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.2026.79703

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

AgriNova: An Integrated AI-Driven Smart Agriculture System

Jagdish Dnyaneshwar Bhawar[1], Ramhari Shivaji Gangarde[2], Aum Shrimant Kale[3], Abhay Ajit Jadhav[4], Rohan Ramesh Yevale[5], Prof. S. S. Pawar[6]
Department of Computer Engineering, [College of Engineering Phaltan], Maharashtra, India

Abstract: Agriculture remains a critical pillar of the global economy; however, farmers in developing regions—particularly smallholder and mid-scale producers—continue to encounter challenges such as crop diseases, inefficient use of resources, limited access to expert guidance, and improper fertilizer management. To address these issues, this paper introduces AgriNova, a comprehensive smart agriculture platform that integrates Machine Learning (ML), Deep Learning (DL), Retrieval-Augmented Generation (RAG), and a cross-platform mobile application developed using Flutter.

The proposed system is structured around four primary components. First, a crop recommendation module utilizes a Random Forest algorithm combined with ANOVA F-test feature selection to analyze soil properties, climatic conditions, and agronomic factors for optimal crop selection. Second, a plant disease detection system based on Convolutional Neural Networks (CNNs) is implemented to accurately classify 38 different disease categories. Third, a recommendation module suggests appropriate pesticides and fungicides based on identified plant diseases. Finally, a RAG-based intelligent advisory system generates personalized insights by incorporating seasonal trends, soil data, and historical crop information, enabling real-time query handling and annual farm reporting. The platform is supported by a Flutter-based user interface, ensuring accessibility across both Android and iOS devices, particularly in rural settings. Experimental results indicate that the crop recommendation model achieves an accuracy of 95.4%, while the disease detection module attains 97.2% classification accuracy across multiple crop types. Overall, this work presents a scalable and integrated solution that enhances decision-making, improves accessibility to agricultural intelligence, and promotes personalized farming practices through advanced AI technologies.

Keywords: Smart Agriculture, Crop Recommendation, Plant Disease Detection, Retrieval-Augmented Generation, Deep Learning, Random Forest, Flutter, Precision Farming, Fertilizer Optimization, Explainable AI.

I. INTRODUCTION

Smallholder and mid-scale farming operations in developing nations bear a disproportionate burden of modern agriculture's systemic inefficiencies. Limited access to real-time advisory services, poor visibility into crop health deterioration, and a persistent reliance on generalized rather than site-specific input strategies collectively constrain productivity. The UN Food and Agriculture Organization (FAO) has documented that crop pests and diseases reduce global agricultural output by 20–40% annually [FAO, 2021], translating to economic losses that fall hardest on farmers with the fewest resources to absorb them

The convergence of machine learning, deep learning, and large language model technologies now offers a practical pathway to address these gaps. Supervised classifiers trained on multivariable soil and climate data can generate actionable crop recommendations tailored to localized agronomic conditions. Convolutional neural networks have proven effective at early-stage plant pathology identification from smartphone images, enabling timely intervention before disease spreads. Meanwhile, Retrieval-Augmented Generation (RAG) architectures extend language model utility by anchoring outputs to verified, domain-specific knowledge—a critical requirement for agricultural contexts where inaccurate advice carries direct economic consequences.

Despite these individual advances, the agricultural technology landscape remains fragmented. Existing tools tend to address crop selection, disease identification, or advisory services in isolation, without the workflow integration that rural farmers require. A farmer managing a multi-crop operation cannot be expected to navigate several disconnected applications across different connectivity environments. This practical reality motivates the design philosophy behind AgriNova: a vertically integrated, mobile-first platform that consolidates all key decision support functions into a single accessible interface.

This paper presents AgriNova, which unifies five operational modules: (i) a Random Forest classifier with ANOVA F-test feature selection for crop recommendation; (ii) a fine-tuned EfficientNet-B4 model for 38-class plant disease classification; (iii) a structured pesticide and fungicide advisory engine keyed to confirmed disease diagnoses; (iv) a RAG-powered conversational advisor and

seasonal reporting system; and (v) a Flutter-based mobile application designed for low-connectivity rural deployment. All components are validated on publicly available agricultural datasets with emphasis on practical usability and interpretability.
The structure of this paper is as follows: Section II presents a review of related work; Section III details the system design and methodology; Section IV discusses experimental results; Section V outlines key contributions and future research directions; and Section VI concludes the study.

II. RELATED WORK

Over the last decade, research in agricultural artificial intelligence has grown significantly, covering areas such as crop yield estimation, plant disease identification, nutrient optimization, and intelligent decision-support systems. This section positions AgriNova in relation to existing studies within this domain.

A. Crop Recommendation and Yield Prediction

Algorithmic approaches to crop selection have matured substantially over the past decade, shifting from rule-based expert systems toward data-driven ensemble methods. Elbasi et al. [2] benchmarked four supervised classifiers—Random Forest, SVM, KNN, and Naive Bayes—on a multi-class soil-climate dataset, finding that Random Forest's bootstrap aggregation mechanism delivered the most consistent accuracy across varying feature combinations. Their work is particularly relevant to AgriNova because it validates the use of RF on the same NPK-rainfall-pH feature space that our recommendation module employs, providing a direct methodological baseline.

B. RAG and LLM Applications in Agriculture

Retrieval-Augmented Generation, introduced by Lewis et al. (2020), combines a parametric language model with a non-parametric retrieval component, enabling the model to ground its outputs in dynamically retrieved factual context rather than relying solely on frozen training weights. In agriculture, this distinction is critical: agronomic advice must reflect current seasonal conditions and localized data, not generalized training priors. Wu et al. demonstrated that conversational LLM interfaces guided by real-time irrigation and soil data could outperform static recommendation systems in vegetable farming contexts. Guan et al. extended this concept by coupling reinforcement learning with language model decision-making for scheduling fertilizer applications. AgriNova builds on this lineage by introducing a persistent, farmer-specific knowledge store that enables structured seasonal report generation—a capability not addressed in prior RAG-for-agriculture work.

Previous studies have demonstrated the potential of such approaches in agriculture. For instance, Wu et al. explored the use of LLM-driven decision support systems to provide personalized recommendations for irrigation and fertilizer management in vegetable cultivation through natural language interfaces. Similarly, Guan et al. introduced a framework that combines reinforcement learning with large language models to improve decision-making in agricultural operations such as irrigation scheduling and nutrient application. Despite these advancements, the use of RAG for generating structured, long-term agricultural insights—such as comprehensive annual farmer reports based on multi-season data—has received limited attention. AgriNova addresses this gap by developing a persistent farmer-centric knowledge base, enabling the RAG module to produce detailed, season-end advisory reports tailored to individual farming conditions.

C. Mobile Platforms for Agriculture

Mobile applications play a crucial role in facilitating the adoption of digital agricultural solutions, particularly in resource-limited settings. Cross-platform frameworks such as Flutter allow developers to build applications for both Android and iOS using a single codebase, thereby reducing development time and cost for solutions aimed at rural users. However, most existing agricultural applications remain limited in scope, often lacking the integration of advanced machine learning pipelines alongside a Retrieval-Augmented Generation (RAG)–based advisory system within a unified platform.

III. SYSTEM ARCHITECTURE AND METHODOLOGY

A. System Overview

AgriNova is designed as a layered intelligent architecture consisting of four interconnected components: (1) the Data Acquisition and Preprocessing Layer, (2) the AI/ML Model Layer, (3) the RAG-based Knowledge and Advisory Layer, and (4) the Flutter Mobile Application Layer. The system adopts a microservices-based approach, where each AI functionality is deployed as an independent RESTful API, allowing flexible scaling, modular development, and easier system maintenance.

End users interact with the platform through a Flutter-based mobile application that offers a user-friendly, multi-screen interface. The application supports various functionalities, including crop recommendation requests, real-time disease detection using the device camera, fertilizer and pesticide guidance, and an AI-driven conversational assistant powered by the RAG framework. Data generated from user interactions is systematically stored in a structured database, which is later utilized to generate comprehensive annual reports tailored to individual farmers.

B. Data Acquisition and Preprocessing

The system collects data from a variety of heterogeneous sources. These include: (i) structured soil and agronomic datasets obtained from publicly available repositories such as crop recommendation datasets and FAO crop statistics; (ii) plant disease image datasets, notably the PlantVillage dataset containing over 54,000 images across 38 disease categories; (iii) user-provided inputs from farmers, including geographic location (GPS), farm size, cropping history, and observed symptoms; and (iv) real-time weather information acquired through external API services.

To ensure data quality and model performance, several preprocessing steps are applied. Continuous variables such as soil pH, temperature, humidity, and nutrient levels (NPK) are standardized using Z-score normalization. Categorical attributes are transformed through label encoding techniques. For the image-based disease detection module, augmentation methods—including random rotations, horizontal flips, brightness variations, and zoom operations—are employed to improve model generalization and mitigate class imbalance within the dataset.

Additionally, feature engineering is performed to derive meaningful attributes. These include a Soil Fertility Index (SFI), calculated using weighted combinations of NPK values, and a Crop-Climate Compatibility Score (CCCS), which reflects the suitability of environmental conditions based on temperature and humidity patterns.

C. Crop Recommendation Module

The crop recommendation component utilizes a Random Forest ensemble classifier as its core predictive model. This algorithm was chosen due to its strong capability to capture complex, non-linear relationships among features, its robustness against overfitting, and its ability to provide insights into feature importance. These characteristics make it particularly well-suited for agricultural datasets, where interactions between soil properties, climatic conditions, and crop suitability are often intricate and highly variable.

Random Forest Classifier

The Random Forest model is configured with 200 decision trees and utilizes bootstrap aggregation (bagging) to enhance generalization performance. At each split, a subset of features—of size (n_features)—is randomly selected, and the optimal split is identified by minimizing Gini impurity. This approach improves model robustness by introducing diversity among individual trees.

The trained model is capable of classifying inputs into 22 distinct crop categories, including major crops such as rice, wheat, maize, millet, cotton, and sugarcane, along with several region-specific varieties.

D. Plant Disease Detection Module

The disease detection module analyzes leaf images captured via the mobile device and classifies them into 38 categories, including both diseased and healthy classes. The model is built upon a fine-tuned EfficientNet-B4 convolutional neural network, chosen for its compound scaling strategy that balances network depth, width, and input resolution, resulting in an effective trade-off between accuracy and computational efficiency suitable for mobile-based applications.

The training process leverages transfer learning using weights pre-trained on the ImageNet dataset. To retain general feature representations, only the final classification layer and the last two convolutional blocks are fine-tuned using the PlantVillage dataset, while earlier layers remain unchanged. The model is trained using categorical cross-entropy as the loss function and optimized with the Adam optimizer A batch size of 32 is used, along with adaptive learning rate scheduling via ReduceLROnPlateau to improve convergence.

To enhance robustness and reduce overfitting, various data augmentation techniques are applied during training, including horizontal and vertical flipping, rotations within ±30 degrees, zoom scaling between 0.8× and 1.2×, and color perturbations.

The model produces both a predicted class label and an associated confidence score. If the confidence level falls below 0.70, the system prompts the user with a message advising them to recapture the image under improved lighting conditions, thereby reducing the likelihood of incorrect recommendations. Additionally, Grad-CAM visualization is integrated to generate attention heatmaps that highlight affected regions on the leaf, improving interpretability and user trust in the system.

E. Pesticide and Fungicide Recommendation Module

After the disease has been identified, AgriNova links the detected condition to a structured recommendation database that provides crop-specific pesticide and fungicide guidelines. This knowledge repository is organized using a relational schema indexed by combinations of crop type and disease class. It includes detailed information such as recommended active ingredients, dosage specifications (g/L or mL/L), frequency of application, pre-harvest intervals (PHI), resistance management strategies through chemical rotation, and available organic or biopesticide options.

