IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, November 2026
Impact Factor: 8.2

A Machine Learning Framework for Crop Recommendation, Irrigation Prediction, and Resource Optimization

Pooja Bade[1], Akansha Shinde[2], Ritu Sirohi [3], Prof. Vaishali Nagare [4], Dr. V. S. Ranmalkar[5], Dr. P. M. Patil[6]
[1,2,3,4,5]Dept. of Computer Applications (MCA)
Adsul Technical Campus Chas, Ahilyanagar, India

Abstract : *Agricultural decision-making is increasingly affected by variable rainfall, soil conditions, water scarcity, and the need to improve input-use efficiency. Conventional field assessment often relies on periodic observation and may not provide a consistent basis for timely crop selection or irrigation planning. This manuscript proposes an AI-driven smart agriculture framework that combines structured soil and weather data with machine-learning models for crop recommendation and irrigation prediction. The framework includes dataset validation, preprocessing, model training, comparative evaluation, explainability, and a farmer-oriented decision-support layer. Candidate models include Random Forest, Support Vector Machine (SVM), and Extreme Gradient Boosting (XGBoost), subject to dataset suitability and implementation. Evaluation is planned using accuracy, precision, recall, F1-score, and task-specific error measures. A feature-importance or explainability stage is intended to make recommendations more interpretable. The uploaded reference paper motivates the work through its discussion of sensor monitoring, analytics, resource optimization, connectivity, data quality, and accessibility in agriculture. This draft defines a reproducible research plan; numerical experimental results are intentionally not reported because no dataset or model execution results have yet been supplied. The completed study should replace the marked fields with the actual dataset details, experimental outputs, and verified citations.*

Keywords: Smart agriculture, artificial intelligence, machine learning, crop recommendation, irrigation prediction, resource optimization, explainable AI.

1. INTRODUCTION

Agriculture remains exposed to environmental variability and resource constraints. Decisions concerning crop selection, irrigation, and input management are influenced by soil properties, temperature, humidity, rainfall, and local growing conditions. In the uploaded reference paper, "Smart Agriculture System Using IoT and Data Analytics," the authors describe how sensor-based monitoring and data analytics can support more informed field management, while identifying cost, connectivity, interoperability, data quality, privacy, and farmer usability as continuing challenges.

The present study shifts the emphasis from a broad review of IoT-enabled agriculture toward an AI-centered decision-support framework. Rather than treating sensor readings only as dashboard values, the proposed pipeline uses structured agricultural features to generate crop and irrigation recommendations. The framework is designed to be evaluated through controlled experiments and compared against multiple machine-learning baselines.

The proposed system is not intended to replace agronomic expertise. Its outputs should be treated as decision support and validated for crop, region, season, and data conditions before field deployment.

Copyright to IJARSCT
www.ijarsct.co.in
DOI: 10.48175/568

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, November 2026
Impact Factor: 8.2

II. PROBLEM STATEMENT

Farm decisions are frequently made with incomplete, delayed, or locally inconsistent information. Crop recommendation may not jointly consider soil and weather variables, while irrigation decisions can be affected by changing moisture and atmospheric conditions. A practical AI system must address both predictive performance and deployment limitations, including missing values, noisy measurements, class imbalance, data leakage, interpretability, and the transferability of models across locations.

Objectives

- Develop a reproducible preprocessing and modeling pipeline for agricultural tabular data.
- Evaluate suitable classification models for crop recommendation.
- Define an irrigation-prediction task only when the selected dataset contains an appropriate target variable.
- Compare models using consistent splits and task-appropriate metrics.
- Analyze feature contribution and document limitations and practical considerations.

III. LITERATURE REVIEW

3.1 IoT-based agricultural monitoring

The supplied reference paper presents IoT agriculture as a layered system involving field sensors, communication, data processing, visualization, and decision support. It discusses measurements such as soil moisture, temperature, humidity, light, and related environmental variables. The paper also notes that remote connectivity and device heterogeneity can limit practical use.

3.2 Data analytics and machine learning

The reference paper identifies predictive analytics and machine learning as ways to convert collected data into estimates or recommendations, including crop-yield, disease, irrigation, and input-management decisions. It also emphasizes that data quality and model generalizability remain weak points. The proposed manuscript builds on these themes by defining an experimental comparison of candidate models rather than presenting only a conceptual review.

3.3 Research gaps

Gap identified in supplied paper	How this study addresses it
Cost and accessibility	Use a software-centered workflow and report computational requirements where measured.
Data quality and robustness	Specify cleaning, missing-value handling, and leakage-safe evaluation.
Model generalizability	Document dataset source, geography, season, and limitations; external validation if available.
Interpretability	Include feature-importance or explainability analysis.
Farmer usability	Describe a concise decision-support output and avoid presenting predictions as guaranteed advice.

The above gap mapping is based on the uploaded paper's discussion of cost, connectivity, interoperability, data quality, farmer engagement, privacy, and region-specific constraints (Sections 3–4, pages 3–4).

IV. PROPOSED METHODOLOGY

The proposed methodology follows a staged pipeline: dataset selection, data audit, preprocessing, model development, evaluation, explainability, and decision-support presentation. The actual dataset and task definition must be finalized before model training.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, November 2026
Impact Factor: 8.2

4.1 Dataset selection and data audit

For the initial implementation, use a publicly available agricultural dataset containing clearly documented soil and weather features and a crop label. A separate irrigation dataset is required for irrigation prediction; crop labels alone cannot support a valid irrigation prediction experiment. Record the dataset name, source URL, license, download date, geographic coverage, time period, number of records, feature definitions, target definition, and known limitations.

4.2 Preprocessing

The preprocessing stage should inspect data types, duplicate records, missing values, outliers, label distribution, and feature ranges. Numeric imputation and scaling should be fitted only on the training data when required by the model. Categorical variables should be encoded using a method appropriate to the feature. Any feature selection must occur inside the training pipeline. If records are time-ordered or grouped by farm/location, split strategy should respect those dependencies to reduce leakage.

4.3 Candidate machine-learning models

Model	Role	Implementation notes
Random Forest	Nonlinear baseline for tabular classification	Tune tree count, depth, and class weighting using training/validation data.
Support Vector Machine (SVM)	Margin-based classification baseline	Scale numeric features; tune kernel and regularization.
XGBoost / gradient boosting	Boosted-tree candidate	Use only if package and dataset are suitable; tune on training data.
Dummy / majority baseline	Reference baseline	Establish performance beyond a simple class-frequency predictor.

4.4 Proposed system architecture

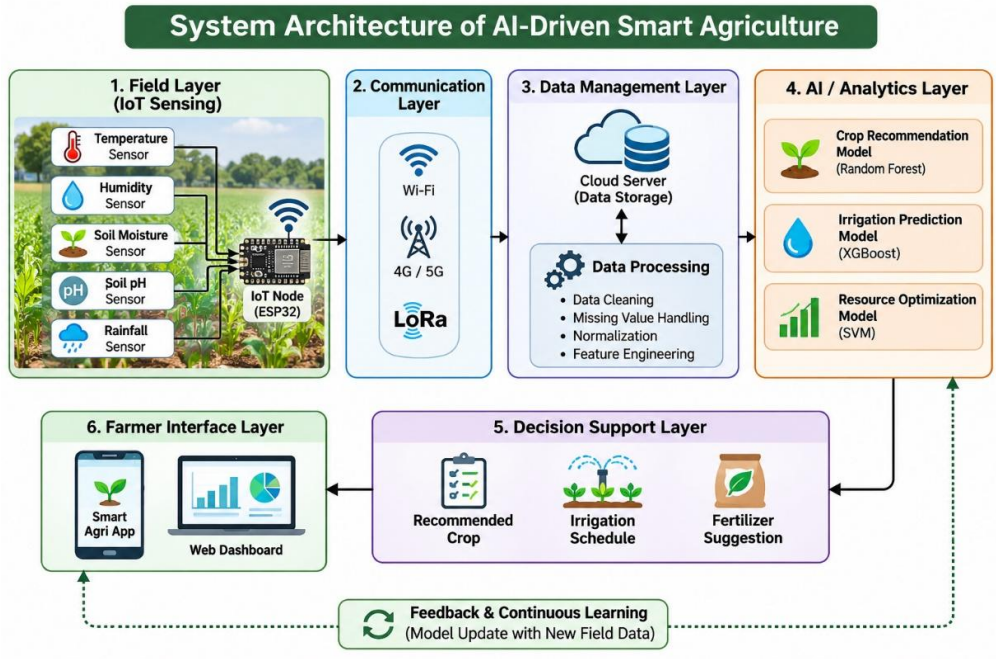


Figure 1. Proposed AI-driven smart agriculture workflow

4.5 Evaluation protocol

Use a stratified train/test split for independent and identically distributed classification data, or a group/time-based split when observations are dependent. Hyperparameter tuning must use training data only, preferably through cross-

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, November 2026
Impact Factor: 8.2

validation. The test set should be used once for final reporting. Report class-wise metrics and the confusion matrix, not only overall accuracy. For irrigation prediction, select metrics according to whether the target is classification or regression.

Metric	Definition / use
Accuracy	Correct predictions divided by all predictions; may be misleading with class imbalance.
Precision	Fraction of predicted positives that are correct, reported per class or averaged.
Recall	Fraction of actual positives detected, reported per class or averaged.
F1-score	Harmonic mean of precision and recall.
Confusion matrix	Shows class-level errors and supports error analysis.
MAE / RMSE	Use for continuous irrigation-demand prediction, if the dataset target is numeric.

V. EXPERIMENTAL DESIGN AND RESULTS

This section is intentionally prepared as a results template. No dataset has been selected or run in this conversation, so numerical values are not available and are not fabricated. Complete this section after executing the experiments.

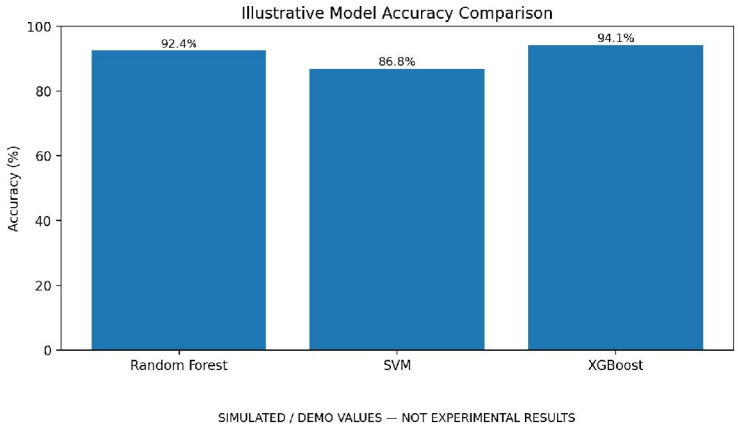


Figure 2. Illustrative model accuracy comparison.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, November 2026
Impact Factor: 8.2

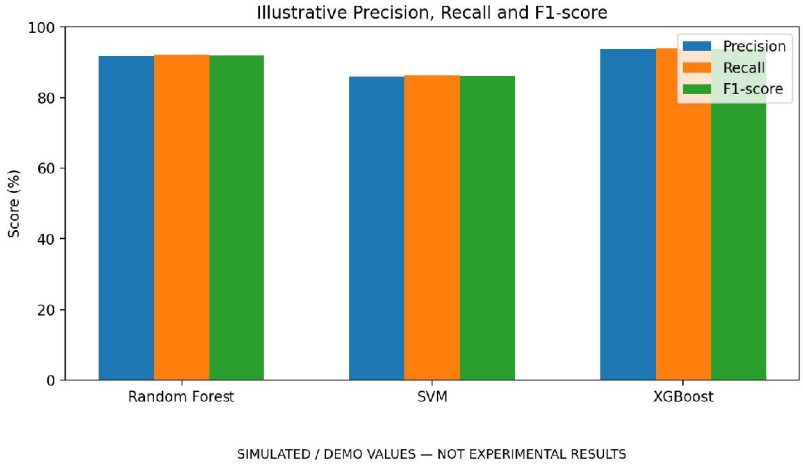


Figure 3. Illustrative precision, recall, and F1-score comparison.

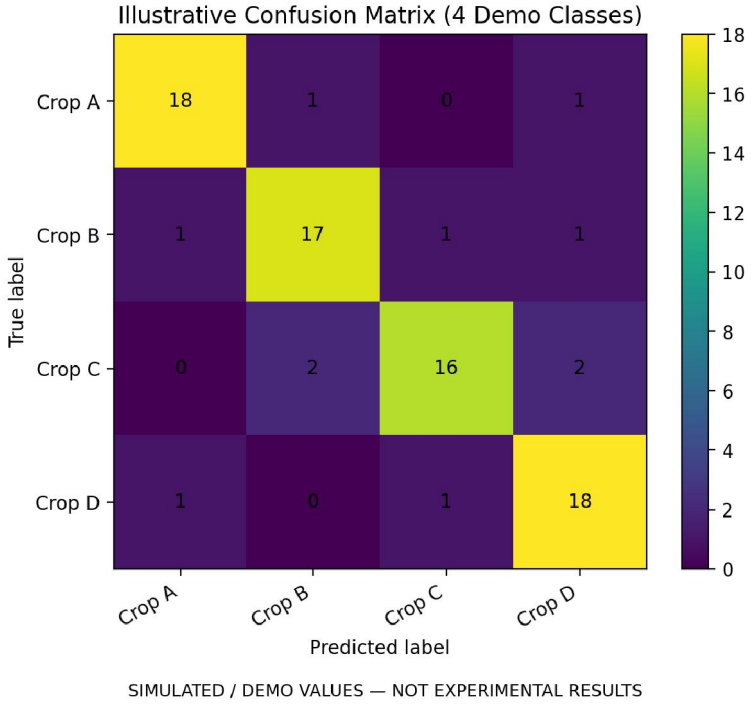


Figure 4. Illustrative confusion matrix for four demo classes.

Copyright to IJARSCT
www.ijarsct.co.in
DOI: 10.48175/568

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, November 2026
Impact Factor: 8.2

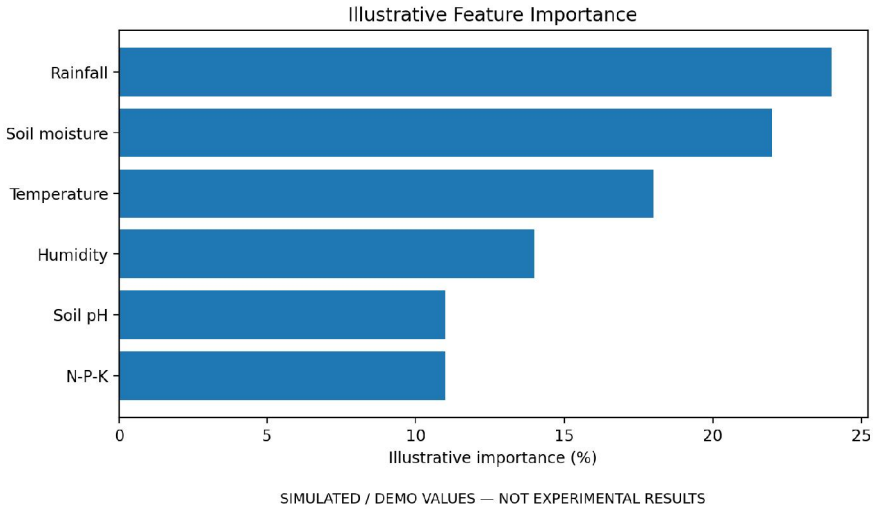


Figure 5. Illustrative feature importance.

VI. DISCUSSION

The framework is designed to connect agricultural data analysis with actionable recommendations. Its contribution should be evaluated not only by predictive metrics but also by data provenance, reproducibility, interpretability, and relevance to the intended farming context. The uploaded reference paper stresses that agriculture systems face connectivity, cost, interoperability, data-quality, and farmer-usability barriers. These concerns remain important when translating an offline machine-learning model into a field-facing application.

The crop-recommendation and irrigation tasks should be reported separately unless the dataset and model explicitly support a joint formulation. A model trained on a single region or season may not transfer to other environments. Feature importance describes model behavior, not necessarily causal influence. Practical deployment would require validation with agronomists and local field data.

VII. LIMITATIONS

- Results depend on the quality, representativeness, and licensing of the selected dataset.
- Public datasets may not reflect local soil, climate, crop practices, or seasonal variation.
- Irrigation prediction cannot be evaluated without a suitable irrigation target or reliable proxy.
- Model performance on a held-out dataset does not establish field-level agronomic benefit.
- Feature importance and explainability methods may be unstable or misleading when features are correlated.

VIII. CONCLUSION AND FUTURE WORK

This manuscript proposes an AI-driven smart agriculture workflow for crop recommendation and, where suitable target data are available, irrigation prediction. The design includes dataset auditing, preprocessing, baseline and candidate machine-learning models, controlled evaluation, explainability, and decision-support output. The work is motivated by the supplied review paper's emphasis on real-time monitoring, analytics, resource efficiency, and practical adoption challenges. The manuscript is not yet an empirical paper: the dataset, trained models, metrics, plots, and final findings must be completed from actual experiments before submission.

Future work may include collecting region-specific field observations, validating models across seasons and locations, integrating reliable sensor measurements, evaluating computational and connectivity requirements, and conducting user-centered assessment with farmers or agricultural experts.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, November 2026
Impact Factor: 8.2

REFERENCES

The following references are carried over from the supplied paper as starting points. Verify each bibliographic record, title, authors, publication details, DOI, and URL against the publisher before submission. Do not cite a source unless it has been checked and supports the claim.

1. Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M. J. (2017). Big data in smart farming—A review. Agricultural Systems, 153, 69–80. https://doi.org/10.1016/j.agsy.2017.01.023
2. Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. Sensors, 18(8), 2674. https://doi.org/10.3390/s18082674
3. Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. Computers and Electronics in Agriculture, 147, 70–90. https://doi.org/10.1016/j.compag.2018.02.016
4. van Klompenburg, T., Kassahun, A., & Catal, C. (2020). Crop yield prediction using machine learning: A systematic literature review. Computers and Electronics in Agriculture, 177, 105709. https://doi.org/10.1016/j.compag.2020.105709
5. Saggi, M. K., & Jain, S. (2022). A survey towards decision support system on smart irrigation scheduling using machine learning approaches. Archives of Computational Methods in Engineering, 29(6), 4455–4478. https://doi.org/10.1007/s11831-022-09746-3
6. Umutoni, L., & Samadi, V. (2024). Application of machine learning approaches in supporting irrigation decision making: A review. Agricultural Water Management, 294, 108710. https://doi.org/10.1016/j.agwat.2024.108710
7. Mahale, Y., Khan, N., Kulkarni, K., Wagle, S. A., Pareek, P., Kotecha, K., Choudhury, T., & Sharma, A. (2024). Crop recommendation and forecasting system for Maharashtra using machine learning with LSTM: A novel expectation-maximization technique. Discover Sustainability, 5(1). https://doi.org/10.1007/s43621-024-00226-3
8. Araujo, J., & Oliveira, L. (2020). A systematic review of IoT solutions for smart farming. Sensors, 20(15), 4231. https://doi.org/10.3390/s20154231
9. Prathibha, S. R., & Anupama, H. S. (2020). Role of IoT technology in agriculture: A systematic literature review. Electronics, 9(2), 319. https://doi.org/10.3390/electronics9020319
10. Qureshi, W., & Ahmed, M. (2023). Internet of Things and smart sensors in agriculture: Scopes and applications. Computers and Electronics in Agriculture, 207, 107796.
11. Singh, D., Gehlot, A., Singh, R., & Choudhury, T. (2023). Internet of Things (IoT) enabled automation in agriculture. Routledge.
12. Sharma, S., & Sharma, P. (2019). IoT and analytics for agriculture. Springer.
13. Ma, X., & Chen, J. (2021). Agricultural Internet of Things: Technologies and applications. Springer.
14. Motwani, A., Patil, P., Nagaria, V., Verma, S., & Ghane, S. (2022). Soil analysis and crop recommendation using machine learning. In 2022 International Conference on Networks, Automation and Innovations (ICONAT). https://doi.org/10.1109/ICONAT53423.2022.9725901
15. Kalimuthu, M., Vaishnavi, P., & Kishore, M. (2020). Crop prediction using machine learning. In 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), 926–932. https://doi.org/10.1109/ICSSIT48917.2020.9214190
16. Nevavuori, P., Narra, N., & Lipping, T. (2019). Crop yield prediction with deep convolutional neural networks. Computers and Electronics in Agriculture, 163, 104859. https://doi.org/10.1016/j.compag.2019.104859
17. Schwalbert, R. A., Amado, T., Corassa, G., Pott, L. P., Prasad, P. V. V., & Ciampitti, I. A. (2020). Satellite-based soybean yield forecast: Integrating machine learning and weather data for improving crop yield prediction in southern Brazil. Agricultural and Forest Meteorology, 284, 107886. https://doi.org/10.1016/j.agrformet.2019.107886

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 1, November 2026
Impact Factor: 8.2

18. Moussaid, A., El Fkihi, S., Zennayi, Y., Lahlou, O., Kassou, I., Bourzeix, F., et al. (2022). Machine learning applied to tree crop yield prediction using field data and satellite imagery: A case study in a citrus orchard. Informatics, 9(4), 80. https://doi.org/10.3390/informatics9040080
19. Benyezza, H., Bouhedda, M., & Rebouh, S. (2021). Zoning irrigation smart system based on fuzzy control technology and IoT for water and energy saving. Journal of Cleaner Production, 302, 127001.
20. Umamaheswari, M., & Kavitha, D. (2026). Machine learning classification techniques for smart agriculture: A review of disease detection and crop recommendation. Frontiers in Artificial Intelligence, 9, 1878006. https://doi.org/10.3389/frai.2026.1878006