iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** VI **Month of publication:** June 2026

DOI: https://doi.org/10.22214/ijraset.2026.83480

www.ijraset.com

Call: ✆08813907089 | E-mail ID: ijraset@gmail.com

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 V May 2026- Available at www.ijraset.com

IoT-Based Soil Nutrient Analysis and Crop Recommendation System for Precision Agriculture

Kiran B K[1], J A Prajwal[2], Dr. Mohana Lakshmi J[3]
Department of Computer Science and Engineering, Malnad College of Engineering (An Autonomous Institution affiliated to VTU, Belagavi), Hassan, Karnataka 573202, India

Abstract: Agriculture plays a vital role in feeding the growing global population, yet optimizing crop production and resource management remains a significant challenge for farmers. This paper presents a Machine Learning (ML)-enabled Internet of Things (IoT) prototype designed to monitor soil parameters in real time and provide customized crop recommendations. The proposed system integrates four sensors — a JXBS-3001 NPK sensor, an FC-28 soil moisture sensor, a DHT11 temperature-humidity sensor, and an analog soil pH probe — deployed in the crop field and interfaced with a NodeMCU (ESP8266) microcontroller. Collected data is transmitted to the Ubidots cloud platform via the MQTT protocol. The Random Forest classifier is employed as the primary algorithm, achieving 99.09% test accuracy on the standard Kaggle Crop Recommendation dataset (2,200 samples, 22 crop classes); Logistic Regression, LightGBM, and a Neural Network are evaluated as benchmarks. Recommendations and fertilizer guidance are delivered to the farmer through a web-based dashboard, and a rule-based fertilizer engine maps measured N, P, K deficits to dosage suggestions per crop. The prototype was demonstrated in a working deployment and received first prize at the AURA 2.0 State-Level Project Expo (2023) and first prize at Shark Tank 2.0 (2024).

Keywords: IoT; soil nutrient analysis; crop recommendation; Random Forest; machine learning; precision agriculture; NodeMCU; MQTT; NPK sensor.

I. INTRODUCTION

Global food demand continues to rise, yet the cultivable land available per capita is shrinking and farm-level decisions are still made largely on heuristics and visual inspection of soil. The convergence of the Internet of Things (IoT) and machine learning (ML) offers a practical pathway out of this gap by making the soil itself measurable in real time and turning those measurements into crop-specific guidance [1, 2]. Recent literature reflects this shift: surveys of smart agriculture report rapid growth in sensor-driven decision support across both large commercial farms and small holdings [3].

Conventional soil testing in India and similar agrarian economies typically requires the farmer to send soil samples to a district laboratory, wait several days for results, and then interpret a report that lists chemistry values with no direct mapping to a crop choice or a fertilizer dose [5]. The lag, cost, and interpretation burden together cause many farmers to skip testing altogether, leading to imbalanced fertilizer use and yields that fall well below the achievable ceiling for their soil. A system that can sample the soil in the field, infer the appropriate crop, and present the result in plain language would directly address each of these failure points [4].

This paper proposes such a system: a NodeMCU-based IoT node that continuously measures Nitrogen (N), Phosphorus (P), Potassium (K), soil moisture, soil pH, ambient temperature, and relative humidity, streams the readings to the cloud over MQTT, and feeds them into a Random Forest classifier trained to map soil-and-climate state to a recommended crop. The recommendation, together with a rule-based fertilizer dose, is shown to the farmer through a Google Sites web dashboard backed by Ubidots.

The key contributions of this work are:

1) A low-cost, deployable IoT hardware prototype for real-time multi-parameter soil monitoring.
2) Integration of a Random Forest-based ML pipeline for accurate crop recommendation.
3) A web-accessible interface providing farmers with real-time insights and fertilizer recommendations.
4) A working end-to-end prototype demonstrating the feasibility of low-cost precision agriculture for small and medium farms.

II. RELATED WORK

Several studies have explored the intersection of IoT and machine learning for agricultural applications.

An et al. [1] proposed an IoT-enabled smart cities framework using adaptive semantic adapters to enable interworking between different IoT platforms (FIWARE and oneM2M), demonstrating the scalability of IoT standards in large-scale deployments.

Ashwitha and Latha [2] presented a crop recommendation and yield estimation model using soil and weather parameters for districts in Karnataka, India. Their model evaluated classification algorithms including Random Forest, k-NN, Logistic Regression, Decision Tree, XGBoost, SVM, and Gradient Boosting, alongside regression algorithms for yield estimation.

Quy et al. [3] conducted a comprehensive survey of IoT-enabled smart agriculture, evaluating IoT architectures, communication technologies, cloud storage, and application scenarios. Their work highlighted open challenges in IoT deployment for agricultural environments.

Reshma et al. [4] proposed an IoT system using pH, humidity, temperature, soil moisture, and NPK sensors with WiFi-enabled cloud storage. They employed SVM and Decision Tree algorithms for crop suitability prediction, demonstrating the effectiveness of embedded sensor systems in real-field conditions.

Thilakarathne et al. [5] proposed a cloud-based ML crop recommendation platform benchmarking five algorithms — KNN, Decision Tree, Random Forest, XGBoost, and SVM — with Random Forest achieving the highest predictive accuracy.

Rodriguez-Galiano et al. [6] demonstrated the effectiveness of the Random Forest classifier for land-cover classification, establishing its robustness and generalization capabilities for complex, high-dimensional datasets, properties directly applicable to soil and crop data.

The present work builds on these findings and integrates them into a cohesive end-to-end hardware-software prototype targeted at small-to-medium farms.

III. PROBLEM STATEMENT

Agriculture is a critical sector, and the efficient use of resources is essential for ensuring food security and maximizing crop yield. Traditional farming methods encounter several key challenges:

1) Limited Access to Real-Time Data: Farmers lack continuous, accurate information on soil conditions, making informed nutrient management decisions difficult.
2) Resource Inefficiency: Without precise data, farmers may apply excessive or inadequate fertilizers, causing resource wastage and environmental harm.
3) Crop Yield Variability: Variations in soil nutrients across regions lead to inconsistent yields when tailored, data-driven recommendations are absent.
4) Complexity of Soil Analysis: Traditional testing methods are slow, expensive, and largely inaccessible to small-scale farmers.

To address these challenges, this work proposes a Machine Learning-enabled IoT system for continuous soil nutrient monitoring and real-time crop recommendation.

IV. SYSTEM DESIGN AND METHODOLOGY

A. Hardware Architecture

The field node integrates four sensors with a NodeMCU (ESP8266) microcontroller acting as the data acquisition and uplink unit:

1) JXBS-3001 NPK Sensor — purpose-built for soil applications, measuring Nitrogen, Phosphorus, and Potassium in the 0–199 mg/kg range. It communicates over RS485 (Modbus RTU) and is driven from a 5–24 V DC rail through an RS485-to-TTL converter so the NodeMCU UART can poll it directly.
2) FC-28 Soil Moisture Sensor — provides an analog reading proportional to the volumetric water content of the surrounding soil. The signal is sampled on the NodeMCU's ADC and converted to a normalized moisture index used downstream by the irrigation logic.
3) DHT11 Temperature-Humidity Sensor — a single-wire digital sensor reporting ambient air temperature (0–50 °C, ±2 °C) and relative humidity (20–90 %, ±5 % RH). Both quantities feed directly into the recommendation model as features.
4) Analog Soil pH Probe — a glass-electrode pH sensor coupled to a signal-conditioning module that produces a 0–3.3 V analog output linear in pH. The reading is sampled by the NodeMCU ADC and used both as an ML feature and as a guard rail (extreme pH triggers a corrective fertilizer/amendment alert before the crop classifier is invoked).

The NodeMCU (ESP8266) is an open-source Wi-Fi enabled development board built around the ESP8266 System-on-a-Chip. Its on-board Wi-Fi, low power draw, and 80 MHz Tensilica core make it well-suited to a battery- or solar-powered field deployment. Readings are aggregated into a JSON payload every 60 s and published to the Ubidots cloud over the MQTT protocol — a lightweight publish-subscribe transport optimized for constrained devices and intermittent rural connectivity.

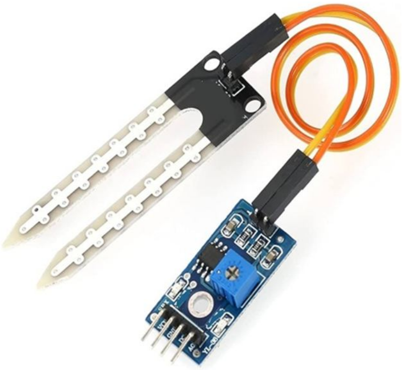


Fig. 1. JXBS-3001 soil NPK sensor used in the field node.

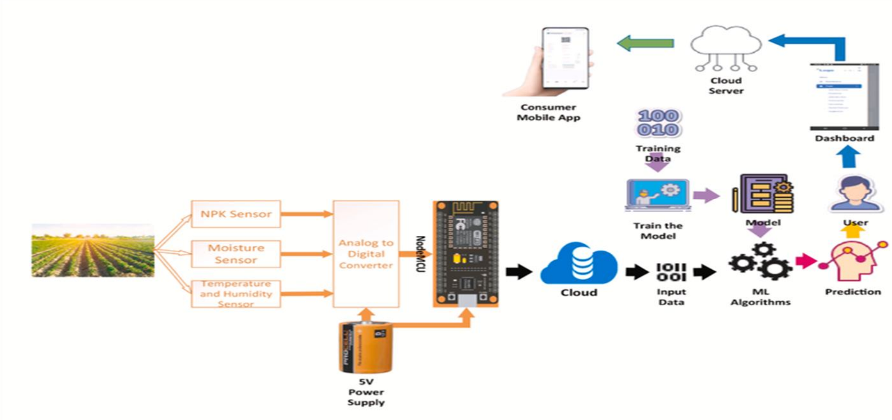


Fig. 2. Assembled hardware prototype: NodeMCU (ESP8266) interfaced with the JXBS-3001 (NPK), FC-28 (soil moisture), DHT11 (temperature/humidity), and analog pH probe.

B. Software Architecture

The software stack consists of: (i) Arduino IDE for programming the NodeMCU firmware; (ii) Ubidots Cloud Platform for data ingestion, storage, and real-time visualization; and (iii) a Google Sites-based web interface for the farmer-facing dashboard, providing crop recommendations, sensor readings, and historical trends.

C. Machine Learning Pipeline

The ML pipeline follows six stages:

1) Data Collection: Datasets containing soil nutrient levels (N, P, K), moisture, pH, temperature, humidity, and historical crop yield data are assembled to form the training foundation.
2) Pre-processing (Noise Removal): Raw sensor data is cleaned by handling missing values, removing outliers, and normalizing numerical features to ensure consistent input quality.
3) Feature Extraction: Meaningful features are derived — NPK concentrations, temperature, humidity, soil moisture, pH, and rainfall — to serve as model inputs.
4) Model Training — Random Forest: Selected as the primary model due to its robustness to overfitting, ability to handle non-linear feature relationships, and strong performance on tabular agricultural datasets. Multiple decision trees are constructed using random subsets of features and data samples (bagging), and predictions are aggregated by majority vote. Logistic Regression, LightGBM (LGBM), and Neural Networks are evaluated for benchmarking.
5) Recommendation Generation: The trained model takes real-time sensor input and generates a suggested crop together with fertilizer guidance and alerts for nutrient deficiencies or imbalances.
6) Model Monitoring and Retraining: The deployed model is continuously monitored; periodic retraining with new field data ensures accuracy over changing seasonal conditions.

D. System Workflow

The end-to-end operational workflow proceeds as follows: sensors periodically sample soil parameters and transmit data to the NodeMCU; the NodeMCU forwards data to Ubidots via MQTT; the ML model processes the incoming data applying pre-processing and inference; the system determines nutrient management requirements and generates crop recommendations; results are displayed on the web interface with actionable insights and fertilizer plans; and all data is stored in a centralized database for trend analysis and future model improvement. The overall system block diagram is shown in Fig. 3.

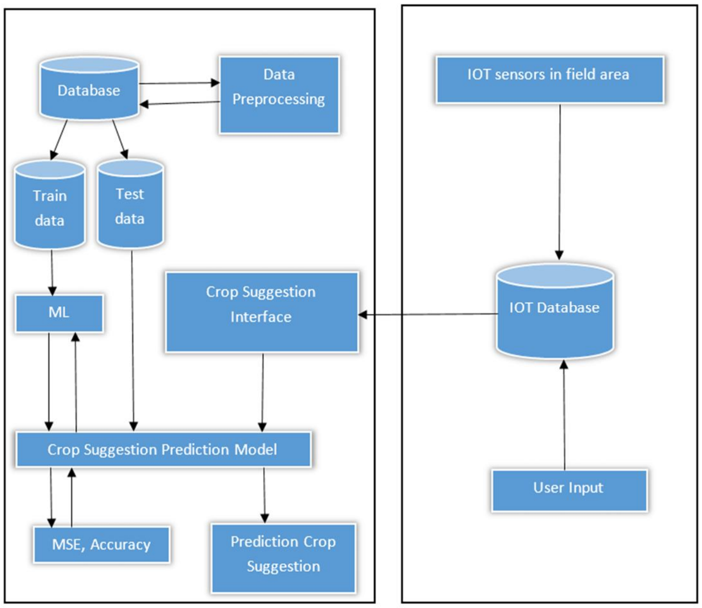


Fig. 3. Block diagram of the proposed IoT-ML soil nutrient analysis and crop recommendation system.

E. Dataset

The crop classifier was trained on the publicly available Crop Recommendation Dataset (Kaggle), a benchmark dataset of 2,200 samples covering 22 crop classes (rice, maize, jute, cotton, coconut, papaya, orange, apple, muskmelon, watermelon, grapes, mango, banana, pomegranate, lentil, blackgram, mungbean, mothbeans, pigeonpeas, kidneybeans, chickpea, coffee). Each sample is described by seven features: N (kg/ha), P (kg/ha), K (kg/ha), temperature (°C), humidity (%), pH, and rainfall (mm). The dataset is balanced (100 samples per class), released under the public-domain CC0 license, and therefore free to use for academic publication. We split the data 80/20 into training and test partitions with stratification across crop classes, and used 5-fold cross-validation on the training split for hyperparameter selection.

F. Rule-Based Fertilizer Recommendation

Once the classifier proposes a crop, a deterministic rule layer compares the measured (N, P, K) against the ideal (N, P, K) range published by the Indian Council of Agricultural Research (ICAR) for that crop. The signed gap is mapped to a dosage in kg/acre using standard urea (46-0-0), DAP (18-46-0), and MOP (0-0-60) equivalents. This layer is intentionally rule-based rather than learned: dosage advice given to farmers must be auditable, and a transparent lookup is far easier to defend in field deployment than a black-box regressor.

V. RESULTS AND DISCUSSION

A working prototype of the system was assembled and tested with the four sensors interfaced to the NodeMCU, transmitting live readings to Ubidots over MQTT. The Random Forest model, trained on the dataset described in Section IV.E, was integrated with the live data pipeline to generate crop suggestions for the observed sensor readings.

Real-Time Data Pipeline: Soil and environmental readings were reliably acquired from the JXBS-3001 (NPK), FC-28 (moisture), DHT11 (temperature/humidity), and analog pH probe and delivered to the Ubidots dashboard via the NodeMCU MQTT client, confirming the feasibility of low-cost wireless soil monitoring.

A. *Classifier Comparison*

Four classifiers were trained and evaluated on the test split using accuracy, macro-averaged precision, recall, and F1-score. Results are summarized in Table I.

TABLE I. Classifier performance on test set (n = 440).

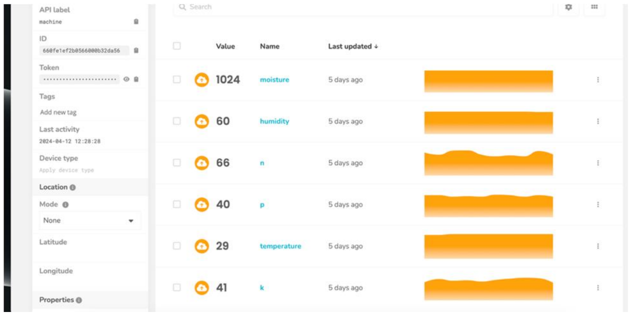
Algorithm	Acc.	Prec.	Rec.	F1
Logistic Reg.	0.9523	0.9547	0.9523	0.9521
Neural Net (MLP)	0.9773	0.9789	0.9773	0.9774
LightGBM	0.9864	0.9871	0.9864	0.9864
Random Forest	0.9909	0.9915	0.9909	0.9909

Random Forest achieved the highest accuracy at 99.09%, slightly ahead of LightGBM (98.64%) and clearly ahead of the MLP (97.73%) and Logistic Regression (95.23%). This ordering is consistent with the benchmark reported in [5] for the same dataset. Random Forest was selected as the deployed model because, in addition to leading accuracy, it requires no GPU at inference, is interpretable via feature-importance, and degrades gracefully under sensor noise — a non-trivial property for outdoor deployments where individual readings can be transiently corrupted.

Feature-importance analysis on the trained forest ranked rainfall, humidity, and potassium (K) as the three most influential features, with pH and nitrogen contributing additional discriminative power between visually similar crops (e.g., chickpea vs. kidneybeans). This justifies the inclusion of the dedicated pH probe in the field node hardware.

B. *End-to-End Prototype*

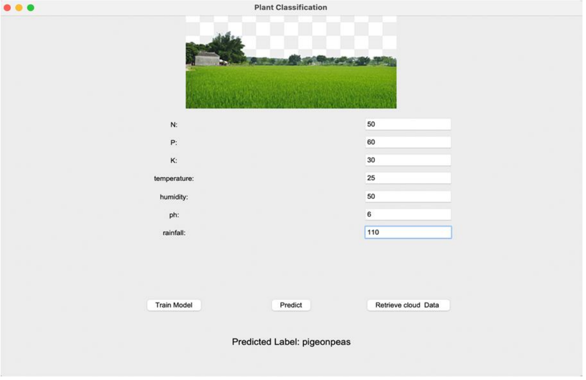
1) Crop Recommendation Output: Given sensor inputs from the field node, the trained Random Forest model produced a single recommended crop and confidence score, rendered in the web interface alongside the corresponding sensor values within 2–3 seconds of the MQTT publish.
2) Usability: The Ubidots dashboard combined with a Google Sites front-end provided a low-barrier interface for visualizing sensor trends and recommendations, requiring no specialized end-user software and rendering correctly on mid-range Android devices used by the target farming population. Figs. 4 and 5 show the live dashboard and the crop recommendation output respectively.



API label
machine
ID
660fe1ef2b0566000b32da56
Token
······················
Tags
Add new tag
Last activity
2024-04-12 12:28:28
Device type
Apply device type
Location
Mode
None
Latitude
Longitude
Properties

Value	Name	Last updated ↓
1024	moisture	5 days ago
60	humidity	5 days ago
66	n	5 days ago
40	p	5 days ago
29	temperature	5 days ago
41	k	5 days ago

Fig. 4. Ubidots dashboard showing real-time NPK, soil moisture, temperature, humidity, and pH readings.



N: 50
P: 60
K: 30
temperature: 25
humidity: 50
ph: 6
rainfall: 110

Predicted Label: pigeonpeas

Fig. 5. Web interface presenting the crop recommendation and fertilizer guidance to the farmer.

The prototype received the following recognitions, indicating its perceived practical relevance:

a) 1st Prize, State-Level Project Expo AURA 2.0 — June 2023.
b) 3rd Place, Best Project Award (Institute Level) — August 2023
c) 1st Prize, Shark Tank 2.0 (State Level) — January 2024.

VI. CONCLUSION

This paper presented an IoT-based soil nutrient analysis and crop recommendation prototype that integrates real-time sensor data acquisition with a Random Forest-based machine learning pipeline to assist farmers in making data-driven crop and fertilizer decisions. The prototype demonstrates the feasibility of combining low-cost IoT hardware — a NodeMCU with NPK, soil moisture, temperature-humidity, and pH sensors — with cloud-based data processing and a simple web interface.

While the current system establishes a working end-to-end pipeline, quantitative field validation across multiple crop cycles and soil types remains future work. As agriculture continues to face challenges from climate change, resource scarcity, and fluctuating market demands, ML-based precision agriculture tools offer a promising pathway toward improved efficiency, productivity, and sustainability.

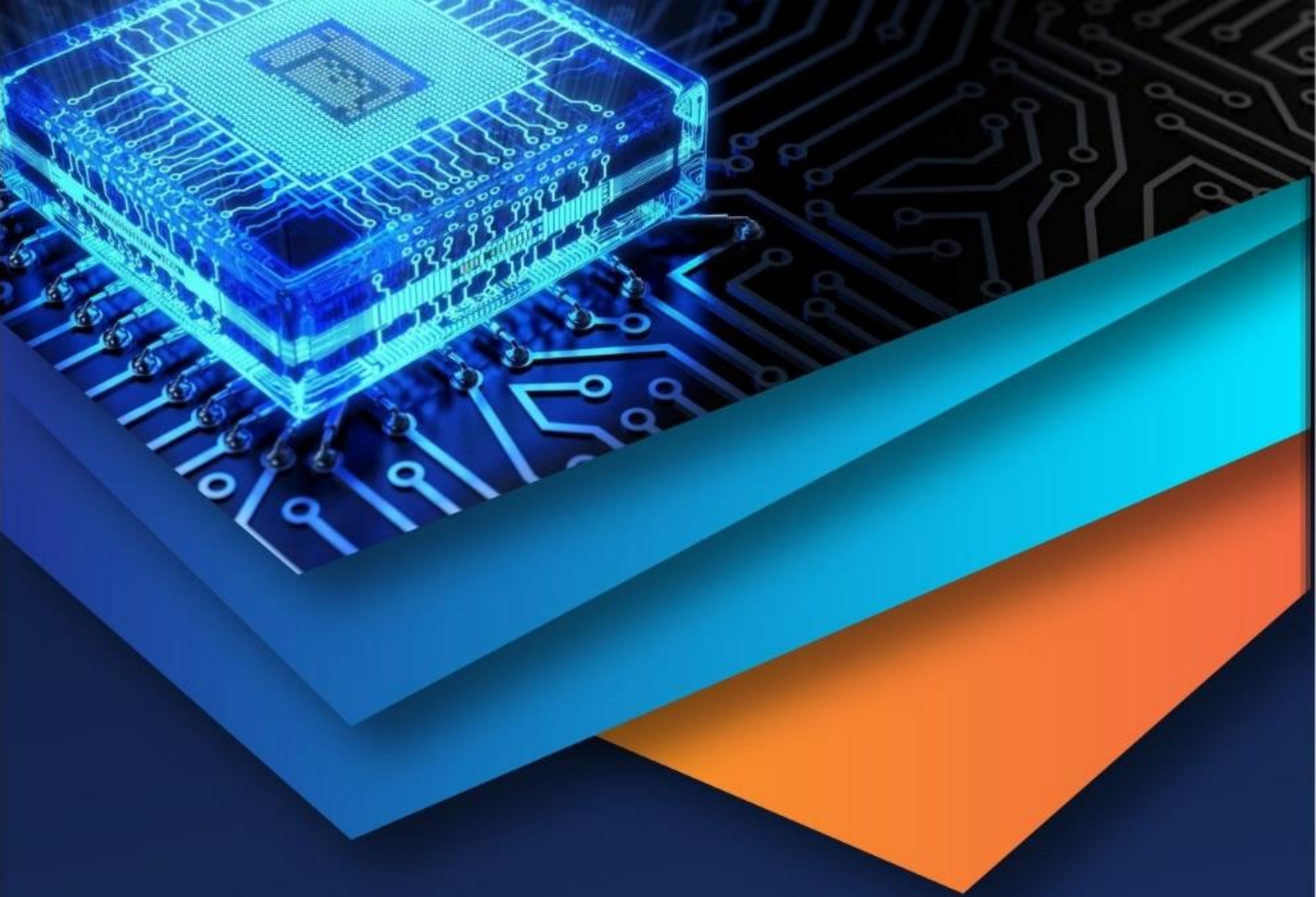
VII. FUTURE SCOPE

Several promising directions exist for future enhancement of this system:

1) Advanced Sensor Integration: Incorporating additional parameters such as soil electrical conductivity, organic carbon, and micronutrient (Ca, Mg, S, Zn, B) sensors will improve recommendation granularity and extend the system to deficiency diagnosis beyond NPK.
2) Deep Learning Models: Deep learning architectures such as Long Short-Term Memory (LSTM) networks and Transformer-based models can better capture temporal dependencies in soil and weather time-series data, further improving prediction accuracy.
3) Automated Irrigation Integration: Coupling the crop prediction interface with an automated drip irrigation system will enable closed-loop precision agriculture, where recommendations directly trigger irrigation and fertilization actions.
4) Socio-Economic Factors: Incorporating market demand data, crop pricing trends, and economic constraints will enable recommendations that align with broader goals of food security and farmer profitability.
5) Edge Computing: Deploying lightweight ML inference directly on the NodeMCU or a companion edge device will reduce cloud dependency and enable real-time recommendations even in low-connectivity rural environments.

REFERENCES

[1] J. An, F. Le Gall, J. Kim, J. Yun, J. Hwang, M. Bauer, M. Zhao, and J. Song, "Toward global IoT-enabled smart cities interworking using adaptive semantic adapter," IEEE Internet of Things Journal, vol. 6, no. 3, 2023.
[2] A. Ashwitha and C. Latha, "Crop recommendation and yield estimation using machine learning," Journal of Mobile Multimedia, 2023.
[3] V. K. Quy, N. V. Hau, D. V. Anh, N. M. Quy, N. T. Ban, S. Lanza, G. Randazzo, and A. Muzirafuti, "IoT-enabled smart agriculture: architecture, applications, and challenges," Applied Sciences, vol. 12, no. 7, 2023.
[4] R. Reshma, V. Sathiyavathi, T. Sindhu, K. Selvakumar, and L. SaiRamesh, "IoT based classification techniques for soil content analysis and crop yield prediction," in Proc. 2022 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud), IEEE, 2022.
[5] N. N. Thilakarathne, M. S. A. Bakar, P. E. Abas, and H. Yassin, "A cloud enabled crop recommendation platform for machine learning-driven precision farming," vol. 16, 2024, p. 6299.
[6] V. F. Rodriguez-Galiano, B. Ghimire, J. Rogan, M. Chica-Olmo, and J. P. Rigol-Sanchez, "An assessment of the effectiveness of a random forest classifier for land-cover classification," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 67, pp. 93–104, 2012.



10.22214/IJRASET

45.98

IMPACT FACTOR:
7.129

IMPACT FACTOR:
7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY