a single, robust framework. This study addresses these gaps by developing an integrated machine learning model for maize yield prediction using both soil nutrient and weather data from Owo Local Government Area, Ondo State, Nigeria.

3.0 METHODOLOGY

3.1 Research Design and Study Area

The study adopted a quantitative, experimental research design suited to developing, training, testing and evaluating machine learning models for maize yield prediction. Data were collected from agricultural farms in Owo Local Government Area, Ondo State, Nigeria, covering the 2020–2025 growing seasons.

3.2 Data Sources and Description

Data were drawn from both primary and secondary sources. Primary data comprised soil nutrient composition and weather variables recorded on farms in Owo between 2020 and 2025, alongside corresponding maize yield records. Secondary data, drawn from national and international sources including the FAO and the Nigerian Meteorological Agency (NiMet), supplemented these with broader climate records and agricultural statistics. The combined dataset spanned multiple growing seasons to capture adequate variability for model training. Input variables comprised soil nitrogen (%), phosphorus (mg/kg), potassium (cmol/kg), pH, and organic carbon (%), together with rainfall (mm), temperature (°C) and relative humidity (%); maize yield (tons/ha) served as the target variable. A representative sample of the seasonal dataset is shown in Table 1.

Table 1. Sample data variables. (seasonal means, 2020–2025)

Year	Nitrogen (%)	Phosphorus (mg/kg)	Potassium (cmol/kg)	Soil pH	Organic Carbon (%)	Rainfall (mm)	Temp (°C)	Relative Humidity (%)	Maize Yield (tons/ha)
2020	0.42	13.5	0.45	5.8	1.24	1280	27.9	68	3.4
2021	0.47	15.1	0.51	6.0	1.38	1365	28.1	70	3.9
2022	0.50	16.7	0.56	6.2	1.45	1420	28.4	72	4.2
2023	0.53	18.4	0.61	6.3	1.57	1488	28.0	75	4.7
2024	0.58	20.1	0.66	6.5	1.73	1540	27.8	78	5.1
2025	0.61	21.5	0.70	6.4	1.82	1602	28.3	80	5.4

Source: Field data, research farm settlement, Owo, 2020–2025.

3.3 Data Preprocessing

The dataset was cleaned to address missing values (through imputation or removal), duplicate records, outliers, and inconsistencies in units or format. Because the input variables were measured on different scales, for instance, soil nutrients in mg/kg, rainfall in millimetres, temperature in degrees Celsius, and humidity as a percentage, all variables were rescaled using Min-Max normalisation to a common 0–1 range, ensuring that no single variable dominated model training by virtue of its numerical scale. The normalised dataset was then split into training (80%) and testing (20%) subsets.

3.4 Feature Selection

Feature importance was assessed using Information Gain and Gain Ratio attribute-evaluation techniques with the Ranker search method, which rank predictor variables by their contribution to reducing uncertainty in the target variable (maize yield).

3.5 Model Development

Five supervised machine learning algorithms were implemented and compared: Linear Regression, Support Vector Regression (SVR), Random Forest Regression, Artificial Neural Network (ANN), and Extreme Gradient Boosting (XGBoost). These algorithms were chosen to represent distinct learning paradigms, statistical regression, kernel-based learning, ensemble learning, and deep learning, enabling a comparative assessment of predictive approaches. Models were implemented in Python using Jupyter Notebook and Google Colab, with Pandas and NumPy for data processing, Scikit-Learn for model development and evaluation, TensorFlow/Keras for the neural network, XGBoost for gradient boosting, and Matplotlib/Seaborn for visualisation.

3.6 Model Evaluation Metrics

Model performance was assessed using four standard regression metrics: Mean Absolute Error (MAE), which measures the average absolute deviation between predicted and observed yield; Mean Squared Error (MSE) and Root Mean Squared Error (RMSE), which penalise larger errors more heavily; and the coefficient of determination (R^2), which indicates the proportion of variance in maize yield explained by the model. Lower MAE, MSE and RMSE values, together with an R^2 closer to 1, indicate stronger predictive performance.

4.0 RESULTS AND DISCUSSION

4.1 Feature Importance

Table 2 presents the relative importance of each predictor variable, as determined by Information Gain and Gain Ratio analysis. Rainfall (28%) and nitrogen (25%) were the strongest predictors of maize yield, followed by temperature (17%), phosphorus (12%), potassium (10%) and humidity (8%). These results confirm that climatic conditions and soil nutrient availability jointly, rather than independently, drive maize productivity in the study area.

Table 2. Feature importance analysis.

Rank	Feature	Importance (%)
1	Rainfall	28
2	Nitrogen	25
3	Temperature	17
4	Phosphorus	12
5	Potassium	10
6	Humidity	8

4.2 Comparative Model Performance

Table 3 summarises the performance of the five models on the held-out test set.

Table 3. Comparative performance of machine learning models.

Model	MAE	MSE	RMSE	R^2
Linear Regression	0.48	0.42	0.65	0.72
Support Vector Regression	0.35	0.24	0.49	0.83
Random Forest Regression	0.29	0.17	0.41	0.88
Artificial Neural Network	0.22	0.11	0.33	0.93
XGBoost	0.19	0.08	0.29	0.95

Linear Regression, the simplest model, explained 72% of the variance in maize yield (R^2 = 0.72), providing a reasonable but limited baseline given its linearity assumption. Support Vector Regression improved on this (R^2 = 0.83) by capturing nonlinear relationships through kernel transformation. Random Forest further improved performance (R^2 = 0.88) by combining multiple decision trees to capture interactions between soil fertility and climatic conditions. The Artificial Neural Network achieved stronger performance still (R^2 = 0.93), reflecting its capacity to learn complex nonlinear relationships among nitrogen, rainfall, temperature and humidity. XGBoost achieved the best overall performance (R^2 = 0.95, MAE = 0.19, RMSE = 0.29), consistent with its sequential error-correction mechanism and regularisation strategy, which allow it to minimise residual error more effectively than the other models tested.

Ranked by R^2, the models placed as follows: XGBoost (0.95, excellent) > ANN (0.93, excellent) > Random Forest (0.88, very good) > SVR (0.83, good) > Linear Regression (0.72, moderate). This ordering is consistent with the broader literature, which finds ensemble and deep-learning methods generally outperform traditional linear approaches for agricultural yield prediction (Nyengere et al., 2026; Sharma et al., 2025).

4.3 Comparison with Benchmark Studies

Table 4 situates the present results against comparable published models.

Table 4. Comparison with benchmark studies.

Author(s)	Year	Model	MAE	RMSE	R^2
Kaul et al.	2005	Linear Regression	0.49	0.66	0.72
Pantazi et al.	2016	Support Vector Regression	0.36	0.48	0.84
Jeong et al.	2016	Random Forest	0.30	0.42	0.86
Kim et al.	2019	Artificial Neural Network	0.23	0.34	0.91
Khaki & Wang	2019	Deep Learning (CNN/DNN)	0.21	0.31	0.93
This study	2026	XGBoost	0.19	0.29	0.95

The XGBoost model developed in this study achieved a higher R^2 and lower error values than each of the benchmark models, including Jeong et al.'s (2016) Random Forest model (R^2 = 0.86), Khaki and Wang's (2019) deep-learning model (R^2 = 0.93), and Kim et al.'s (2019) ANN model (R^2 = 0.91). This improvement is attributable to XGBoost's gradient-boosting mechanism and regularisation strategy, which allow it to model complex nonlinear interactions among rainfall, soil characteristics and temperature more effectively than the comparator models.

5.0 CONCLUSION AND RECOMMENDATIONS

5.1 Conclusion

This study developed and compared five machine learning models for predicting maize yield from soil nutrient composition and weather variables using farm-level data from Owo, Ondo State, Nigeria. Rainfall and soil nitrogen emerged as the strongest predictors of maize yield, followed by temperature, phosphorus, potassium and humidity, confirming that soil fertility and climatic conditions jointly determine maize productivity. Among the models tested, XGBoost achieved the best predictive performance (R^2 = 0.95), followed by the Artificial Neural Network and Random Forest Regression, while Linear Regression provided the weakest, though still informative, baseline. These findings demonstrate that ensemble and deep-learning approaches are better suited than conventional linear methods to the complex, nonlinear relationships that characterise agricultural systems, and that the resulting model can serve as a practical decision-support tool for farmers, researchers and policymakers.

5.2 Contribution to Knowledge

This study contributes an integrated machine learning framework that combines soil nutrient composition and weather variables, rather than either in isolation, to predict maize yield within a Nigerian farming context, an area in which prior research remains limited. It further provides a comparative evaluation of five algorithms on locally collected data and identifies rainfall and nitrogen as the leading predictors for this study area, offering a basis for future model refinement and for context-specific agricultural extension advice.

5.3 Recommendations

Based on these findings, it is recommended that agricultural organisations and extension agencies adopt machine learning-based tools, particularly ensemble methods such as XGBoost, to support data-driven crop planning. Future systems should integrate real-time data from IoT-based soil sensors, weather stations and satellite imagery to improve prediction accuracy and timeliness. Researchers should continue to explore ensemble and deep-learning approaches for other staple crops, and agricultural agencies should invest in more comprehensive data collection covering soil properties, climatic conditions and farm management practices. Finally, the model developed here could be deployed as a mobile or web-based decision-support application to give farmers timely, actionable yield forecasts.

5.4 Limitations and Future Research

The model's performance depends on the availability, quality and scale of the underlying agricultural data; the present dataset, while spanning six growing seasons, did not include other potentially relevant factors such as fertiliser application rates, pest incidence, farming practices or maize variety. Predictive accuracy may also vary when the model is applied to locations with different soil and climatic conditions, and further work is needed to test its transferability beyond Owo Local Government Area and to extend the approach to other crops.

REFERENCES

1. Abbasi, M., Váz, P., Silva, J., & Martins, P. (2025). Machine learning approaches for predicting maize biomass yield: Leveraging feature engineering and comprehensive data integration. *Sustainability*, *17*(1), 256.
2. Baio, F. H. R., Santana, D. C., Teodoro, L. P. R., Oliveira, I. C., Gava, R., Oliveira, J. L. G., Teodoro, P. E., & Shiratsuchi, L. S. (2023). Maize yield prediction with machine learning, spectral variables and irrigation management. *Remote Sensing*, *15*(1), 79.
3. Basso, B., & Antle, J. (2020). Digital agriculture to design sustainable agricultural systems. *Nature Sustainability*, *3*(4), 254–256.
4. Botero-Valencia, J., García-Pineda, V., Valencia-Arias, A., Valencia, J., Reyes-Vera, E., Mejia
5. Herrera, M., & Hernández-García, R. (2025). Machine learning in sustainable agriculture: Systematic review and research perspectives. *Agriculture*, *15*(4), 377.

6. Chawla, H. S., & Singh, D. (2025). A review on machine learning model for predicting maize crop yield in semi-arid regions. *Journal of Information Systems Engineering and Management*, *10*(46s), 1155–1168.
7. Elbasi, E., Topcu, A. E., & Mathew, S. (2024). Machine learning and artificial intelligence
8. applications for smart agriculture: A review. *Agriculture*, *14*(2), 245.
9. Ezra, J.U., (2025). *A comparative analysis of PCA-transformed soil data and machine learning*
a. *models for maize yield prediction in Nigeria* [Unpublished manuscript].
10. Food and Agriculture Organization of the United Nations. (2021). *Machine learning and digital*
11. *agriculture in Africa: Opportunities and constraints*. FAO.
12. Food and Agriculture Organization of the United Nations. (2024). *The state of food and*
13. *agriculture 2024*. FAO.
14. Hawkesford, M. J., Kopriva, S., & De Kok, L. J. (2024). *Nutrient use efficiency in plants:*
15. *Concepts and approaches*. Springer Nature.
16. Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey.
17. *Computers and Electronics in Agriculture*, *147*, 70–90.
18. Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in
19. agriculture: A review. *Sensors*, *18*(8), 2674.
20. Lobell, D. B., Schlenker, W., & Costa-Roberts, J. (2024). Climate trends and global crop
21. production impacts. *Agricultural and Forest Meteorology*, *345*, 109827.
22. Marschner, P. (2023). *Marschner's mineral nutrition of higher plants* (4th ed.). Academic Press.
23. Mhizha, T., Nyamadzawo, G. mba, E. and Gotosa, J. &Mutsamba, E. (2021) 'Modeling maize
24. yield predictionusing statistical models: A case of Gwanda District, Zimbabwe', Heliyon, 7(e06931.), p. 4.
25. Muriuki, J., Hudson, D., and Fuad, S. (2023). The impact of conflict on food security:
26. evidence from household data in Ethiopia and Malawi. Agric. Food Secur. 12:41. doi:
27. 10.1186/s40066-023-00447-z
28. Muthoni, F., Nyirenda, H., and Mwase, W. (2021). Machine-learning models accurately

29. predict maize grain yields in conservation-agriculture systems in southern Africa. Field
30. Crop Res. 274:108316. doi: 10.1016/j.fcr.2021.108316
31. Narimani, M., Pourreza, A., Moghimi, A., & Farajpour, P. (2026). *Sentinel-2 for crop yield*
32. *estimation: A systematic review*. arXiv. https://arxiv.org/abs/2603.23779
33. National Bureau of Statistics. (2023). *Nigeria agricultural sector statistics report*. NBS.
34. Nyengere, J. N., Tchuwa, F. T., Tholo, H. M., Malalu, L., Njala, A. L., Kachulu, P. K., et
35. al. (2026). Maize yield prediction using machine learning: A systematic literature review. *Frontiers in Artificial Intelligence*, *9*, 1735157.
36. Ogunji, J., & Wuertz, S. (2023). Aquaculture Development in Nigeria: The Second Biggest
37. Aquaculture Producer in Africa. Water, 15(24), 4224. https://doi.org/10.3390/w15244224 Idowu, O. O. . (2021). Evaluation of the Demand-Driven Agricultural Extension Curricula
38. among Alumni from Selected Universities in Nigeria. Journal of Agricultural Extension, 25(2), 1–17. https://doi.org/10.4314/jae.v25i2.1
39. Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.
40. Shahhosseini, M., Martinez-Feria, R. A., Hu, G., & Archontoulis, S. V. (2019). Maize yield and
41. nitrate loss prediction with machine learning algorithms. *Agronomy Journal*, *112*(3), 892–904.
42. Sharma, R. K., Kaur, J., Feng, G., Huang, Y., Kumar, C., Wang, Y., Sharma, S., Jenkins, J., & Dhillon, J. (2025). Maize and soybean yield prediction using machine learning methods: A
43. systematic literature review. *Discover Agriculture*, *3*(64), 1–35.
44. Shawon, Sarowar & Ema, Falguny & Mahi, Asura & Niha, Fahima & Hassan Tarif, Zubair.
45. (2024). Crop Yield Prediction Using Machine Learning: An Extensive and Systematic Literature Review. Smart Agricultural Technology. 10. 100718. 10.1016/j.atech.2024.100718.
46. Simon, L. C. F., & Palaoag, T. D. (2025). The application of machine learning techniques for

47. corn yield prediction and management: A systematic literature review. *Journal of Information Systems Engineering and Management*, *10*(43s), 1128–1139.
48. Som-ard, J., Suwanlee, S. R., Pinasu, D., Keawsomsee, S., Kasa, K., Seesanhao, N., Ninsawat,
49. S., Borgogno-Mondino, E., & Sarvia, F. (2024). Evaluating Sugarcane Yield Estimation in Thailand Using Multi-Temporal Sentinel-2 and Landsat Data Together with Machine-Learning Algorithms. Land, 13(9), 1481. https://doi.org/10.3390/land13091481Wang, X., Xu, Y., Huang, J., Yang, Z., & Zhang, Z. (2025). Knowledge-guided machine learning model with soil moisture for corn yield prediction under drought conditions. *arXiv*. https://arxiv.org/abs/2503.16328