AI in Agriculture: Techniques and Applications

Bhagirathi Halalli*, Deepa Garag**, Vinay Kumar V. & Kavita Hurakdli*******

DOI: https://doi.org/10.5281/zenodo.21206189

ABSTRACT

The role of agriculture in providing food security globally cannot be overstated, but it is associated with various complex issues, and agricultural researchers have found that Machine Learning (ML), as a subset of Artificial Intelligence (AI), is helpful in precision decision-making through learning from multiple agricultural data types. This paper reviews various state-of-the-art ML algorithms applied to crop yield, disease identification, soil/water management, and decision support systems. The paper summarizes various ML models, applications, and trends, as well as research challenges and gaps, aiming to offer insights to guide future research and applications.

KEYWORDS: *Artificial Intelligence (AI), Machine Learning (ML), Agriculture, Techniques, Applications*

Introduction

Agriculture is a major factor in global food security; however, the situation is becoming challenging owing to increased population rates, climate change, and water scarcity. The global population is expected to rise to nearly 10 billion by the year 2050; hence, food requirements will increase immensely without an increase in arable land.[1] This necessitates intelligent practices in increasing food production without compromising the environment.

Under traditional agricultural decision-making approaches, farm experience and conventional statistical models have shown several limitations in addressing the nonlinear relationships between weather conditions, soil properties, crop physiology, and management practices. Modern developments in acquiring climate, weather, and crop data using techniques like remote sensing, Internet of Things (IoT) technology, unmanned aerial vehicles, and geographic information systems (GIS) have helped in collecting agricultural data on large scales and in high resolution in real-time.[2]

At the same time, advances in computational techniques have made it possible to analyze the collected data. A sub-discipline of AI, machine learning (ML) has proven itself an important tool in extracting insightful information from diverse data sets in agriculture. Unlike traditional systems, which employ rule-based systems, ML has proven effective in efficiently learning complex relationships between input data sets and agriculture. This has ensured effective prediction using intelligent decision systems.[3]

Supervised learning using techniques such as Random Forest, Support Vector Machines, and regression algorithms has proven effective in agriculture, from crop yield modeling, soil fertility analysis, and crop classification, compared to traditional statistical methods.[4] Furthermore, deep learning methods, particularly CNNs, have registered significant achievements in various image-related agricultural practices that include plant disease recognition, pest recognition, and weed species detection by means of both leaf and field images.[5] This ensures that farmers are able to easily diagnose plant diseases early to avoid losses.

Moreover, various machine learning-based intelligent irrigation systems are designed by incorporating soil moisture sensors and weather forecasts to optimize water requirements for farmers, thereby significantly reducing water losses.[6] Using intelligent algorithms, crop price and demand forecasting can also be achieved in the realm of agricultural market intelligence.[7]

* **Bhagirathi Halalli,** Assistant Professor, Government First Grade College, Navanagar, Bagalkot.

** **Deepa Garag,** Assistant Professor, Government First Grade College, Navanagar, Raibag.

*** **Vinay Kumar V.,** Assistant Professor, Government First Grade College, Navanagar, Chintamani.

**** **Kavita Hurakdli,** Assistant Professor, Government First Grade College, Navanagar, Bagalkot.

All these applications discussed so far can be collectively considered as part of the base in precision agriculture, where different quantities of water, fertilizers, and pesticides need to be applied to the fields. In conclusion, the role of machine learning in the transition from traditional to modern and precision agriculture can be viewed as follows. Machine learning plays a crucial role in the transition from traditional to modern and precision agriculture, which is capable of coping with the complexities and improving the sustainability and efficiency of the agricultural sector. The application of modern sensing and communication technologies with machine learning is of vital importance to tackle the challenges of the agricultural sector in the present and future.

Background and Literature Review

Machine learning in agriculture has rapidly evolved from simple regression models to highly advanced deep learning techniques for numerous applications within the last decade. A review of comprehensive work in ML research at large categorizes the applications of ML in agriculture into domains like crop management-yield and disease, livestock management, soil, and water management, thus proving applicability in almost all agricultural needs. Moreover, research points out that deep neural networks, specifically Convolutional Neural Networks, allow for a better performance in image-related applications, such as disease detection or plant identification.

Table 1. Review of Literature

Authors & Year	Dataset	Problem Statement	Key Results / Accuracy	Research Gap
A. Morales, F. J. Villalobos (2023)[8]	DSSAT simulated wheat & sunflower data	ML models for crop yield forecasting	Random Forest outperformed ANN and linear models (RMSE~35–38% vs baseline ~42%)	Real weather forecast dependence not considered
S. Iniyan, V. A. Varma, C. T. Naidu (2023)[9]	Crop/environmental dataset	Crop yield prediction using ML	Feature engineering- based LSTM is the most efficient model, whichgives the best accuracy 86.3%,	Mixed environmental features likesoil type require expansion
S. P. Mohanty, D.P. Hughes, M. Salathé[10]	Plant disease image sets	Image-based disease detection using CNNs	High classification accuracy (>99%) demonstrated	Lack of cross- crop generalization
H. Fernando, T. Ha, K. A. Nketia, A. Attanayake & S. Shirtliffe (2024)[11]	Soil datasets	Satellite-based subfield canola yield prediction	Random Forest & XGB performed well Validation R^2 up to ~18-46	Needs validation across crop types
M. Radwan, A. A. Alhussan, A. Ibrahim, S. M. Tawfeek (2024)[12]	Multi-institution ML dataset	Potato leaf disease classification Optimized ML + feature selection effective classification	Accuracy 98.3%	Limitations in real-field varied lighting environments
J. Yao, S. N. Tran, S. Sawyer, et al. (2023)[13]	Review compendium	Leaf disease classification methods	Summarizes ML approaches for plant diseases	Real-time deployment challenges

Recent developments in machine learning (ML) and deep learning (DL) techniques have greatly enhanced the accuracy of agricultural prediction systems and plant health monitoring. Morales and Villalobos (2023) have shown the potential of RF models for DSSAT-simulated wheat and sunflower crop yield forecasting compared to ANN and linear regression models, yielding a minimum RMSE of 35-38%, though incorporating other weather changes was not considered. Iniyan, Varma, and Naidu (2023) have applied LSTM for crop yield prediction using a feature engineering approach, yielding a maximum accuracy of 86.3%, but environmental inputs like crop yields vary.

While different machine learning techniques have been applied for plant health monitoring, Mohanty, Hughes, and Salathé have shown classification accuracy above 99% using CNNs for detecting crop leaf diseases, though accuracy diminishes for other crop leaves. Fernando et al. (2024) have applied RF models for satellite-based canola crop yield forecasting, achieving a validation R-squared value within the range of 0.18-0.46, though crop validation is essential.

The authors Radwan et al. (2024) attained an accuracy of 98.3% for the classification of potato disease in leaves using optimized ML with feature selection. There is also the issue of testing ML classification under different lighting in the field. Additionally, Yao et al. (2023) [13] pointed out the issue of implementing ML classification in the detection of disease in plants, thereby further validating the practicality of the application of ML in agriculture.

Common Research Methodology of ML for Agriculture

In the diagram, different stages are well defined, and every stage signifies an essential step in the creation of an efficient agricultural-based machine learning system. Here, the process starts with Problem Definition, in which the agricultural-related problem is well identified, such as crop yield prediction, disease detection, and optimization of irrigation systems.

Data Collection involves multi-source data in agriculture: weather conditions, soil characteristics, and crop growth records, among IoT sensor readings, as well as imagery taken by either UAV or satellites. High-quality and diverse datasets are crucial for robust model performance. The Data Pre-processing includes cleaning of raw data, normalization, and transformation of the data. This includes handling missing values, removal of noise, image augmentation-in case of disease detection, time-series alignment-in case of forecasting tasks.

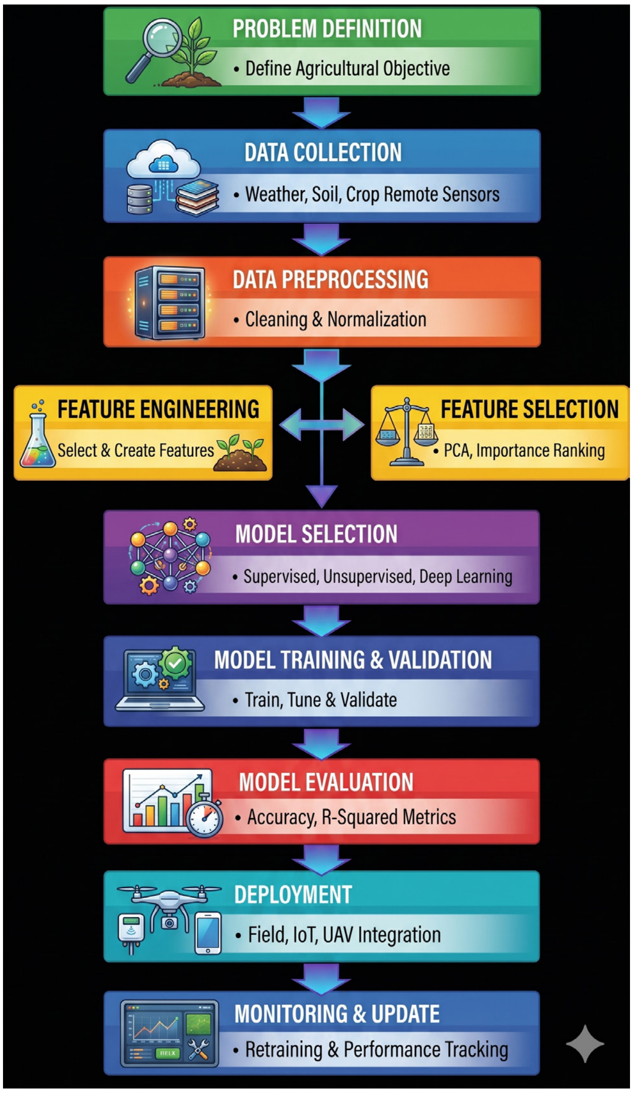

Figure 1. Machine Learning Pipeline for Agriculture

The pipeline then proceeds to Feature Engineering and Feature Selection, where meaningful agronomic indicators (e.g., NDVI, soil nutrients, rainfall trends) are extracted and irrelevant or redundant variables are removed using techniques like PCA or importance ranking. It is followed by model selection, where appropriate algorithms are chosen: supervised models, whether Random Forest or SVM, unsupervised methods such as K-means, deep learning models such as CNN and LSTM, or reinforcement learning for autonomous decision-making.

The Model Training & Validation training continues with a tuning stage, where it will be evaluated against the metrics: Accuracy, RMSE, and R^2. After validation is achieved, the model progresses to Deployment, which incorporates with the IoT devices, mobile apps, or UAV systems for real-time application. Finally, Monitoring & Updates provide for continuous retraining and performance tracking, adapting the system to seasonal and environmental changes that support the pathway to sustainable precision agriculture.

Machine Learning Techniques in Agriculture

There are three kinds of general research models for agricultural artificial intelligence based on their capability of handling different kinds of data. The data type ranges from structured tabular data, including soil and climate, to unstructured image data like leaf and field images.

- **Traditional ML Models:** These models find wide usage in establishing baselines or in agricultural data processing where structured data is of utmost importance.
- **Generic Models:** Random Forest (RF), which is commonly employed in crop yield prediction and soil modeling owing to its robustness against overfitting and its capacity to successfully tackle non-linear relationships between features.
- **Support Vector Machines (SVM):** It is commonly employed due to its proficiency in handling high-dimensional problems, especially for diseases classification and regression such as yield prediction.
- **Gradient Boosting Algorithms:** Algorithms like XGBoost, LightGBM, CatBoost are preferred for recent research works due to the accuracy and efficiency of the models in dealing with tabular data.
- **Deep Learning Architectures:** Deep learning architectures have a dominant role in tasks for which complex spatial or temporal patterns are essential.
- **Convolutional Neural Networks (CNNs):** Conventional CNN architectures like AlexNet, VGG, ResNet, and DenseNet are predominantly used for image-based classification of diseases and land cover extraction.
- **Recurrent Neural Networks & LSTMs:** The research focus of Long Short-Term Memory (LSTM) networks is discussed in the context of temporal dependencies or the effects of weather changes on growth in a season of cultivation.
- **Ensemble & Hybrid Architectures:** Also, as indicated in the present research literature, it has been shown that better results are obtained when an ensemble method is employed as against the “best” models.
- **Use of Stacked Ensemble Models:** Heterogeneous sub-learners were used in these models. For example, Deep Neural Networks and XGBoost/CatBoost algorithms were used to attain the best level of accuracy. In millet yield forecasting, for example, a level of 0.9910 R^2 is achieved.
- **CNN-Transformer Hybrids:** New models, such as AttCM-Alex and CRFormer, enable local feature extraction via CNNs and global context awareness via self-attention of Vision Transformers, with improved reliability in messy field settings.

Advanced Autonomous and Interpretability Frameworks

- **Reinforcement Learning (RL):** MDP-based agents, such as RL algorithms applied to Q-learning, PPO, and Soft Actor-Critic (SAC), are explored for autonomous applications involving task decisions such as precision irrigation and resource management.
- **Explainable AI (XAI) Frameworks:** With SHAP and LIME being incorporated into predictive models, it helps provide more insights into what drives yield predictions, e.g., rainfall or nitrogen.
- **Domain Adaptation (DA) Models:** Domain Adversarial Neural Networks (DANN) and Self-Supervised Learning (SSL) frameworks like SimCLR and MAE are used to enable the generalization of AI systems to new areas or seasons where labeled image data is insufficient.

Key Agricultural Applications of ML

One of the transformative contributions of machine learning (ML) in the current form of agriculture is precision farming, prediction, and automation. The uses of ML are varied and pertain to crop production, plant health management, water and soil management, and decision support systems at the farms.

- ***Crop Yield Prediction:*** One of the most commonly explored agricultural application of ML is crop yield prediction. In the ML approach for crop prediction, past yields of the crops are incorporated as inputs along with environmental factors like temperature, humidity, rainfall, and soil types. The ML algorithms used for prediction include Linear Regression, Support Vector Machine (SVM), and popular data mining algorithms known as Random Forest and Gradient Boosting or XGboost. Another popular approach is the Long Short-term Memory (LSTM) network. The accurate prediction of crop yields is helpful in making better decisions and thus optimizing the usage of resources.

- ***Disease and Pest Detection:*** Detection of disease and pests in plants has been made more efficient using ML-based image analysis techniques. It has been found that most researchers have successfully employed the CNN approach for the classification of plant leaf diseases using images with high classification accuracy. Real-time monitoring in fields has been achieved through the use of mobile application-based systems and UAV images processed with ML-based techniques. Improved techniques like GAN assist in improving image dataset diversity and efficiency even under changing lighting and other environmental conditions.
- ***Soil and Water Management:*** ML models predict soil properties, estimate moisture levels, and recommend irrigation schedules using sensor and satellite data. These approaches optimize fertilizer and water use, improving sustainability and conserving natural resources.
- ***Decision Support Systems:*** It can be done through the use of AI-powered decision support technologies with site-specific recommendations regarding crop choices, nutrient management, and farming practices with the help of climate, soil, as well as historical data.
- ***Precision Farming & Variable Rate Technology (VRT):*** ML can help learn site-specific crop management using data collected from the spatial variability offered by satellites, UAV, and IoT sensors. Reduction in input costs for fertilizers, pesticides, and irrigation.
- ***Weed Detection & Smart Spraying:*** The vision model can automatically identify weed presence in real-time, thereby allowing smart spraying of herbicide to occur, which brings concepts of sustainability to entirely new heights.
- ***Livestock Monitoring and Health Management:*** ML models process sensor data, images, and behavioral patterns to monitor health, diagnose disease, forecast milk production, and optimize nutrition. IoT-enabled wearables and machine learning techniques enhance animal productivity and well-being.
- ***Smart Irrigation Systems:*** Reinforcement learning and predictive ML models can be used to optimize the irrigation schedule considering factors like soil moisture, weather forecasts, and crop stages.
- ***Post-Harvest Quality Assessment:*** ML methods evaluate the quality of the crops using image processing and sensor inputs for grading and shelf life prediction and storage management.
- ***Climate Risk and Weather Forecasting:*** Machine Learning models can forecast extreme weather events, droughts, and the effect of climate variability on crops.
- ***Agricultural Robotics and Automation:*** Autonomous tractors, harvesting equipment, and drone-based systems employ ML techniques for navigation, detection of objects, as well as instantaneous decisions.
- ***Supply Chain and Market Price Forecasting:*** The machine learning models study historical market, demand, and supply chain trends and patterns, thereby giving forecasts on crop prices and supply chain activities.

Challenges and Limitations

Despite the progress made in agricultural ML, there are still challenges:

- **Data Scarcity and Quality:** Available data varies in quality, which is a limitation for applying models.
- **Interpretability:** Complex models such as deep neural networks tend to be non-transparent.
- **Scalability:** High computational needs are a barrier in low-resource environments.
- **Integration Challenges:** Seamless integration of ML into current agricultural workflows is still an important challenge.

Future Research Directions

Future research should prioritize:

- Explainable AI (XAI): Techniques for making decisions transparent to end users.
- Low-Cost and Edge Deployment: Optimizing ML models for resource-constrained platforms.
- Multimodal Data Integration: Combining Satellite, Sensor, and Field Data for Richer Insights.
- Region-Specific Models: Models designed for specific agro-ecological regions and types of farming.

Conclusion

The focused research papers, taken as a group, establish the prospect of using the application of the aforementioned machine learning approaches to improve agricultural productivity significantly. In that regard, it is found that high prediction accuracy in terms of future productivity forecasts and diseases is an indicator of the reliability of data-driven agricultural systems.

Notwithstanding the opportunities created by the text-mined research papers, limitations such as restricted applicability across different crops, inadequate incorporation of dynamic variables based on the environment, or issues concerning real-time applicability remain to be addressed. As such, it is recommended that future research should consider the applicability of such approaches to different crops, live implementation with the incorporation of real-time weather forecasts, or model implementation despite factors such as the environment.

References:

1. Food and Agriculture Organization of the United Nations, The Future of Food and Agriculture – Trends and Challenges, FAO, Rome, 2017. [Online]. Available: https://www.fao.org
2. A. Kamilaris and F. X. Prenafeta-Boldú, “Deep learning in agriculture: A survey,” Computers and Electronics in Agriculture, vol. 147, pp. 70–90, 2018. doi: 10.1016/j.compag.2018.02.016
3. K. G. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, “Machine learning in agriculture: A review,” Sensors, vol. 18, no. 8, pp. 1–29, 2018. doi: 10.3390/s18082674
4. A. Chlingaryan, S. Sukkarieh, and B. Whelan, “Machine learning approaches for crop yield prediction and nitrogen status estimation,” Sensors, vol. 18, no. 11, pp. 1–27, 2018. doi: 10.3390/s18113706
5. S. P. Mohanty, D. P. Hughes, and M. Salathé, “Using deep learning for image-based plant disease detection,” Frontiers in Plant Science, vol. 7, pp. 1–10, 2016. doi: 10.3389/fpls.2016.01419
6. M. Smith, J. Mullins, and R. Smith, “Machine learning methods for irrigation decision support,” Agricultural Water Management, vol. 216, pp. 1–10, 2019. doi: 10.1016/j.agwat.2019.01.004
7. Y. Liu, Y. Wang, and C. Yang, “Short-term agricultural price forecasting using LSTM networks,” IEEE Access, vol. 8, pp. 1–12, 2020. doi: 10.1109/ACCESS.2020.2987654.
8. Morales A, Villalobos FJ. Using machine learning for crop yield prediction in the past or the future. Front Plant Sci. 2023 Mar 30;14:1128388. doi: 10.3389/fpls.2023.1128388. PMID: 37063228; PMCID: PMC10097960
9. D. J. Reddy and M. R. Kumar, “Crop Yield Prediction using Machine Learning Algorithm,” 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2021, pp. 1466-1470, doi: 10.1109/ICICCS51141.2021.9432236.
10. Mohanty SP, Hughes DP, Salathé M. Using Deep Learning for Image-Based Plant Disease Detection. Front Plant Sci. 2016 Sep 22;7:1419. doi: 10.3389/fpls.2016.01419. PMID: 27713752; PMCID: PMC5032846.
11. Fernando, H., Ha, T., Nketia, K.A. et al. Machine learning approach for satellite-based subfield canola yield prediction using floral phenology metrics and soil parameters. Precision Agric 25, 1386–1403 (2024). https://doi.org/10.1007/s11119-024-10116-1
12. Radwan, M., Alhussan, A.A., Ibrahim, A. et al. Potato Leaf Disease Classification Using Optimized Machine Learning Models and Feature Selection Techniques. Potato Res. 68, 897–921 (2025). https://doi.org/10.1007/s11540-024-09763-8
13. Yao, J., Tran, S.N., Sawyer, S. et al. Machine learning for leaf disease classification: data, techniques and applications. Artif Intell Rev 56 (Suppl 3), 3571–3616 (2023). https://doi.org/10.1007/s10462-023-10610-4