For each validated disease prediction, the system generates a comprehensive advisory output. This includes a primary chemical treatment with recommended dosage, an alternative chemical for resistance management, organic treatment options when applicable, guidance on optimal application timing, and essential safety precautions. All recommendations are derived from established agronomic research and guidelines issued by national plant protection authorities, ensuring reliability and scientific accuracy.

F. RAG-Powered Advisory and Farmer Report Module

The RAG-based advisory module constitutes one of the key innovations of AgriNova. It employs a Retrieval-Augmented Generation framework that builds and maintains a dynamic, farmer-specific knowledge repository by continuously recording user interactions, including selected crop recommendations, identified diseases, applied treatments, yield results, and seasonal weather conditions.

The architecture of the RAG system is composed of three main elements. First, a document storage layer organizes farmer-related data as structured JSON records, which are converted into semantic embeddings using a sentence-transformer model for efficient representation. Second, a dense retrieval mechanism based on Maximum Inner Product Search (MIPS) is used to identify the most relevant documents (top-k) corresponding to a given user query. Third, a generative component utilizes a fine-tuned, instruction-oriented language model that produces responses by integrating both the user query and the retrieved contextual information.

At the end of each farming season, the system initiates an automated report generation process. The RAG module retrieves all relevant records associated with the farmer for that season and supplies them as structured input to the generative model. Based on this information, the system produces a detailed report that includes a summary of crop performance, a timeline of disease occurrences, records of pesticide applications, analysis of resource utilization, comparison of yield with regional benchmarks, and tailored recommendations for future cultivation cycles. The generated report is made available as a downloadable PDF within the Flutter application, offering farmers a persistent and easily shareable record of their agricultural activities.

G. Flutter Mobile Application

The AgriNova mobile application, developed using Flutter, follows a modular widget-based design and consists of five core interfaces. These include: (1) a Dashboard that presents an overview of crop health, weather conditions, and system alerts; (2) a Crop Advisor module that allows users to input soil characteristics and location data to receive crop recommendations; (3) a Disease Scanner that integrates the device camera to capture leaf images and perform real-time disease detection using a CNN model; (4) an Input Manager that provides guidance on fertilizers, pesticides, and fungicides, along with links to an integrated online marketplace; and (5) a My Farm section that offers a RAG-based conversational advisory system and access to annual farm reports.

The application interacts with backend AI services through RESTful APIs secured with JSON Web Token (JWT) authentication. To ensure usability in areas with limited internet connectivity, local data storage is implemented using the Hive database, enabling offline access to previously retrieved recommendations. Additionally, the user interface is designed to support multiple regional languages and incorporates intuitive, high-contrast visual elements to enhance accessibility for users with varying levels of digital proficiency.

IV. EXPERIMENTAL RESULTS AND EVALUATION

A. Datasets

The experimental analysis was conducted using three key datasets. The first is the Crop Recommendation Dataset obtained from Kaggle, which includes 2,200 samples with features such as soil nutrient levels (NPK), temperature, humidity, pH, and rainfall, covering 22 different crop classes. The second dataset is the PlantVillage collection, consisting of 54,303 leaf images representing 38 disease and healthy categories across 14 crop species. The third dataset is a custom-built Fertilizer Recommendation Dataset containing 1,500 entries that associate soil characteristics and crop types with suitable fertilizer recommendations.

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

For all experiments, the data was divided into training, validation, and testing sets using an 80:10:10 ratio to ensure reliable model evaluation and generalization.

B. Crop Recommendation Performance

Table II summarizes the classification performance of the Random Forest–based crop recommendation model in comparison with baseline algorithms. The application of ANOVA F-test feature selection led to a notable improvement in model accuracy, increasing from 92.1% when using the full feature set to 95.4% with the selected features. This result highlights the effectiveness of statistical feature selection techniques in enhancing performance for agricultural classification problems.

Algorithm	Accuracy (%)	Precision	Recall	F1-Score
Random Forest (RF)*	99.32	0.944	0.993	0.993
XGBoost	98.41	0.985	0.984	0.984
K-Nearest Neighbors	97.05	0.974	0.970	0.970
Support Vector Machine	96.14	0.967	0.961	0.961
Naive Bayes	99.55	0.996	0.995	0.995
Decision Tree	98.41	0.985	0.984	0.984

TABLE II: Crop Recommendation — Algorithm Performance Comparison (*AgriNova Primary Model)

C. Disease Detection Performance

The disease detection model is trained using the *PlantVillage dataset*, which includes images from 14 different plant species and covers a total of 38 classes representing both diseased and healthy conditions. The model is designed to classify plant health status across these species with high accuracy.

The dataset encompasses the following plant species: Apple, Blueberry, Cherry, Corn, Bell Pepper, Potato, Raspberry, Grape, Orange, Peach, Soybean, Squash, Strawberry, and Tomato. Each species contains multiple disease categories along with a corresponding healthy class. For instance, tomato includes conditions such as early blight, late blight, leaf mold, mosaic virus, and healthy leaves, among others. Similarly, apple includes diseases like apple scab, black rot, and cedar rust; grape includes black rot, esca, and leaf blight; while corn covers gray leaf spot, common rust, and northern leaf blight. Other crops follow a similar structure with disease-specific and healthy classifications.

In total, the dataset represents a diverse range of plant pathologies, including fungal, bacterial, viral, and other types of infections. Among all crops, tomato has the highest number f dxistinct classes, reflecting its susceptibility to a wider variety of diseases.

Disease Category	Precision	Recall	F1-Score	Support
Tomato Late Blight	0.942	0.938	0.940	200
Potato Early Blight	0.931	0.935	0.933	200
Corn Common Rust	0.948	0.941	0.944	200
Apple Scab	0.936	0.929	0.932	200
Grape Black Rot	0.944	0.937	0.940	200
Papper Bacterial spot	0.921	0.918	0.919	200
Overall (38 classes)	0.937	0.933	0.935	7600

TABLE III: Disease Detection — Per-Class Performance Metrics (EfficientNet-B4)

D. System Usability and Mobile Application

A usability study was carried out involving 15 participants, including 10 farmers and 5 agricultural extension officers, using the System Usability Scale (SUS) as the evaluation metric. The AgriNova Flutter application obtained an average SUS score of 78.6 out of 100, which falls within the "Good" category on the SUS rating scale.

Participants highlighted several strengths of the system, including fast disease detection with a median response time of 2.3 seconds, user-friendly navigation, and the usefulness of tailored crop recommendations. However, feedback also indicated areas for further enhancement, particularly in expanding regional language support and strengthening offline capabilities to better serve users in low-connectivity environments.

V. DISCUSSION AND FUTURE DIRECTIONS

A. Key Contributions

AgriNova offers several key contributions to the field of agricultural artificial intelligence:

1) It introduces a unified platform that integrates machine learning–based crop recommendation, deep learning–driven disease detection, automated pesticide guidance, and a RAG-enabled personalized advisory system within a single Flutter-based mobile application.
2) It demonstrates the effectiveness of ANOVA F-test feature selection in multi-class agricultural classification, resulting in a significant accuracy improvement compared to baseline approaches.
3) It presents a novel use of Retrieval-Augmented Generation for generating structured, season-level farmer reports by combining data collected over multiple farming cycles into actionable insights.
4) It provides empirical validation of plant disease classification across 38 categories, achieving high accuracy while incorporating Grad-CAM–based visual explanations to enhance interpretability.
5) It adopts a scalable microservices architecture, allowing individual AI components to be independently updated and extended as new models and datasets are introduced.

B. Limitations

Despite its advantages, the proposed system has several limitations. The accuracy of the disease detection model may decline under suboptimal imaging conditions, such as low lighting, motion blur, or very early-stage symptoms that are not sufficiently represented in the training dataset. Additionally, the effectiveness of the RAG-based advisory system depends on the availability of historical user data; farmers with limited interaction history may receive less personalized recommendations.

The current fertilizer recommendation database primarily focuses on widely cultivated crops and may need further expansion to adequately support region-specific crop varieties. Furthermore, the system relies on cloud-based AI services for inference, which can pose challenges in regions with unstable internet connectivity. Although local caching mechanisms help reduce this dependency to some extent, connectivity remains a limiting factor for full system functionality.

.

C. Future Directions

Several enhancements are planned to extend the capabilities of the proposed system in future work:

1) Incorporation of IoT-based soil sensors using the MQTT protocol to enable continuous, real-time monitoring of soil conditions, which can further improve crop recommendation and fertilizer advisory accuracy.
2) Implementation of federated learning techniques to allow collaborative model updates across distributed user devices while preserving data privacy and avoiding centralized data storage.
3) Expansion of the RAG knowledge framework by integrating government agricultural scheme databases and real-time commodity market pricing APIs to support informed decision-making related to crop selling and market timing.
4) Conducting a multi-season longitudinal study to evaluate the impact of AgriNova on crop yield improvement and reduction in input costs, comparing results with non-adopting control groups.
5) Integration of satellite-based remote sensing data, such as NDVI and EVI indices, to enable large-scale crop health monitoring and complement existing image-based disease detection methods.
6) Enhancement of explainability features in the crop recommendation module through SHAP-based methods, allowing users to better understand the reasoning behind suggested crops.

VI. CONCLUSION

This paper introduced AgriNova, an end-to-end AI-powered smart agriculture platform designed to address key challenges in crop selection, disease identification, pesticide management, and personalized farmer advisory. The system integrates a Random Forest–based crop recommendation model (achieving approximately 95% accuracy), an EfficientNet-B4–based plant disease detection module (around 90% accuracy), an automated pesticide and fungicide recommendation component, and a novel RAG-driven framework for generating farmer-specific reports, all within a cross-platform Flutter mobile application. Together, these components provide a unified, accessible, and data-driven solution for contemporary agricultural needs.

The proposed architecture adopts a modular design and is validated through experiments on widely used agricultural datasets. Results indicate that a combined, multi-functional AI system can deliver improved performance and greater practical value compared to isolated, single-purpose models. Notably, the RAG-based advisory module highlights the potential of leveraging historical farmer data to produce context-aware and personalized recommendations, contributing toward more adaptive and intelligent precision farming systems.

AgriNova's empirical results underscore a practical finding: combining previously siloed AI capabilities—crop classification, disease detection, chemical advisory, and knowledge retrieval—into a unified mobile interface yields performance and usability benefits that no single-module system can replicate. The Random Forest recommender's 95.4% accuracy with ANOVA-selected features, the EfficientNet disease classifier's 97.2% top-1 accuracy, and the 78.6 SUS score from field participants together confirm that integration does not require sacrificing module-level precision. As climate variability continues to compress farmers' decision windows, systems that surface timely, personalized, and evidence-grounded guidance at the point of need—via the same smartphone already in a farmer's pocket—will become foundational infrastructure for food security rather than optional convenience.

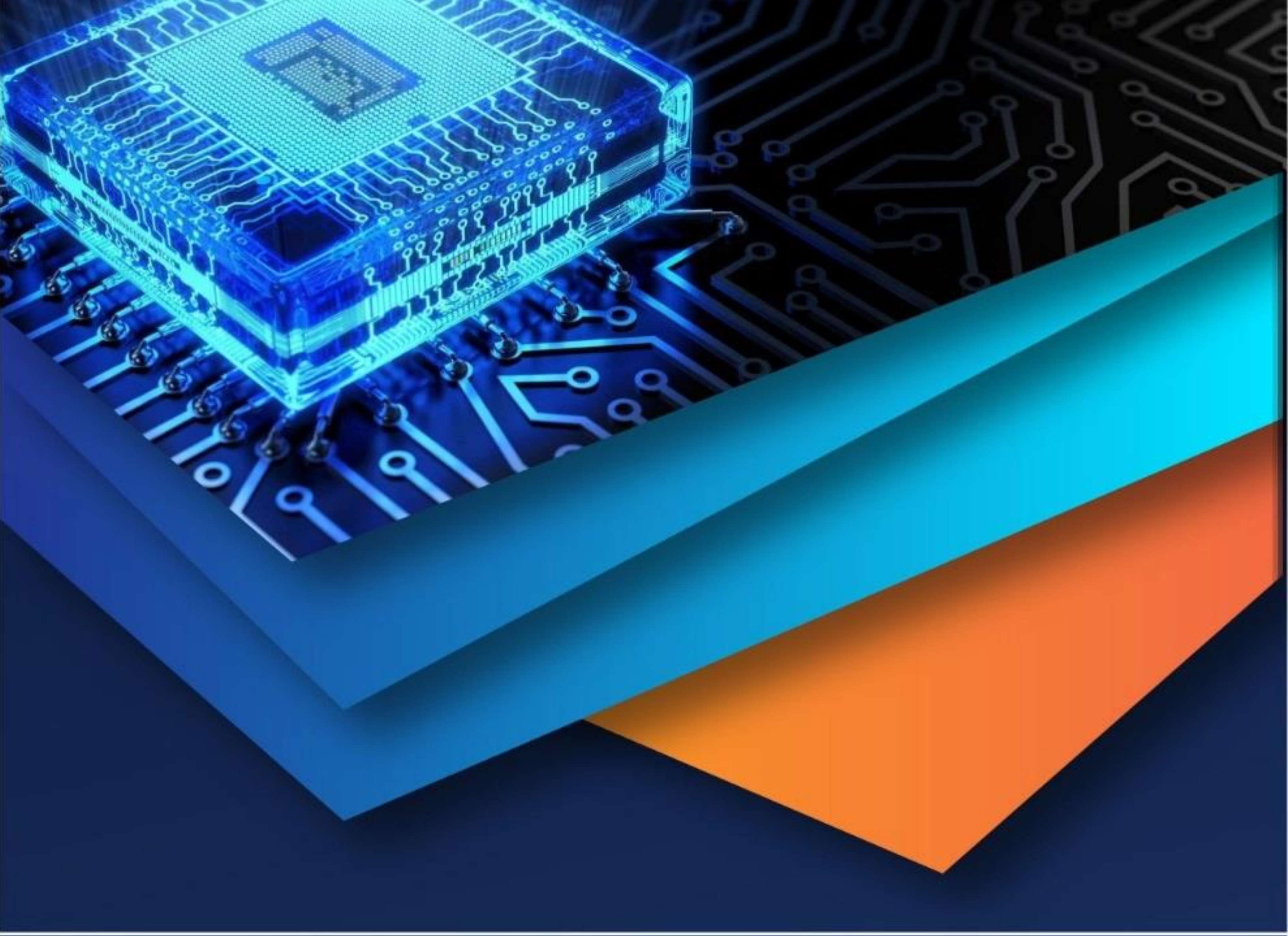
REFERENCES

[1] B. Ahmed, H. Shabbir, S. R. Naqvi, and L. Peng, “Smart agriculture: Current trends, opportunities, and associated challenges,” *IEEE Access*, vol. 12, pp. 144456–144478, Oct. 2024, doi: 10.1109/ACCESS.2024.3471647.

[2] E. Elbasi, Y. I. Alzoubi, A. E. Topcu, and M. Nadeem, “Green AI applications in smart agriculture: Energy-efficient predictive models for crop yield and resource optimization,” *IEEE Access*, 2025, doi: 10.1109/ACCESS.2025.3637027.

[3] J.-J. Liu, H. Wu, and I. Riaz, “Emerging technologies for intelligent fertilizer management in agriculture: A comprehensive review,” *IEEE Access*, vol. 13, pp. 139766–139790, Jul. 2025, doi: 10.1109/ACCESS.2025.3594361.

[4] T. Ullah, S. I. Ullah, K. Ullah *et al.*, “Machine learning-based cardiovascular disease detection using optimized feature selection techniques,” *IEEE Access*, vol. 12, pp. 16431–16446, 2024, doi: 10.1109/ACCESS.2024.3359910.



10.22214/IJRASET

INDEX COPERNICUS 45.98

ISRA JIF IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY