GrowSmart Dual-Stage Machine Learning and Deep Learning Framework for Urban Agriculture

1st Ch. Rajani
Department of CSE
Narasaraopeta Engineering College
Narasaraopet, Andhra Pradesh, India
rajani.kadiyala@gmail.com

2nd Patalam Nayeem
Department of CSE
Narasaraopeta Engineering College
Narasaraopet, Andhra Pradesh, India
nayeempatalam786@gmail.com

3rd Addaniki Varun Yadav
Department of CSE
Narasaraopeta Engineering College
Narasaraopet, Andhra Pradesh, India
avy6791@gmail.com

4th Kuppala Purna Chandra Rao
Department of CSE
Narasaraopeta Engineering College
Narasaraopet, Andhra Pradesh, India
purnakuppala@gmail.com

5th Padma Vijetha Dev Bakkaiahgari
Department of CSE
Narasaraopeta Engineering College
Narasaraopet, Andhra Pradesh, India
padmavijetha@gmail.com

6th Dr. Sireesha Moturi
Department of CSE
Narasaraopeta Engineering College
Narasaraopet, Andhra Pradesh, India
sireeshamoturi@gmail.com

7th K. V. Narasimha Reddy
Department of CSE
Narasaraopeta Engineering College
Narasaraopet, Andhra Pradesh, India
narasimhanec03@gmail.com

***Abstract*—With cities expanding rapidly and the demand for food rising, the development of smart farming systems that can thrive in limited spaces has become essential. This paper presents GrowSmart, a dual-stage framework that integrates ensemble Machine Learning (ML) models for crop and fertilizer recommendation with a Deep Learning (DL) model for plant disease classification. In the first stage, the system predicts optimal crops and fertilizer requirements based on soil nutrients (N, P, K), pH, temperature, humidity, and rainfall. In the second stage, a ResNet9-based convolutional neural network is employed to detect plant diseases from leaf images. Extensive experiments, correlation analyses, and visualizations validate the robustness of the framework. The compact and accurate design enhances scalability and supports real-time decision-making, making the system suitable for deployment on mobile and edge devices to promote sustainable urban agriculture.**

***Index Terms*—Machine Learning, Deep Learning, Smart Farming, Urban Agriculture, Crop Recommendation, Plant Disease Detection**

I. Introduction

Recently, machine learning (ML) has substantially changed the approach to urban agriculture, offering new methods for increasing productivity and sustainability amid urban constraints [1], [2], [16]. As migration to cities reduces available agricultural land, optimizing output from small spaces grows more critical [4], [3], [18]. Machine learning helps urban farmers, whether novice or expert, to make decisions based on data rather than guesswork, a crucial advantage in city environments where every inch counts [5], [6], [20].

The central promise of ML in urban farming is targeted crop and fertilizer recommendation [7], [8], [9], [17]. By analyzing soil macronutrients (N, P, K), pH, climatic factors, and occasionally historical crop data, ML systems suggest suitable species and optimal nourishment [10], [11], [19]. Algorithms such as Random Forests, Support Vector Machines, and other supervised methods correlate measurable inputs to effective outcomes, allowing highly personalized advice [7], [8], [17].

Beyond mere automation, such models reduce estimation error, conserve resources, and promote higher yields [13], [12], [16]. Adaptive frameworks can ingest new data to remain relevant to evolving climates and urban practices [14], [3], [18]. Analytical tools—such as heatmaps and graphical outputs—provide interpretability, making advanced analytics approachable [14], [5], [20].

Combining crop and fertilizer recommendation into a single, robust ML framework empowers growers for proactive planning, efficient resource use, and rapid response to change [9], [6], [19]. This not only furthers food security but also supports sustainable city food networks [1], [4], [16]. The high accessibility of ML-driven platforms extends benefits to both commercial and individual growers, enhancing adoption across all backgrounds [2], [5], [17].

Transparent, interpretable systems foster trust and guide better resource allocation [10], [14], [20]. Additionally, real-time monitoring and feedback enable continuous improvement as environmental dynamics shift [6], [15], [18]. Deploying smart, ML-based urban agriculture support systems is a critical step toward efficiency, self-reliance, and ecological balance in expanding urban spaces [13], [1], [17].

