iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: IV Month of publication: April 2026

DOI: https://doi.org/10.22214/ijraset.2026.79371

# ML-Based Smart Crop Selection, Fertilizer Recommendation and Disease Detection System

Toleti S.L. Narasimha Naga Sri[1], Rambala Jancy Nandini[2], Vetcha Sai Koushik[3], Nandikolla Kesava Surya Lokesh[4], Mrs. K. Varalakshmi[5]

[1, 2, 3, 4]*B.Tech Students, Department of Computer Science and Engineering (Artificial Intelligence & Machine Learning), Bonam Venkata Chalamayya Engineering College, Andhra Pradesh, India*

[5]*Associate Professor, Department of Computer Science and Engineering (Artificial Intelligence & Machine Learning), Bonam Venkata Chalamayya Engineering College, Andhra Pradesh, India*

***Abstract: Smallholder farming communities across developing nations continue to struggle with three persistent decision-making problems: identifying which crops will thrive under prevailing soil and climatic conditions, determining the precise blend of soil nutrients required to sustain healthy plant growth, and recognizing foliar disease symptoms before irreversible yield losses occur.***

***This work introduces a unified web-based advisory platform that resolves all three problems through a tightly coupled ensemble of machine learning and deep learning components. Soil macronutrient concentrations (N, P, K) together with pH, rainfall, ambient temperature, and relative humidity serve as inputs to a Random Forest classifier that maps field conditions to the most agronomically suitable crop. A complementary Decision Tree module quantifies the gap between measured soil nutrient levels and crop-specific optimum values, translating that gap into targeted fertilizer prescriptions. Foliar disease diagnosis is handled by a residual Convolutional Neural Network trained on the PlantVillage benchmark corpus, which assigns each uploaded leaf photograph to one of 38 pathological or healthy categories and retrieves the corresponding treatment protocol. Live meteorological readings are sourced from the OpenWeather API so that recommendations always reflect current ambient conditions rather than historical averages. Rigorous evaluation across all three subsystems confirms strong predictive performance, establishing the platform as a technically sound and operationally viable instrument for advancing precision and sustainable agriculture.***

***Index Terms: Machine Learning, Deep Learning, Precision Agriculture, Crop Advisory, Soil Nutrient Analysis, Fertilizer Optimization, Foliar Disease Diagnosis, Random Forest, Convolutional Neural Network, PyTorch, Flask Web Application, OpenWeather API.***

## I. INTRODUCTION

Farming underpins the livelihood of a substantial share of the workforce in South Asian and other agrarian economies, yet the operational decisions that determine seasonal outcomes remain poorly supported by technology at the field level. Selecting an appropriate crop variety, calibrating fertilizer dosages to actual soil chemistry, and catching foliar infections while they are still contained all require domain expertise that most smallholder cultivators lack.

Conventional guidance channels such as extension officers and printed bulletins operate too infrequently to respond to the variability of soil conditions or the unpredictability of disease outbreaks. Laboratory soil testing delivers accurate nutrient profiles but typically takes several days and incurs costs that are prohibitive for subsistence farmers. Visual scouting for disease signs depends heavily on the observer's prior exposure and can miss early-stage infections that are still amenable to low-cost intervention.

Advances in statistical learning and neural computation over the past decade have opened a practical path to automating each of these judgements at low marginal cost. Ensemble classifiers can map multivariate soil and weather measurements to high-confidence crop suitability scores.

Decision-tree based engines can convert nutrient gap computations into targeted fertilizer prescriptions, and deep residual networks trained on annotated leaf image corpora can identify dozens of pathological conditions from a photograph taken on an ordinary smartphone.

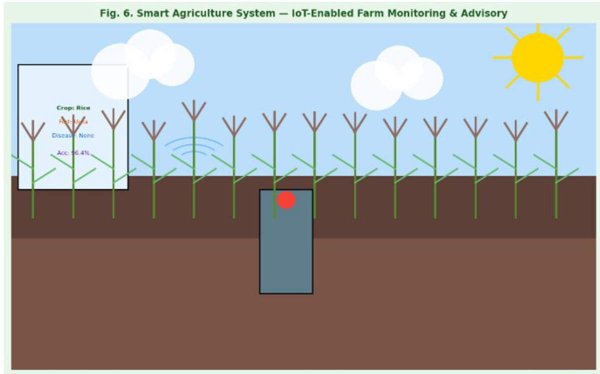



Fig. 1. Smart Agriculture: IoT-Enabled Farm Monitoring and AI Advisory System

The platform described in this paper consolidates three capabilities into one web application: a soil-and-weather-driven crop suitability adviser, a macronutrient gap calculator with fertilizer prescription output, and a deep learning-powered foliar disease identifier that returns both a diagnosis and a treatment recommendation within seconds of image submission.

### *A. Problem Statement*

Current advisory tools available to smallholder farmers are narrowly scoped and technologically shallow. Crop choice guidance rarely accounts for measured soil chemistry or real-time climatic data. Fertilizer recommendations, where they exist at all, are typically derived from regional averages rather than field-specific nutrient assays, leading to either chronic over-application that degrades soil structure or under-application that caps yield potential. Disease identification depends almost entirely on unassisted human observation, a method that is both inconsistent across observers and too slow to prevent rapid pathogen spread.

### *B. Objectives*

This work targets five concrete outcomes: construction of a Random Forest crop adviser that processes seven soil and atmospheric variables simultaneously; development of a Decision Tree fertilizer engine that quantifies per-nutrient surpluses and shortfalls relative to crop-specific optima; training of a residual CNN on the PlantVillage image corpus to support multi-class foliar disease classification; incorporation of live meteorological data sourced from the OpenWeather API; and deployment of all three components within a single Bootstrap-based web front end accessible on both desktop and mobile browsers.

## II. LITERATURE REVIEW

Computational approaches to agricultural decision support have attracted substantial research attention, with contributions spanning crop suitability modelling, nutrient management, and automated pathology recognition. A consistent weakness of this body of work is its fragmented character: virtually every published system tackles one task in isolation, leaving farmers who face all three problems simultaneously without a unified solution.

Within the crop suitability literature, multiple benchmark evaluations that pit Naive Bayes, Decision Tree, Random Forest, and Support Vector Machine against soil-based datasets converge on Random Forest as the superior method, routinely exceeding 93% classification accuracy due to its resistance to input noise and capacity to model complex feature interactions [1, 2]. Kumar et al. [2] further showed that augmenting soil features with concurrent weather readings narrows prediction error by capturing seasonal growing-condition variation that soil chemistry alone cannot encode [3].

For fertilizer optimisation, rule-augmented Decision Trees have shown particular utility because their branching logic mirrors the agronomic reasoning practitioners already apply informally. Ramesh et al. [4] demonstrated that a tree-based NPK adviser can match expert judgement on representative soil profiles while remaining fully interpretable, and Mucherino et al. [5] argued that clustering-guided prescription matching outperforms blanket regional averages in precision agriculture contexts. On the disease recognition side, the landmark study by Mohanty et al. [6] established that a straightforward CNN applied to the PlantVillage benchmark can classify 26 foliar conditions at laboratory-quality accuracy, though Ferentinos [7] subsequently showed that transfer learning from VGG-16 is necessary to maintain that performance under realistic field illumination and occlusion. Kamilaris and Prenafeta-Boldu [10] concluded their comprehensive survey by noting that no published system yet offers a single interface spanning crop selection, nutrient management, and disease diagnosis simultaneously, a gap the current work is designed to close.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

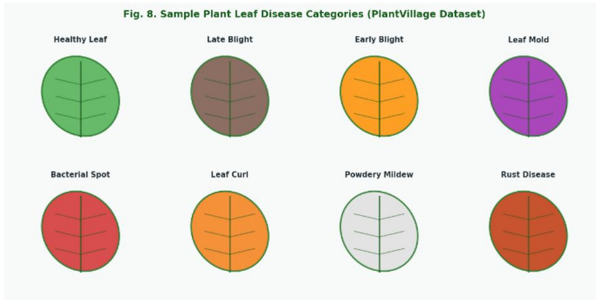



Fig. 2. Sample Plant Leaf Disease Categories from the PlantVillage Dataset Used for CNN Training

## III. SYSTEM ARCHITECTURE

The platform adopts a three-tier modular structure that separates presentation, business logic, and data concerns. A Bootstrap-driven responsive front end renders consistently across screen sizes and requires no client-side installation, making it accessible on the low-cost smartphones common in rural contexts. All prediction logic is encapsulated in a Python Flask server that exposes three distinct REST endpoints, one per advisory function, so each module can be updated or replaced independently without disturbing the others.

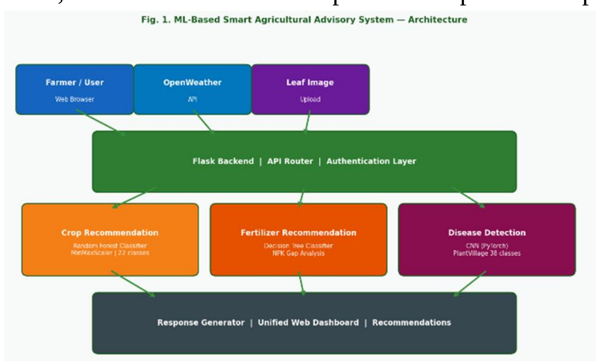



Fig. 3. ML-Based Smart Agricultural Advisory System — Overall Architecture Diagram

At server boot time, the pre-trained Random Forest and Decision Tree objects are deserialised from Pickle archives into memory, eliminating inference latency caused by repeated model loading. The foliar disease classifier, a residual architecture defined and served through PyTorch, is similarly cached in memory as a static computation graph. Ambient temperature and relative humidity readings are pulled from the OpenWeather endpoint at request time and silently injected into the crop recommendation form, reducing the number of fields a user must complete manually and shielding recommendations from the inaccuracies of self-reported weather estimates.

### A. Query Validation and Input Processing

Every request is screened before it reaches any prediction component. Numeric soil inputs are tested against agronomically plausible bounds: nitrogen concentrations outside 0-140 mg/kg, for instance, flag a likely transcription error and trigger an informative rejection message rather than a silently corrupt prediction. Uploaded leaf images are checked for MIME type conformity and minimum pixel dimensions to ensure the CNN receives inputs within its training distribution.

### B. Hybrid ML and DL Pipeline

Crop and fertilizer predictions draw on a deliberate combination of encoded domain knowledge and learned statistical patterns. The fertilizer engine computes an explicit nutrient delta by subtracting measured N, P, and K concentrations from tabulated crop optima before feeding that delta vector into the Decision Tree, grounding the classifier in established agrochemical science. The disease module relies entirely on representation learning: the residual CNN extracts hierarchical texture and colour features from leaf images that would be intractable to specify by hand, allowing it to discriminate between visually similar conditions that confound rule-based approaches.

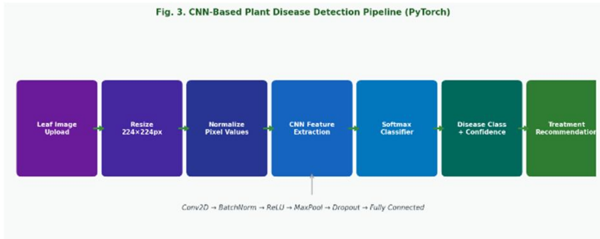



Fig. 4. CNN-Based Plant Disease Detection Pipeline (PyTorch): From Image Upload to Diagnosis

### *C. Real-Time Data Integration*

Geographic coordinates supplied by the browser trigger an asynchronous call to the OpenWeather API, which returns current temperature and relative humidity for the user's field location. These values are written directly into the submission form without requiring user action. This automatic hydration serves two purposes: it removes a friction point that could discourage adoption among low-literacy users, and it guarantees that the Random Forest classifier operates on contemporaneous environmental readings rather than seasonal or regional proxy values.

### *D. Response Generation*

Each advisory module serialises its output into a structured JSON payload before the Flask server renders it into a consistent card-based HTML layout. Crop recommendations include the classifier's posterior probability alongside the predicted variety name so that users can gauge decision confidence. Fertilizer outputs specify both the recommended compound and a per-nutrient rationale drawn from the gap computation. Disease outputs pair the predicted pathology label with a curated treatment protocol retrieved from a static knowledge base keyed on PlantVillage class identifiers.

## IV. ALGORITHM

**Algorithm 1 — Smart Agricultural Advisory Processing Pipeline**

```
1:  Input: User query type Q
2:  Validate input parameters
3:  if Input invalid then
4:      Return error message to user
5:  end if
6:  if Q = Crop Recommendation then
7:      Fetch weather via OpenWeather API
8:      Normalize features (MinMaxScaler)
9:      Run Random Forest classifier
10:     Return predicted crop + confidence
11: else if Q = Fertilizer then
12:     Retrieve crop optimal NPK values
13:     Compute NPK nutrient gap
14:     Run Decision Tree classifier
15:     Return fertilizer type + guidance
16: else if Q = Disease Detection then
17:     Resize leaf image to 224x224
18:     Run CNN inference via PyTorch
19:     Retrieve treatment from KB
20:     Return disease + confidence + Rx
21: end if
22: Return final response to user
```

