International Journal of Emerging Research in Science, Engineering, and Management
Vol. 2, Issue 1, pp.175-180, January 2026.
www.ijersem.com eISSN – 3107-9075

AgroIntelligence: A Farmer Centric ML Solution for Smart Agriculture

[1]Pathipati Haarika Naidu, [2]Chanumallu Charan Jetty, [3]D. Janani, [4]Mothukuri Mounika,[5]Vuggam Krishna Vamsi, [6]Muktha Nithish

[1,2,4,5,6]UG Scholar, Department of CSM, Siddartha Institute of Science and Technology, Puttur, India.
[3]Assistant Professor, Department of CSE, Siddartha Institute of Science and Technology, Puttur, India.

[1]haarikanaidup@gmail.com, [2]ccharanccharan16@gmail.com, [3]janani.jaanu38@gmail.com,
[4]mothukurimounika22@gmail.com, [5]vamsivuggam85@gmail.com, [6]mukthanithish@gmail.com

Abstract: Unpredictable climate patterns and declining soil health have intensified agricultural distress in India, posing serious challenges to food security and farmer livelihoods. This paper introduces AgroIntelligence, a machine-learning– based framework for real-time crop and fertilizer recommendation in precision agriculture. The system employs an optimized Random Forest classifier to model soil– climate relationships and uses K-Nearest Neighbors (KNN) imputation to handle missing soil nutrient data. To improve rural applicability, the framework supports two input modes: a Manual Mode for direct farmer input and an Auto Mode that leverages geolocation to obtain district-level NPK, pH, rainfall, and real-time weather information through API integration. A bilingual AI chatbot is incorporated to enhance accessibility for diverse farmer communities. Experimental evaluation on a custom dataset of 18,240 records from 26 districts of Andhra Pradesh reports a classification accuracy of 98%. The proposed framework supports sustainable precision agriculture by reducing unnecessary crop and fertilizer input usage and enabling data-driven agricultural planning.

Keywords: Smart Agriculture, Crop Recommendation, Fertilizer Recommendation, Weather integration, Local Language Support.

1 INTRODUCTION

Agriculture plays a central role in India's economy, employing a significant portion of the workforce and sustaining rural livelihoods [1]. However, increasing challenges such as unpredictable climate patterns and declining soil health have contributed to farmer distress and reduced agricultural productivity [2], [3]. While laboratory-based soil testing offers accurate nutrient analysis, delayed report availability and limited accessibility in rural regions reduce its effectiveness for timely crop planning decisions [4]. These limitations highlight the need for automated, interpretable, and low- bandwidth agricultural decision-support systems that can be effectively deployed in rural environments [5].

Several existing machine learning–based crop recommendation systems depend heavily on manual data entry and relatively simple predictive models, which restrict their scalability in large multi-class recommendation scenarios involving complex, non-linear soil–climate interactions [6], [7]. Among traditional machine learning techniques, Random Forest—a tree- based ensemble learning method—has demonstrated strong robustness in capturing non-linear relationships between soil parameters and weather conditions [8]. To address usability and contextual relevance gaps, AgroIntelligence proposes a dual-input recommendation framework comprising a Manual Mode for direct farmer input and an Auto Mode that leverages geolocation to retrieve district-level NPK, pH, rainfall, and real-time weather data through API integration [9], [10].

The framework further employs K-Nearest Neighbors (KNN) imputation to handle missing soil nutrient attributes [11] and integrates a bilingual AI chatbot within a lightweight, low-bandwidth web interface to enhance accessibility, advisory clarity, and adoption among rural farmers [12]. Large-scale evaluation conducted on a custom agricultural dataset confirms the reliability of the proposed classifier, demonstrating its feasibility for scalable and interpretable smart-agriculture advisory in rural settings. Key contributions include:

- A dual-input crop recommendation framework supporting both Manual input and geolocation-based Auto Mode.
- Integration of district-level soil parameters with real-time weather data through API-based retrieval.
- Incorporation of a bilingual AI chatbot to improve accessibility and farmer adoption.
- Large-scale evaluation on 18,240 records from 26 districts of Andhra Pradesh, achieving 98% classification accuracy using Random Forest.
- Deployment of a scalable, interpretable, and low- bandwidth web-based decision-support system for rural agriculture.

2 LITERATURE REVIEW

Crop recommendation systems have evolved significantly to support agricultural decision-making processes [1], [2]. Early advisory approaches were primarily based on static crop calendars and seasonal schedules developed by agricultural institutions, providing generalized sowing guidelines without real-time climate adaptability [3]. Such systems often become unreliable during irregular or delayed monsoon conditions, adversely affecting planting schedules and overall crop productivity [4].

With the advancement of machine learning, data-driven analysis of soil and environmental parameters has gained prominence through models such as Decision Trees and Random Forest. Ensemble learning techniques, particularly Random Forest, improve generalization capability and reduce overfitting when compared to single-tree models [5], [6]. While traditional classifiers perform well in binary or low-class prediction tasks, their effectiveness diminishes in complex multi-class crop recommendation scenarios involving a large number of crop categories [7], [8]. Although deep learning approaches can achieve high predictive accuracy, they require substantial computational resources, longer training durations, and complex inference pipelines, making them less suitable for rural and edge-based agricultural deployments [9].

Despite these advances, many existing crop advisory platforms rely on rigid manual entry of soil parameters, lack real-time weather integration, exclude fertilizer recommendation modules, and provide support only in English, thereby limiting accessibility and adoption among rural farming communities [10], [11]. To overcome these challenges, AgroIntelligence introduces a dual-input framework that combines Manual input with a geolocation-enabled Auto Mode for retrieving district-level soil enhance inclusivity, usability, and practical adoption in rural agricultural settings [12], [13].

3 PROPOSED SYSTEM

This section describes the proposed AgroIntelligence framework, an ML-driven system designed for real-time crop and fertilizer recommendation in geolocation-aware precision agriculture. The framework supports both Manual and Auto input modes, applies K-Nearest Neighbors (KNN) imputation and Min–Max scaling for data preprocessing, and performs multi-class crop classification using an optimized Random Forest model. The overall system architecture and data flow are illustrated in Fig. 1.

3.1. System Overview

The system operates in four stages:

i) Data Acquisition,
ii) Preprocessing & Imputation,
iii) Multi-Class Classification, and
iv) Recommendation & Alerts.

The four stages are implemented through different layers. They are presented below.

a. Input Layer: Accepts soil (N, P, K, pH) and environment (temperature, humidity, rainfall) inputs. Supports Manual Mode for direct entry and Auto Mode for geolocation-based district soil enrichment and live weather retrieval via REST APIs [3].
b. Preprocessing: Uses KNN imputation for missing soil values and Min-Max scaling (0– 1) for normalization.
c. Inference: An optimized Random Forest (500 trees, depth 30) performs multi-class crop prediction.
d. Decision Support: The Suggestion Engine estimates fertilizer dosage and retrieves relevant government schemes, reducing input wastage.

This modular design ensures scalability, interpretability, and low-bandwidth usability, making the framework suitable for rural smart-agriculture deployment.

3.2. Data Preprocessing Algorithms

The Agricultural datasets are often affected by noise and missing values. To address these challenges, two preprocessing techniques are employed.

i) K-Nearest Neighbors (KNN) Imputation: Missing soil attributes are estimated using inverse distance weighting based on the k-nearest neighbors. This approach prevents the exclusion of incomplete farmer inputs and ensures that a complete feature vector is provided to the classifier.
ii) Feature Scaling: All numerical attributes are normalized using Min Max scaling: Normalization ensures balanced learning across features with different value ranges, such as rainfall and soil pH, preventing dominance by high-magnitude attributes [4].

International Journal of Emerging Research in Science, Engineering, and Management
Vol. 2, Issue 1, pp.175-180, January 2026.
www.ijersem.com eISSN – 3107-9075

3.3. Classification Using Random Forest

The Random Forest classifier models complex non- linear soil–weather interactions using ensemble learning principles based on bagging and majority voting [5], [14].

a. Bagging: Multiple bootstrap samples are generated from the training dataset to construct diverse decision trees.
b. Split Criterion: Information Gain is maximized through entropy reduction. Since urea contains 46% nitrogen by weight, this formulation converts nutrient gaps into actionable fertilizer quantities, enabling practical field-level adoption and efficient input utilization [15].
c. Voting: Final crop prediction is obtained using majority voting across all trees. This ensemble-based approach provides robust and interpretable predictions for complex, large-scale multi- class crop recommendation tasks.
d. Rule-based Fertilizer Logic: Following crop prediction, soil nutrient deficiencies are translated into fertilizer dosage recommendations using rule-based logic. The required quantity of urea based on soil nitrogen deficiency is estimated.

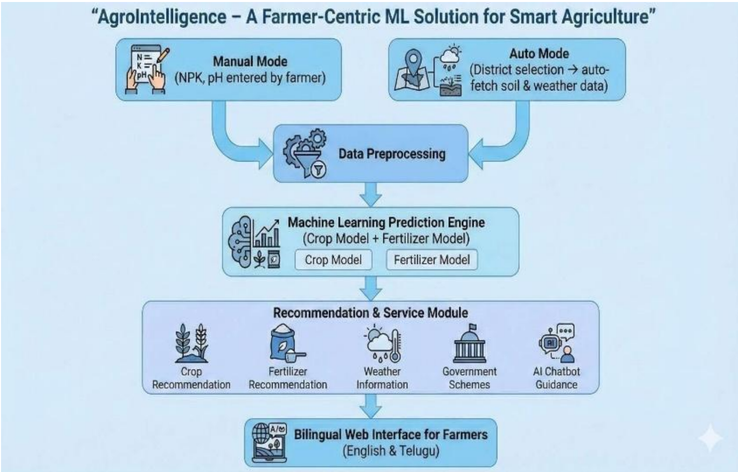

Fig. 1. Architecture of the proposed AgroIntelligence System

4 RESULTS AND DISCUSSION

4.1. Experimental Setup

The proposed framework is evaluated on a custom dataset of 18,240 agricultural records collected from districts of Andhra Pradesh. An 80:20 train–test split is used, along with 10-fold cross-validation to ensure robustness. Multi- class crop classification is performed using an optimized Random Forest model configured with 500 trees, a maximum depth of 30, and standard regularization parameters to maintain model stability and generalization.

4.2. Quantitative Evaluation

The Random Forest classifier is compared with Logistic Regression, Support Vector Machines, and Decision Tree models under identical experimental conditions. Performance is assessed using Accuracy, Precision, Recall, and F1-score. As shown in Table 1, Random Forest achieves the highest performance, with a classification accuracy of 98%, consistently outperforming the baseline models across all metrics. This result highlights the effectiveness of ensemble learning in capturing complex soil–climate interactions and confirms the suitability of the proposed framework for real-time, multi-class crop recommendation.

Table 1. Model Performance Comparison of AgroIntelligence System

Model	Accuracy	Precision	Recall	F1 score
Logistic Regression	84.5%	0.84	0.83	0.83
Decision Tree	91.2%	0.92	0.91	0.91
Support Vector Machines	96.8%	0.96	0.96	0.96
Random Forest	98%	0.98	0.98	0.98

4.3. Training and Validation Analysis

Fig. 2 shows training and validation accuracy. This exhibits fast convergence in early trees and stable generalization at 500 trees, validating scalability for multi-class crop classification.

International Journal of Emerging Research in Science, Engineering, and Management
Vol. 2, Issue 1, pp.175-180, January 2026.
www.ijersem.com eISSN – 3107-9075

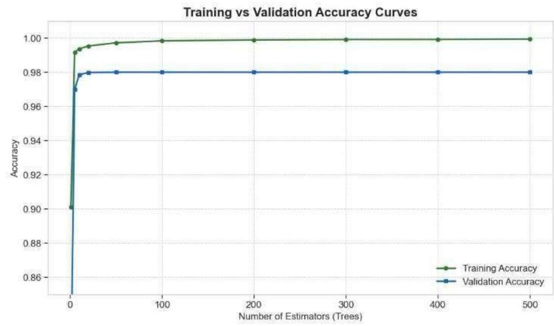

Fig. 2. Training and Validation Accuracy Curves

Fig. 3. Log-loss shows a steep decline for initial trees and early stabilization, indicating improved prediction confidence and well-calibrated model inference.

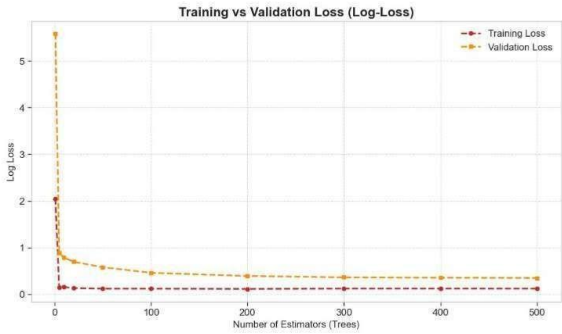

Fig. 3. Training and Validation Loss Curve

4.4. Confusion Matrix

The confusion matrix demonstrates strong agreement between actual and predicted crop classes, with most predictions concentrated along the diagonal. Minimal off-diagonal values indicate low misclassification and effective separation among multiple crop classes. The optimized Random Forest model, configured with 500 trees and a depth of 30, provides stable and efficient inference, supporting reliable real-time crop recommendation for tabular agricultural data.

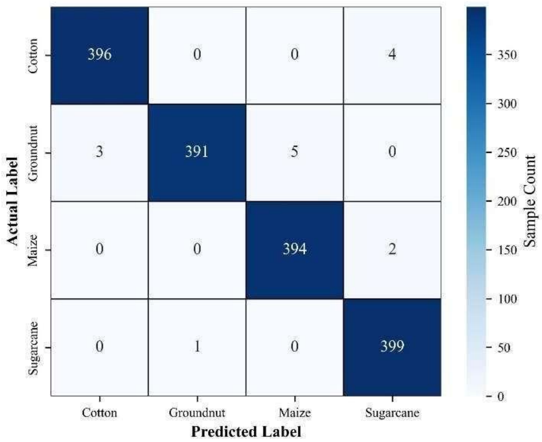

Fig. 4. Sample confusion matrix showing class-level prediction agreement for key crops

4.5. Computational Efficiency

The proposed AgroIntelligence framework demonstrates practical feasibility for deployment in low-bandwidth rural environments. The optimized Random Forest model enables efficient inference, while the geolocation-based Auto Mode retrieves contextual soil and weather information with minimal delay. Overall, the system supports responsive web-based crop advisory without imposing significant computational overhead, making it suitable for real- time agricultural decision support.

5 Conclusion

This paper presents AgroIntelligence, a real-time crop and fertilizer recommendation framework aimed at supporting rural precision agriculture through data-driven decision-making. The proposed system integrates KNN-based imputation and Min–Max scaling to effectively preprocess soil and climate data, followed by an optimized Random Forest classifier for accurate multi-class crop prediction. Experimental evaluation on a custom dataset of 18,240 agricultural records collected from Andhra Pradesh demonstrates a predictive accuracy of 98%, indicating reliable model performance under diverse agronomic conditions. Beyond crop prediction, the framework provides interpretable fertilizer recommendations, incorporates real-time weather awareness, and offers bilingual conversational support to address the practical needs of farmers operating in low-bandwidth rural environments. By combining automated recommendations with localized agricultural insights, the system bridges the gap between traditional farming practices and machine learning–assisted advisory services. Future work may focus on integrating crop disease detection and IoT-based soil health monitoring to further enhance field-level decision support and system adaptability. Overall, AgroIntelligence validates the applicability of ensemble-based machine learning for scalable and sustainable smart-agriculture decision support, while enhancing usability, trust, and adoption among rural farming communities.

Funding Information

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Ethics Statement

This study did not involve human or animal subjects and, therefore, did not require ethical approval.

Statement Of Conflict Of Interests

The authors declare that they have no conflicts of interest related to this study.

Licensing

This work is licensed under a Creative Commons Attribution 4.0 International License.

References

[1] Nurmalitasari, Nurchim, and R. D. Lestari, "Artificial intelligence-driven solar smart irrigation for sustainable agriculture: Trends, challenges, and SDG implications – A systematic review," *Smart Agricultural Technology*, vol. 12, p. 101665, Nov. 2025, doi: 10.1016/j.atech.2025.101665.

[2] A. Rustemi and F. Dalipi, "Synergizing IoT, AI, and blockchain for smart agriculture: Challenges, opportunities, and future directions," *Internet of Things*, vol. 34, p. 101778, Sep. 2025, doi: 10.1016/j.iot.2025.101778.

[3] O. V. Ocama *et al.*, "A review on Advancing Technologies in Precision Agriculture: Applications, challenges, and the Way forward," *Procedia Computer Science*, vol. 265, pp. 572–577, Jan. 2025, doi: 10.1016/j.procs.2025.07.221.

[4] S. Shukla, K. Chaudhary, Aahana, S. Phutela, R. Bhutani, and S. K. Shukla, "Smart crop varieties and Precision agriculture: a way ahead for climate-resilient sustainable agriculture," in *Elsevier eBooks*, 2025, pp. 435–466. doi: 10.1016/b978-0-443-26520-4.00027-5.

[5] R. Chowdhury, F. N. Nur, M. N. Islam, Md. N. Islam, P. Das, and A. S. Afridi, "SPAS-Dataset-BD: Dataset for smart precision agriculture system in Bangladesh," *Data in Brief*, vol. 61, p. 111727, May 2025, doi: 10.1016/j.dib.2025.111727.

[6] V. Sharma, G. Kaur, S. S, V. Chhabra, and R. Kashyap, "Smart irrigation systems in agriculture: An overview," *Computers and Electronics in Agriculture*, vol. 239, p. 111008, Sep. 2025, doi: 10.1016/j.compag.2025.111008.

[7] W. Wang and Q. Li, "Smart farming revolution: Leveraging machine learning for sustainable agriculture," *Journal of Cleaner Production*, vol. 527, p. 146434, Sep. 2025, doi: 10.1016/j.jclepro.2025.146434.

[8] F. R. De Avila and J. L. V. Barbosa, "Smart environments in digital agriculture: a systematic review and taxonomy," *Computers and Electronics in Agriculture*, vol. 236, p. 110393, Apr. 2025, doi: 10.1016/j.compag.2025.110393.

[9] N. N. Thilakarathne, M. S. A. Bakar, P. E. Abas, and H. Yassin, "Internet of things enabled smart agriculture: Current status, latest advancements, challenges and countermeasures," *Heliyon*, vol. 11, no. 3, p. e42136, Jan. 2025, doi: 10.1016/j.heliyon.2025.e42136.

[10] M. S. H. Eyasin, M. E. Sobhani, S. Nasrin, A. S. A. Rafi, and A. K. M. M. Islam, "CropSynergy: Harnessing IoT solutions for smart and efficient crop management," *Crop Design*, p. 100127, Dec. 2025, doi: 10.1016/j.cropd.2025.100127.

[11] M. Roy and A. Medhekar, "Transforming smart farming for sustainability through agri-tech Innovations: Insights from the Australian agricultural landscape," *Farming System*, vol. 3, no. 4, p. 100165, Jul. 2025, doi: 10.1016/j.farsys.2025.100165.

[12] V. Choudhary, P. Guha, G. Pau, and S. Mishra, “An overview of smart agriculture using internet of things (IoT) and web services,” *Environmental and Sustainability Indicators*, vol. 26, p. 100607, Jan. 2025, doi: 10.1016/j.indic.2025.100607.

[13] Y. Chandervanshi, P. Mandal, and S. Tewari, “Next generation nanobioformulation: A fascinating field for smart and sustainable agriculture,” *Plant Nano Biology*, vol. 13, p. 100191, Aug. 2025, doi: 10.1016/j.plana.2025.100191.

[14] H. Abdel-Gawad, I. A. Matter, M. I. Abo-Alkasem, N. G. A.- Balakocy, and O. M. Darwesh, “Phyco-biosynthesis of Chlorella-CuO-NPs and its Immobilization on Polyester/Cotton Blended Textile Waste Activated by Cellulase Enzymes for Application as Wastewater Disinfection Filter,” *Egyptian Journal of Chemistry*, vol. 67, no. 7, pp. 609–621, May 2024, doi: 10.21608/ejchem.2024.258806.9102.

[15] L. H. Abdel-Rahman, B. S. Al-Farhan, D. A. El-Ezz, M. a. A. Sayed, M. M. Zikry, and A. M. Abu-Dief, “Green biogenic synthesis of silver nanoparticles using aqueous extract of moringa oleifera: access to a powerful antimicrobial, anticancer, pesticidal and catalytic agents,” *Journal of Inorganic and Organometallic Polymers and Materials*, vol. 32, no. 4, pp. 1422–1435, Jan. 2022, doi: 10.1007/s10904-021-02186-9.