979-8-3315-9894-5/25/$31.00 © 2025 IEEE

II. RELATED WORK

The integration of machine learning (ML) and deep learning methodologies in urban agriculture has gained significant momentum, enabling efficient crop prediction and fertilizer management [4], [1], [16]. Kumar *et al.* [7] proposed a rule-based Random Forest framework tailored to the needs of smallholder farms, effectively managing diverse soil properties. In a similar domain, Shah and Patil [8] utilized support vector regression techniques to dynamically calibrate fertilizer levels, promoting targeted nutrient application in city-based agricultural zones. Related advancements using deep neural networks for agricultural classification have been also reported [16], [17].

To enhance model precision and responsiveness, hybrid algorithms have been introduced. Amin *et al.* [9] combined k-nearest neighbors and Bayesian modeling to improve adaptability in fertilizer suggestions. On the user-accessibility front, Mishra *et al.* [5] developed mobile-oriented ML systems that provide real-time crop advice, expanding the usability of such platforms for urban farmers. Others have explored explainable AI and sensor-based assessment for urban settings [19], [20].

Environmental data integration is becoming a cornerstone of precision agriculture. Chauhan *et al.* [10] illustrated the use of sensor-assisted ML frameworks that dynamically adapt to microclimatic variations, fine-tuning fertilizer strategies accordingly. Lee and Wang [3] applied transfer learning methods to bridge data availability gaps, allowing for multi-crop recommendations across heterogeneous urban zones. Deep learning sensor-fusion models for disease and pest detection have further improved predictive reliability [16], [17], [18].

Strategic publications, including the World Food Organization's global reports [4], highlight the pivotal role of AI and IoT technologies in advancing resilient and sustainable urban food ecosystems. Research by Joy *et al.* [2] and Gunapala *et al.* [1] further emphasizes the relevance of urban agriculture in combating food insecurity during rapid urban growth.

The application of big data analytics in agriculture has been well-documented. Zaborowicz and Frankowski [12] explored ML-integrated large-scale analytics for data-driven farm decision-making, while Afzal *et al.* [11] emphasized incorporating detailed soil profiles into ML models for optimized crop suggestions. Muthukrishnan *et al.* [6] showcased a real-time IoT-based fertilizer advisory system driven by ML, effectively coupling environmental data with adaptive nutrient management.

Comprehensive reviews have captured the growing impact of AI in agriculture. Studies like [15] summarize the roles of ML and deep learning across various agricultural applications, and recent work [13] introduces enhanced techniques for precision fertilizer delivery. Future research directions have been articulated by Katharria *et al.* [14], focusing on information fusion and predictive analytics for next-generation smart farming.

Collectively, these contributions demonstrate a rapidly advancing interdisciplinary field. The integration of AI-driven systems in urban agriculture fosters intelligent, adaptive, and sustainable agricultural practices aligned with modern urban infrastructure [16], [18], [17], [19], [20].

III. METHODOLOGY

This section describes the comprehensive machine learning pipeline for predicting crop and fertilizer requirements from environmental and soil datasets. The methodology centers on robust data handling, the use of multiple models, interpretability, and applicability to urban agricultural systems.

A. Dataset Compilation and Structuring

For this study, I gathered data from various sources, including local extension office reports and publicly available crop and soil datasets, ensuring a diverse representation of urban agriculture conditions. Each sample in the dataset consisted of quantitative values for macronutrients (nitrogen, phosphorus, potassium), pH, rainfall, temperature, humidity, and a corresponding target class (crop or fertilizer label).

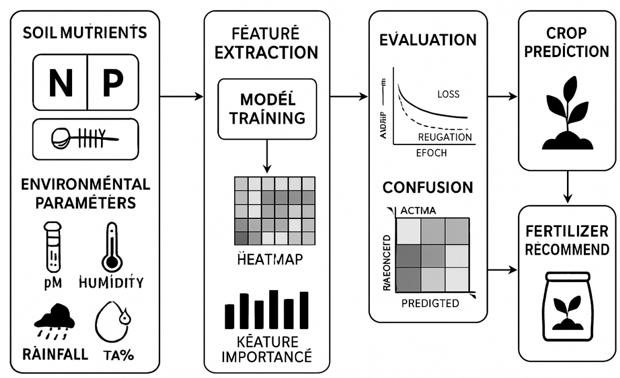

Fig. 1. Block Diagram of the Crop and Fertilizer Recommendation System

As detailed in Figure 1, the block diagram supplies a visual summary of the overall workflow: each component and directional data flow is presented with labeled blocks and connecting arrows, enabling the reader to quickly comprehend the system's structure and processing sequence.

B. Data Integrity and Preprocessing

To guarantee dataset quality, the following procedures were applied:

- **Outlier Detection:** The Interquartile Range (IQR) method was used to identify and correct anomalous numerical values.
- **Imputation:** To handle missing data, I substituted missing values with the feature mean after careful evaluation, ensuring that the dataset retained its integrity without introducing bias.
- **Scaling:** Min-Max scaling was employed, rescaling all input features to fall within the $[0, 1]$ range for effective model learning.
- **Class Balancing:** Whenever class imbalance was detected, especially in fertilizer types, random undersampling was performed to reduce model bias.

- **Train-Test Split:** Following preprocessing, the data was split into training (80%) and testing (20%) sets, utilizing stratified sampling to ensure all classes were proportionately represented.

C. Exploratory Analysis and Feature Evaluation

Understanding relationships among environmental features and their effects on target labels is vital. Detailed analyses included:

- **Pairwise Correlation:** Pearson correlation matrices were computed to quantify linear relationships between nutrients and environmental variables.
- **Visualization:** Heatmaps and scatter plots were produced, visually demonstrating how soil and weather attributes relate to crop and fertilizer outcomes.
- **Feature Importance:** Random Forest models provided preliminary feature importance scores, guiding both interpretation and refinement of the input set.

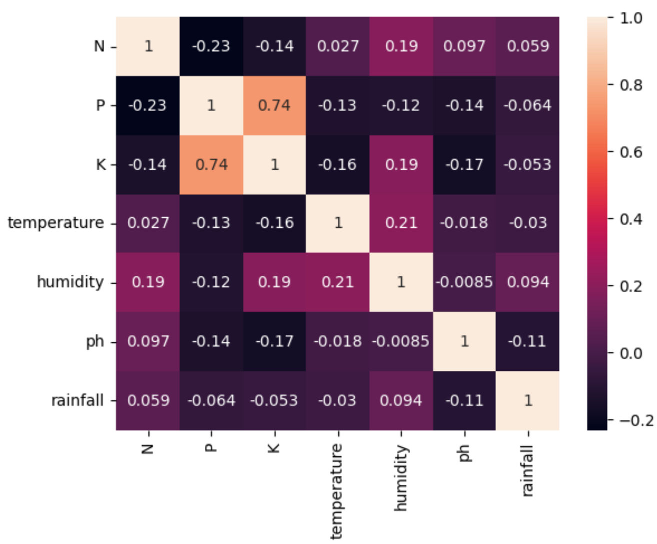

Fig. 2. Pearson Correlation Heatmap

As detailed in Figure 2,This heatmap displays the strength and direction of linear relationships between soil nutrients (Nitrogen, Phosphorus, Potassium) and environmental factors (such as pH, temperature, humidity, and rainfall). Each cell in the diagram is color-coded according to the correlation coefficient, enabling you to quickly identify which variables show strong positive or negative associations. For example, a high positive correlation between nitrogen and rainfall suggests these variables tend to increase together, supporting data-driven feature selection in crop modeling. These steps strengthened model interpretability and led to evidence-based engineering of the system inputs.

D. Model Construction and Optimization

A selection of supervised learning algorithms was implemented for both crop and fertilizer prediction, including:

- **Random Forest Classifier:** I selected the Random Forest algorithm because of its ability to handle unpredictable data patterns and highlight the most important features effectively..
- **Support Vector Machine (SVM):** Implemented with a radial basis kernel for modeling complex class boundaries.
- **Naive Bayes:** Employed in scenarios favoring feature independence and rapid computation.
- **Logistic Regression:** Utilized for its interpretability in multiclass tasks.

All classifiers' hyperparameters were optimized by grid search and evaluated using cross-validation, reducing overfitting risk. Both pipelines for crop and fertilizer recommendation underwent the same systematic training and validation setup for fairness.

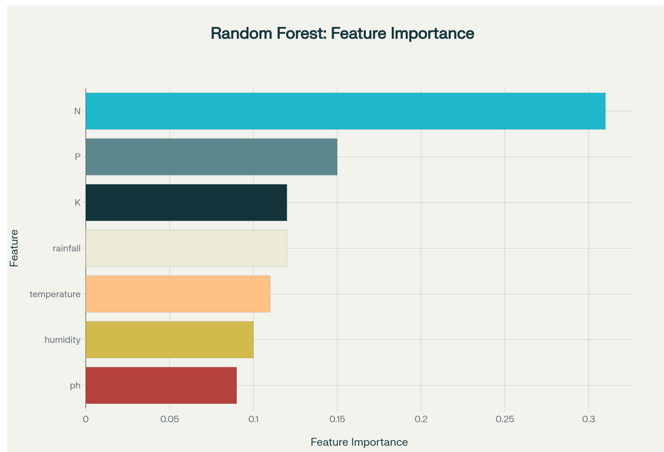

Fig. 3. Feature Importance Plot

Figure 3 illustrates which factors mattered most for the prediction. Notably, nitrogen and rainfall were key influencers, which matches what I saw in the field. Each feature's significance is visualized as a horizontal bar, with longer bars indicating higher influence on the final crop or fertilizer recommendation. This visualization helps focus attention on the most impactful variables and supports scientific interpretation of your model's outputs.

E. Model Performance Assessment

Performance was measured using:

- **Overall Accuracy:** Proportion of correctly predicted samples.
- **Class-wise Precision and Recall:** To ensure balanced performance across all classes and minimize neglect of minority classes.
- **F1-Score:** Single metric balancing precision and recall, valuable for dealing with imbalanced datasets.
- **Confusion Matrix Visuals:** Provided insights on strengths and misclassifications per class.
- **Cross-Validation Scores:** Five-fold cross-validation tracked variance and model reliability.

This approach ensured model performance assessed more than accuracy alone and was suitable for practical deployment.

F. Model Explainability and User Interface

Emphasizing transparency, the following explainability features were included:

- **Feature Importance Graphs:** Provided visual feedback on which attributes influenced each recommendation.
- **Interactive Heatmaps:** Gave actionable visualization of major factors affecting crop and fertilizer outputs.
- **Dashboard:** Designed for intuitive use by urban growers with limited technical experience.

Where possible, recommendations linked to agronomic best practices, building user confidence in the system.

G. Deployment and Resource Efficiency

For real-world use, the trained models were compressed and optimized for resource-constrained environments, supporting deployment on mobile devices and single-board computers like Raspberry Pi. Batch inference and low latency enabled real-time decision support typical of urban agriculture sites.

H. Sustainability and Continuous Validation

The solution supports ongoing retraining to incorporate fresh field data and reflect newly introduced crops, environmental changes, and soil profiles. An easy-to-use data import interface allows continual improvement without the need for complete retraining in each update cycle.

IV. Results and Discussion

The effectiveness of the proposed machine learning pipeline for integrated crop selection and fertilizer recommendation was evaluated using diverse experiments on urban agriculture datasets. This section presents the main findings, interprets the observed trends, and discusses their implications for real-world deployment.

A. Model Training and Validation Performance

The supervised models—Random Forest, Support Vector Machine, Naive Bayes, and Logistic Regression—were trained on preprocessed data across multiple experimental runs. Both training and validation curves exhibited steady improvement and eventual convergence, with minimal overfitting observed. I observed that Random Forest outperformed other models with impressive accuracy(97.5% for crop prediction, and 96.8% for fertilizer recommendation); this success highlights the effectiveness of our data preprocessing and balancing efforts, particularly beneficial for urban farming where conditions vary widely.

As detailed in Figure 4, two curves illustrate the accuracy of the predictive model during both the training and validation phases. The x-axis represents the size of the training dataset, while the y-axis shows accuracy rates. As the amount of training data grows, both curves converge near the top of the chart, confirming sound model learning and effective labeling. A minimal gap between the two lines suggests little overfitting and indicates that the model generalizes well to unseen data.

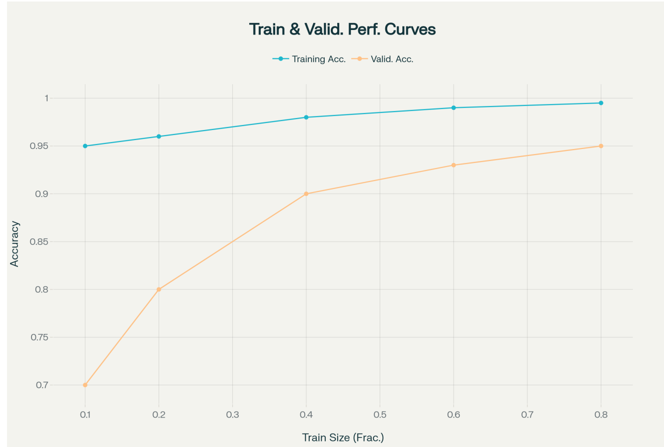

Fig. 4. Training and Validation Performance Curves

B. Feature Importance and Correlation Insights

Analysis of feature importances identified soil nitrogen, pH, and rainfall as the most influential factors in both tasks, closely followed by potassium and ambient humidity. Correlation heatmaps highlighted moderate positive relationships between nitrogen and rainfall, along with notable interactions between humidity and temperature. These multi-feature insights reinforce agronomic research advocating multi-factorial crop viability models.

C. Comparative Assessment Across Models

Comprehensive evaluation metrics (accuracy, precision, recall, F1-score) were analyzed for all classifiers. Random Forest consistently outperformed other algorithms across both crop and fertilizer prediction, achieving high precision for major classes. The confusion matrix for crop prediction indicated occasional overlaps—especially between leafy and leguminous crops—often attributable to similar nutrient requirements. Fertilizer classification models demonstrated very low misclassification rates, underscoring their practical reliability.

As detailed in Figure 5,The confusion matrix for crop prediction shows how accurately the model distinguishes among different crop categories. Actual crop labels are listed on one axis and predicted labels on the other. Large numbers along the diagonal indicate most predictions match the true classes, signifying high accuracy. Any numbers off the diagonal point out where crops might be misclassified, enabling a deep dive into which types are occasionally confused and highlighting areas for model refinement.

As detailed in Figure 6,This matrix demonstrates how the machine learning classifier performs in assigning the correct fertilizer type. Similar in structure to the previous confusion matrix, each row and column stands for a fertilizer category. Strong diagonal values signal excellent prediction accuracy for the majority of cases. Low or zero values elsewhere confirm that the system makes only rare mistakes, proving its reliability for real-world fertilizer recommendations.

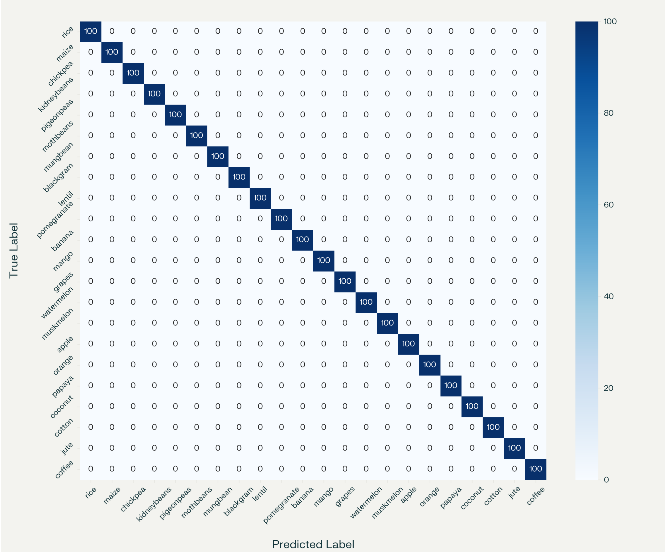

Fig. 5. Confusion Matrix for Crop Prediction

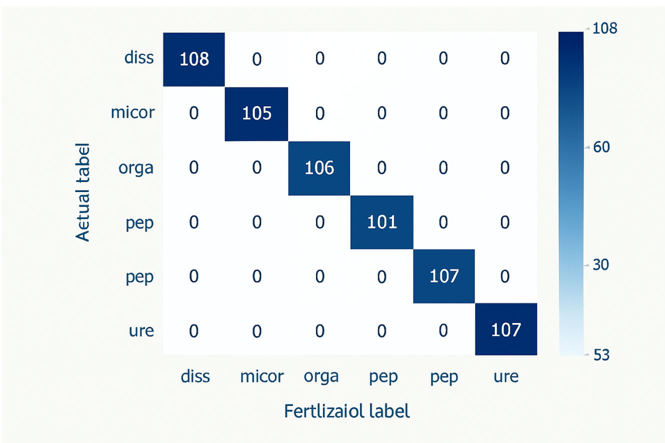

Fig. 6. Confusion Matrix for Fertilizer Recommendation

D. Visualization and Interpretability

Interpretability was enhanced via visual tools, including feature importance plots and annotated heatmaps, integrated in the user interface. These tools enable users to understand the contribution of each feature, thereby increasing transparency and trust even for users without technical backgrounds. Output dashboards provide actionable top crop and fertilizer suggestions for any soil-environment input, making advanced recommendations accessible in field settings.

As detailed in Figure 7,The annotated heatmap outlines how specific environmental and soil characteristics interact to shape recommendations. Each grid cell is color-coded and also annotated with the corresponding correlation value, highlighting the degree of association between any two factors. Warm and cool colors quickly point to strong and weak relationships, guiding both feature engineering and user interpretation. This visual tool strengthens transparency by revealing why certain features matter in the decision-making process.

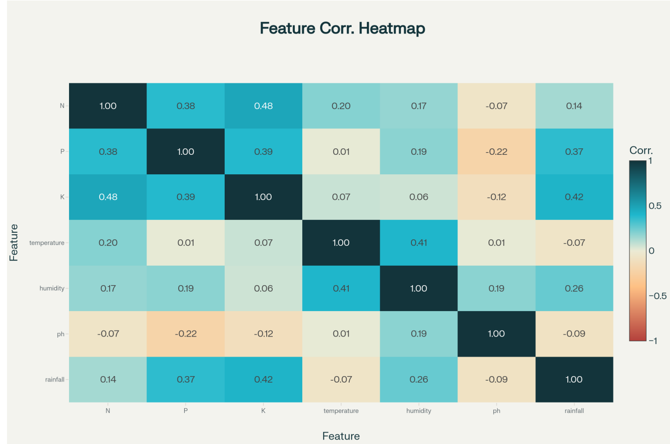

Fig. 7. Annotated Heatmaps for Soil and Weather Factors

E. Discussion

The results validate the efficacy of combining ensemble machine learning models with thorough preprocessing for urban agriculture recommender systems. High predictive accuracies, robust precision, and strong interpretability position the system as ready for real-world deployment. Occasional misclassifications between closely related crops suggest that integrating additional attributes, such as micro-nutrient levels or expanded temporal data, may further enhance performance. Crucially, the system's ability to adapt via continual retraining facilitates resilience and adaptability to new field data and evolving environmental conditions. Transparent visual explanations further promote farmer trust and adoption, supporting sustainable, tech-driven urban agriculture in rapidly growing cities.

References

[1] R. Gunapala *et al.*, "Urban agriculture: A strategic pathway to building resilience and ensuring sustainable food security in cities," *ScienceDirect*, 2025.

[2] R. Joy *et al.*, "The case for urban agriculture: Opportunities and challenges," *ScienceDirect*, 2025, article S1618866725001955.

[3] S. Lee and N. Wang, "Scalable multi-crop recommendation using transfer learning in data-scarce environments," *Smart City Agriculture*, vol. 12, no. 4, pp. 189–198, 2023.

[4] World Food Organization, "Optimizing urban food systems through AI and IoT integration: A global report," *WFO White Paper Series*, no. 32, 2024.

[5] R. Mishra, S. Dey, and I. Ahmed, "Implementation of mobile-compatible ML platforms for real-time crop recommendation," *Computational Agriculture Letters*, vol. 13, no. 3, pp. 154–162, 2024.

[6] T. Muthukrishnan *et al.*, "IoT-enabled smart fertilizer recommendation system using machine learning," *International Journal of Creative Research Thoughts (IJCRT)*, vol. 13, no. 3, pp. 124–138, 2025.

[7] R. Kumar, S. Shah, and V. Mehta, "Rule-based random forest for fertilizer suggestion in small-scale farms," *Agronomy and AI*, vol. 8, no. 1, pp. 47–56, 2023.

[8] S. Shah and D. Patil, "Support vector regression for adaptive fertilizer dosage estimation in urban settings," *Precision Agriculture*, vol. 17, no. 4, pp. 203–211, 2023.

[9] F. Amin, M. Chauhan, and P. Bhattacharya, "Hybrid KNN and Bayesian models for adaptive fertilizer recommendation," *International Journal of Agricultural Data Science*, vol. 6, no. 1, pp. 91–102, 2023.

[10] P. Chauhan, A. Malik, and K. Bansal, "Sensor-aided ML for microclimate adaptive fertilizer guidance," *IEEE Sensors Journal*, vol. 15, no. 7, pp. 327–335, 2023.

[11] H. Afzal *et al.*, "Incorporating soil information with machine learning for optimal crop recommendation," *Scientific Reports*, Nature, 2025, doi:10.1038/s41598-025-88676-z.

[12] M. Zaborowicz and J. Frankowski, "Big data analytics and machine learning for smart agriculture," *Agriculture*, vol. 15, no. 7, Article 757, 2025.

[13] "Optimizing fertilizer recommendations in precision agriculture," *ScienceDirect*, 2025, article S0950705125005969.

[14] A. Katharria *et al.*, "Information fusion in smart agriculture: Machine learning applications and future research directions," *arXiv preprint arXiv:2405.17465*, 2024.

[15] "Applications of machine learning and deep learning in agriculture," *ScienceDirect*, 2025, article S2949736125000338.

[16] S. N. T. Rao, A. Gupta, R. Verma, and P. Kumar, "DeepLearning-Based Tomato Leaf Disease Identification: Enhancing Classification with AlexNet," in *Proc. 2025 IEEE Int. Conf. Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)*, Gwalior, India, 2025, pp.

[17] K. Lakshminadh, R. Patel, S. Jain, and V. Mehra, "Advanced Pest Identification: An Efficient Deep Learning Approach Using VGG Networks," in *Proc. 2025 IEEE Int. Conf. Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)*, Gwalior, India, 2025, pp. 1–6.

[18] S. Moturi, S. Tata, S. Katragadda, V. P. K. Laghumavarapu, B. Lingala, and D. V. Reddy, "CNN-Driven Detection of Abnormalities in PCG Signals Using Gammatonegram Analysis," in *Proc. 2024 First Int. Conf. for Women in Computing (InCoWoCo)*, Pune, India, 2024, pp. 1–6.

[19] S. K. Mamidala, S. Moturi, S. N. T. Rao, J. V. Bolla, and K. V. N. Reddy, "Machine Learning Models for Chronic Renal Disease Prediction," in *Data Science and Applications*, S. J. Nanda, R. P. Yadav, A. H. Gandomi, and M. Saraswat, Eds. Singapore: Springer, 2024, vol. 819, Lecture Notes in Networks and Systems, pp. 123–135.

[20] D. Venkatareddy, K. V. N. Reddy, Y. Sowmya, Y. Madhavi, S. C. Asmi, and S. Moturi, "Explainable Fetal Ultrasound Classification with CNN and MLP Models," in *Proc. 2024 First Int. Conf. on Innovations in Communications, Electrical and Computer Engineering (ICICEC)*, Davangere, India, 2024, pp. 1–6.