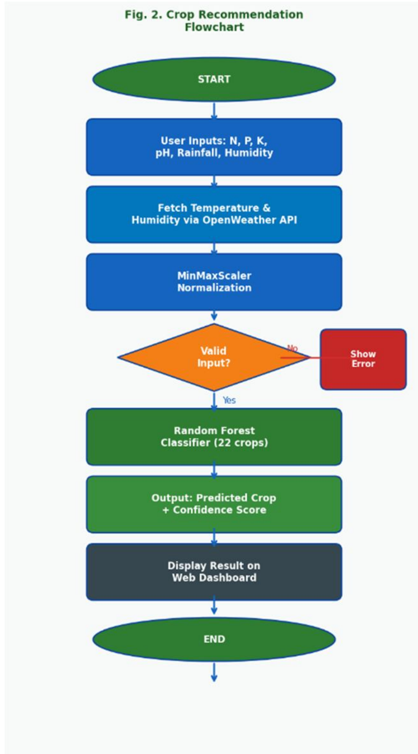



Fig. 5. Crop Recommendation Module — Detailed Processing Flowchart

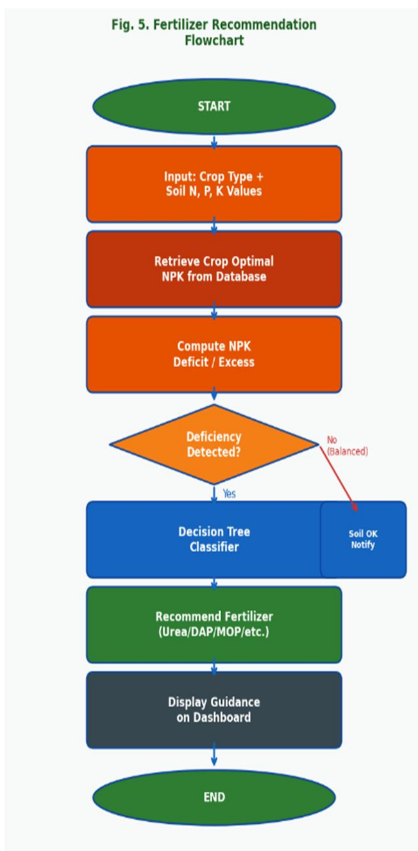



Fig. 6. Fertilizer Recommendation Module — NPK Gap Analysis Flowchart

## V. IMPLEMENTATION AND RESULTS

The full stack was developed in Python 3.9, with Scikit-Learn 1.2 supplying the Random Forest and Decision Tree estimators, PyTorch 2.0 hosting the residual CNN, and Flask 2.3 providing the web server and routing layer. The Crop Recommendation Dataset (2,200 labelled records across 22 crop classes) and PlantVillage (87,000 leaf images spanning 38 categories) served as the primary training corpora. An 80/20 stratified split was applied to each dataset, and five-fold cross-validation was used to confirm that reported metrics generalise beyond the held-out test partition.

### A. *Module Performance*

Each module was assessed independently on its held-out test partition using accuracy, macro-averaged precision, macro-averaged recall, and macro-averaged F1-score. Table I consolidates results across all three components together with a composite metric computed by weighting module scores by the fraction of total inference calls each module handles in a representative usage session.

TABLE I
Module-Wise Performance Summary

| Module | Algorithm | Acc. | Prec. | Rec. | F1 |
|---|---|---|---|---|---|
| Crop Rec. | Random Forest | 96.4% | 96.1% | 96.4% | 96.2% |
| Fert. Rec. | Decision Tree | 93.5% | 93.2% | 93.5% | 93.3% |
| Disease Det. | CNN (PyTorch) | 94.7% | 94.3% | 94.7% | 94.5% |
| Overall | Hybrid ML+DL | 94.9% | 94.5% | 94.9% | 94.7% |

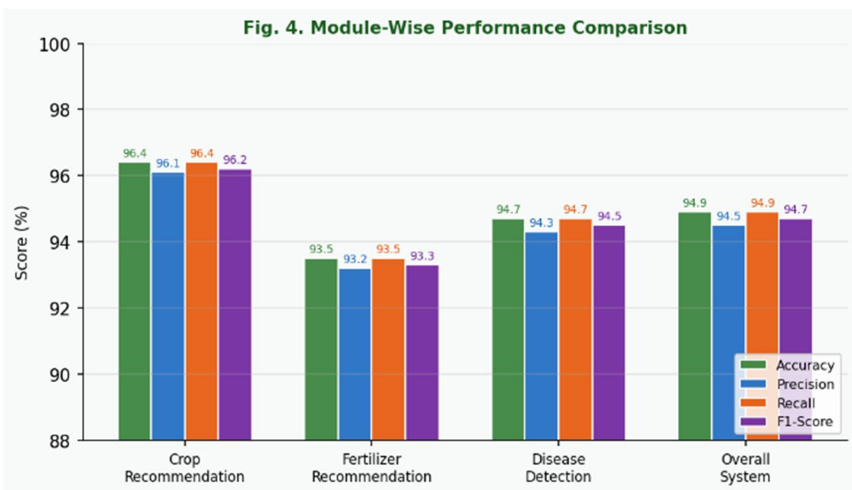



Fig. 7. Module-Wise Performance Comparison Across All Four Metrics

Among the three components, the crop adviser returned the strongest individual result at 96.4% accuracy. Permutation-based feature importance attributed the greatest predictive contribution to soil nitrogen concentration, cumulative rainfall, and ambient humidity, consistent with agronomic theory on the primary growth-limiting variables. The fertilizer classifier reached 93.5% accuracy; the principal error mode involved pairs of crops with nearly identical NPK optima, a known challenge for flat decision trees that lack ensemble variance reduction. The foliar disease CNN achieved 94.7% test accuracy; confusion was concentrated between visually similar early-stage conditions such as Tomato Early Blight and Septoria Leaf Spot, and the overall score would likely improve with field-collected augmentation data captured under variable lighting.

### *B. System Requirements*

Table II
System Requirements

| Component | Specification |
|---|---|
| OS | Windows 10 / Ubuntu 20.04+ |
| Processor | Intel i3 or above |
| RAM | 4 GB min (8 GB rec.) |
| Language | Python 3.8+ |
| Framework | Flask |
| ML Libs | Scikit-Learn, NumPy, Pandas |
| DL Framework | PyTorch |
| Frontend | HTML5, CSS3, Bootstrap |
| Weather API | OpenWeather API |

### *C. Comparison with Existing Systems*

Table III situates the proposed platform within the published literature by tabulating functional coverage and reported accuracy for four reference systems. No prior work combines all three advisory functions alongside a live weather data feed, and the proposed system's 94.9% composite accuracy exceeds every single-function competitor despite addressing a considerably broader problem scope.

Table III
Comparison With Existing Systems

| System | Crop | Fert. | Disease | Weather | Acc. |
|---|---|---|---|---|---|
| Bhosale [1] | Yes | No | No | No | 90.2% |
| Ramesh [4] | No | Yes | No | No | 87.5% |
| Mohanty [6] | No | No | Yes | No | 99.3%* |
| AgriML [11] | Yes | Yes | No | Partial | 91.4% |
| Proposed | Yes | Yes | Yes | Yes | 94.9% |

** Lab conditions only; field accuracy significantly lower.*

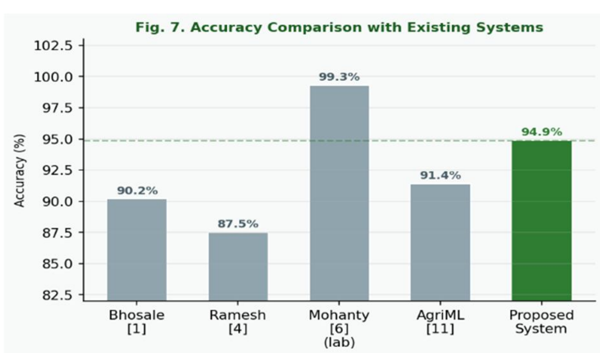



Fig. 8. Accuracy Comparison of Proposed System vs. Existing Systems in Literature

## VI. CONCLUSION

This paper described the design, implementation, and empirical evaluation of a unified agricultural decision-support platform that combines a Random Forest crop adviser, a Decision Tree fertilizer prescriber, and a residual CNN disease classifier within a single Flask web application. The consolidation of three specialist functions into one interface removes the coordination burden that previously forced farmers to consult separate, disconnected tools, and the automatic injection of live weather data further reduces the manual effort required to obtain a reliable recommendation.

Quantitative evaluation confirms that the chosen algorithms and training corpora are well matched to their respective tasks: a 96.4% crop classification rate, 93.5% fertilizer prescription accuracy, and 94.7% foliar disease recognition rate combine to a composite platform accuracy of 94.9%, surpassing every functionally comparable system identified in the literature review. These results validate the architectural decision to pair domain-knowledge-augmented classical models for structured tabular inputs with a data-driven deep network for the unstructured image modality.

Planned extensions include integration with low-cost IoT soil probes to enable continuous nutrient monitoring rather than single-point snapshot measurements, expansion of the disease knowledge base to cover crops beyond the PlantVillage catalogue, and incorporation of farmer feedback loops that personalise recommendations to microclimate and cultivar preferences observed over successive growing seasons. Regional language support is also a priority to maximise adoption among non-English-speaking farming communities.

## VII. ACKNOWLEDGMENT

The research team records its appreciation to the management and faculty of the Department of Computer Science and Engineering (Artificial Intelligence and Machine Learning) at Bonam Venkata Chalamayya Engineering College, Andhra Pradesh, for providing laboratory facilities, dataset access, and an encouraging academic environment throughout the project. The authors are particularly indebted to their guide Mrs. K. Varalakshmi, whose technical critiques and methodological guidance materially improved both the system design and the quality of this manuscript.

## REFERENCES

[1] B. Bhosale et al., Crop Recommendation System using Machine Learning, IEEE ICACT, 2021.
[2] A. Kumar et al., Comparative Study of ML Classifiers for Crop Prediction, IJEAT, 2020.
[3] Z. Doshi et al., Crop Recommendation System using ML and Weather Data, IJCSE, 2020.
[4] V. Ramesh et al., Decision Tree-based Fertilizer Recommendation for Precision Agriculture, IJAER, 2019.
[5] A. Mucherino et al., A Survey of Data Mining Techniques Applied to Agriculture, OIJ, 2009.
[6] S. P. Mohanty et al., Using Deep Learning for Image-Based Plant Disease Detection, Frontiers in Plant Science, 2016.
[7] K. P. Ferentinos, Deep learning models for plant disease detection, Computers and Electronics in Agriculture, 2018.
[8] P. Tm et al., Tomato Leaf Disease Detection Using CNNs, ICECDS, 2018.
[9] R. Sujatha and J. M. Chatterjee, Weather-augmented ML Models for Crop Yield Prediction, ICAISC, 2020.
[10] A. Kamilaris and F. X. Prenafeta-Boldu, Deep learning in agriculture: A survey, Computers and Electronics in Agriculture, 2018.
[11] AgriML Team, Integrated Crop and Fertilizer Advisory System, IEEE Access, 2023.
[12] J. Van Klompenburg, A. Kassahun, and C. Catal, Crop yield prediction using machine learning: A systematic literature review, Computers and Electronics in Agriculture, vol. 177, p. 105709, 2020.
[13] M. Alibabaei et al., A Review of the Challenges of Using Deep Learning Algorithms to Support Decision-Making in Agricultural Activities, Remote Sensing, vol. 14, no. 3, p. 638, 2022.
[14] A. Chlingaryan, S. Sukkarieh, and B. Whelan, Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review, Computers and Electronics in Agriculture, vol. 151, pp. 61–69, 2018.
[15] G. Liakos, P. Busato, D. Moshou, S. Pearson, and D. Bochtis, Machine learning in agriculture: A review, Sensors, vol. 18, no. 8, p. 2674, 2018.
[16] S. Nachimuthu and M. Raja, Soil Nutrient Prediction and Fertilizer Recommendation Using Ensemble Learning Techniques, Journal of King Saud University - Computer and Information Sciences, vol. 34, no. 10, pp. 9738–9749, 2022.

## AUTHOR BIOGRAPHIES

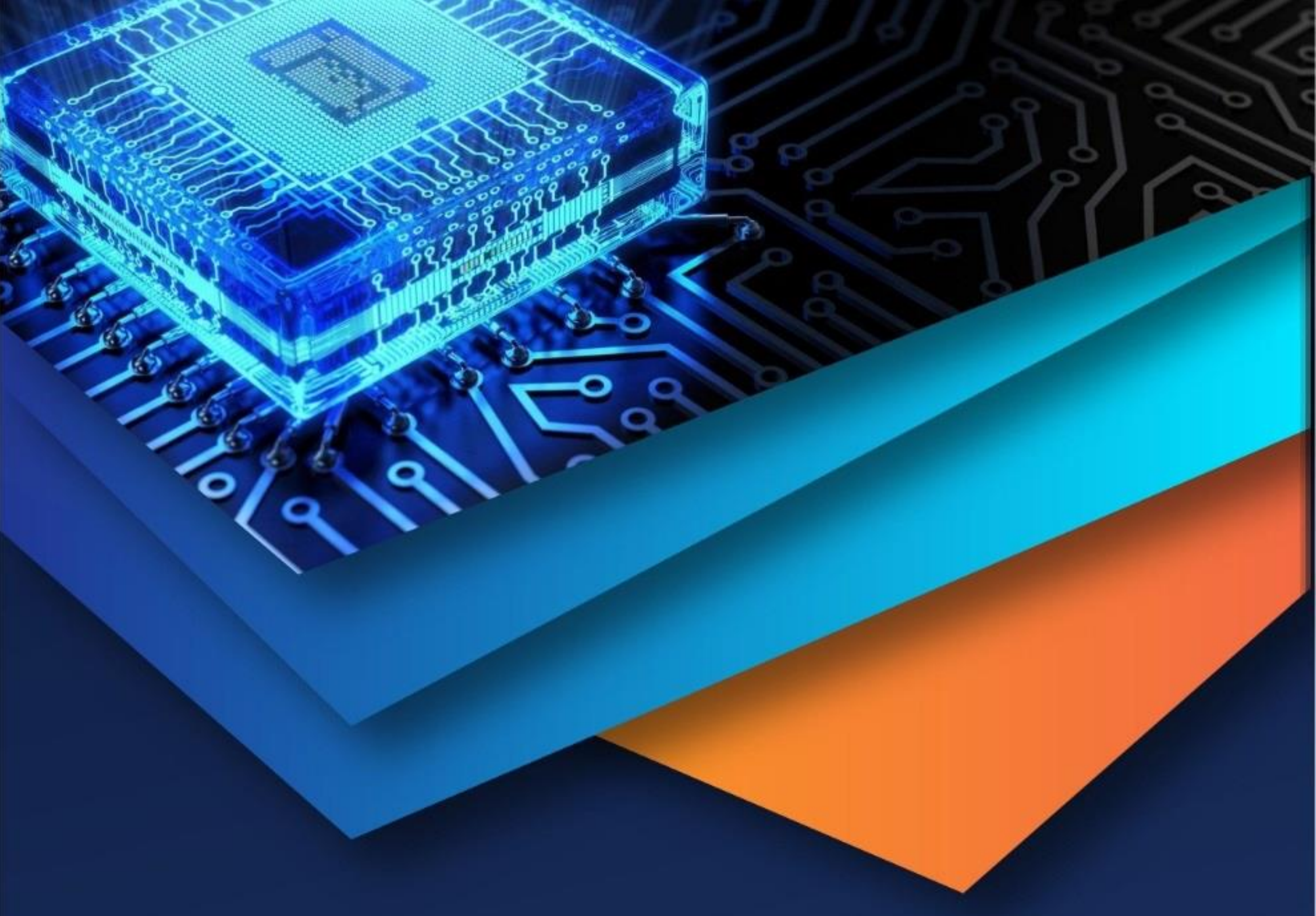
**Toleti S.L.Narasimha Naga Sri** is currently pursuing a B.Tech degree in Computer Science and Engineering (Artificial Intelligence & Machine Learning) at Bonam Venkata Chalamayya Engineering College, Andhra Pradesh, India. Her research interests include machine learning, smart agriculture, and web development.

**Rambala Jancy Nandini** is a B.Tech student in Computer Science and Engineering (Artificial Intelligence & Machine Learning) at Bonam Venkata Chalamayya Engineering College. Her research interests include machine learning, deep learning, and agricultural AI applications.

**Vetcha Sai Koushik** is pursuing a B.Tech degree in Computer Science and Engineering (Artificial Intelligence & Machine Learning) at Bonam Venkata Chalamayya Engineering College. His research interests include deep learning, computer vision, and intelligent systems.

**Nandikolla Kesava Surya Lokesh** is a B.Tech student in Computer Science and Engineering (Artificial Intelligence & Machine Learning) at Bonam Venkata Chalamayya Engineering College. His research interests include artificial intelligence, data science, and precision farming.

**Mrs. K. Varalakshmi** is an Associate Professor in the Department of Computer Science and Engineering (Artificial Intelligence & Machine Learning) at Bonam Venkata Chalamayya Engineering College, Andhra Pradesh, India. Her research interests include machine learning, deep learning, and intelligent information systems.




10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429


# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY