BORNEO JOURNAL OF SCIENCES AND TECHNOLOGY (BJoST)
VOL.: 8, ISSUE: 2, 106-121
DOI: http://doi.org/10.35370/bjost.2026.8.2-09
e-ISSN: 2672-7439, UTS Publisher.

University of Technology Sarawak

ORIGINAL ARTICLE

Self-Powered, Context-Aware and Machine Learning Enhanced Solution for Precision Agriculture

[1]Andrew Ballang Han, *[1]Keh-Kim Kee, [1]Ramli Rashidi, [1]Owen Kwong-Hong Kee, and [2]Simon Buong-Yew Lau

[1]School of Engineering and Technology, University of Technology Sarawak Sibu 96000, Malaysia
[2]Faculty of Science, University of Nottingham Malaysia, Selangor Broga 43500, Malaysia

ABSTRACT - Facing challenges from climate change, resource scarcity, and a growing population, agriculture requires sustainable and data-driven solutions. Rural and conventional farming usually lacks access to crucial data on soil, weather, and environmental conditions, while also having limited access to electricity. Many existing systems are limited to simple monitoring or rely on static data which lack integrated intelligence. This study develops a self-powered, context-aware, and machine learning-enhanced (SPoCAML) platform that seamlessly integrates IoT-based sensing with robust ML models to provide real-time crop recommendations and soil classification for precision agriculture. The system uses three-layer IoT architecture: perception layer with sensors in the field gather data, an edge layer with local hub performs data preprocessing and transmission, and a cloud-based portal lets users view and analyze the results. Designed for remote, energy-scarce environments, it is powered by a 10W solar panel and lithium-ion batteries. For its intelligence, machine learning models were built using a 2200-instance public dataset to recommend suitable crops and classify soil condition. Comparing algorithms—Logistic Regression, Neural Networks, Random Forest, and XGBoost, the Random Forest model achieved highest accuracies of 99.63% for crop recommendation and 95.65% for soil classification. Field validation confirmed the system provides over 14 hours of continuous operation, directly connecting real-time data collection to actionable farming decisions. This integrated and energy-independent platform establishes a new benchmark for intelligent and sustainable precision agriculture in resource-limited settings.

ARTICLE HISTORY
Received: 08 Oct 2025
Revised: 16 Mar 2026
Accepted: 13 Apr 2026

KEYWORDS
Agriculture,
Machine Learning,
Internet of Things,
Crops and Soils Prediction,
Soil Monitoring.

INTRODUCTION

Global food security presents one of the most urgent challenges of our time. With projections indicating a global population of 9.6 billion by 2050, food production must surge by approximately 70% compared to 2005 levels [1]. This rising demand is further intensified by the impacts of climate change, environmental degradation, and the limited supply of arable land and freshwater resources [2-4]. Conventional agricultural techniques, frequently dependent on manual effort, inherited knowledge, and the uniform application of inputs like water and fertilizers, are increasingly revealed as inefficient, unsustainable, and resource-intensive [5]. Such practices risk depleting soil nutrients, accelerating water scarcity, and diminishing harvests, thereby undermining agricultural sustainability and worsening food insecurity [6], [7].

To address these issues, precision agriculture has emerged as a revolutionary new paradigm. This approach focuses on managing variability within a field to maximize the efficiency of inputs and conserve environmental resources [8]. The convergence of the Internet of Things (IoT) and Machine Learning (ML) serves as the technological backbone of modern precision agriculture. IoT enables the real-time, granular

*Corresponding Author: Keh-Kim Kee. University of Technology Sarawak (UTS), email: kkkee@uts.edu.my

monitoring of field conditions through a network of sensors collecting data on critical parameters such as soil moisture, nutrient levels (Nitrogen, Phosphorus, Potassium), pH, temperature, and meteorological data [9].

However, the large volume and complexity of these data often exceed conventional human analytical capabilities, creating a need for intelligent processing. Therefore, ML approach provides a pivotal advantage which can learn from historical and real-time data to classify complex patterns and relationships that are not apparent to the human eye [10]. They can power predictive models for tasks such as crop yield prediction [11], disease detection [12], and crucially, for providing data-driven recommendations on optimal crop selection and soil management practices [13]. By analyzing current soil health and weather conditions, an ML model can advise a farmer on the most suitable crop to plant, thereby maximizing the probability of a successful harvest and optimizing resource use.

Despite this promise, a significant gap often exists between data collection and actionable intelligence in existing systems. Many research efforts focus on either:

1. IoT-based monitoring systems that collect real-time data but lack advanced analytical capabilities to translate it into actionable insights [14-15], or
2. ML-based recommendation models that are trained on static historical datasets and are not integrated with live data streams, limiting their practical applicability in dynamic farm environments [16-18].

Furthermore, the deployment of such technology in remote, off-grid agricultural areas is frequently hampered by the lack of a reliable power infrastructure and internet connectivity [19].

To address these research gaps, this paper presents the design, development, and validation of a self-powered and machine learning enhanced context-aware (SPoCAML) intelligent platform for precision agriculture. The core contribution of this work is the seamless integration of a robust ML-driven decision-support engine with a real-time IoT sensor network, all powered by a sustainable solar-energy module. This system moves beyond simple monitoring to provide intelligent, actionable recommendations directly to farmers. Specifically, this study:

- Develops and implements a highly accurate Random Forest-based model for crop recommendation and soil classification using real-time sensor data.
- Conducts a comparative analysis to demonstrate the superiority of the chosen algorithm against other established ML techniques.
- Designs and deploys a sustainable architecture that enables continuous off-grid operation, making it suitable for remote farming applications.
- Validates the entire system's performance and practicality through a successful field deployment.

LITERATURE REVIEW

The technological evolution of agriculture has seen a significant shift towards data-driven methodologies, primarily through the integration of the Internet of Things (IoT) for data acquisition and Machine Learning (ML) for advanced analytics. This section reviews the current state of research in IoT-based monitoring systems, ML-driven recommendation models, and the emerging trend of their integration, while identifying the specific research gaps this study aims to address.

IoT-Based Soil Monitoring Systems

A significant amount of scholarly work is dedicated to applying IoT technology for real-time agricultural monitoring. These systems constitute the essential data acquisition foundation for precision agriculture. For instance, Gaikwad et al. [15] introduced an IoT-enabled solution for continuous agro-field monitoring, delivering farmers with real-time information on soil condition. Although proficient at collecting information, the system’s dependence on wired infrastructure greatly restricts its deployment flexibility, particularly in remote and large-scale agricultural settings. Similarly, Shukla et al. [14] developed an IoT-based user-authenticated soil monitoring system utilizing sensors to measure vital parameters like pH and moisture content. Their work demonstrates the capability for immediate response to changing soil conditions. However, a key limitation acknowledged is the lack of sophisticated data analysis techniques; the system collects data but does not transform it into predictive insights or actionable knowledge, leaving the interpretation and decision-making burden entirely on the farmer.

ML-Driven Agricultural Recommendation

In parallell with IoT development, machine learning (ML) has become a powerful tool for solving agricultural problems by detecting data patterns for predictive recommendations. Sheeba et al. [20] applied an Extreme Learning Machine to classify soil types with 94% accuracy, but the model's design for a specific region in India questions its generalizability to other regions with different soil compositions and climatic conditions.

Further advancing crop-specific advice, Ikram et al. [21] employed advanced models like Random Forest to recommend optimal crops based on soil properties. Thilakarathne's cloud-enabled platform achieved 97.3% accuracy, while Ikram's model reached 97.4%. These systems provide personalized recommendations that can significantly enhance productivity. However, their primary limitation lies in their operational design; they are largely based on static, historical datasets. They lack integration with real-time, continuously flowing IoT data, which is crucial for adapting to dynamic environmental changes such as sudden rainfall, drought, or rapid nutrient depletion. This disconnection between the recommendation model and live field conditions can reduce the practical utility and timeliness of the advice provided.

The Integration of IoT and ML

The optimal trajectory for contemporary precision agriculture involves a consolidated strategy that integrates the real-time monitoring capacities of the Internet of Things with the prognostic capabilities of machine learning (ML). This integration facilitates the development of a closed-loop architecture, which is proficient not only in monitoring but also in executing intelligent responses to dynamic agro-environmental variables.

The IoT components, as demonstrated in [14-15], provide the continuous stream of real-time data on soil and environmental conditions—a fundamental requirement for any intelligent system. Meanwhile, the ML approaches reviewed in [20-22], show the proven capability to process this data, identify complex patterns, and generate predictive analytics. When combined, these technologies create a powerful feedback loop: IoT sensors feed live data to ML models, which then provide data-driven decisions. These decisions can either be presented to the farmer via a dashboard or used to automatically control actuators (e.g., irrigation valves, fertilizer dispensers) for immediate action, thereby continuously enhancing system performance and resource optimization.

However, this integration presents its own set of challenges, including ensuring robust data security, managing the complexity of system architecture, and overcoming the constraints of IoT devices themselves, such as high energy consumption and connectivity issues in remote areas [23]. Furthermore, a review of the literature reveals that while many studies propose such integration, few demonstrate a fully operational, field-tested system that combines self-powered IoT sensing with a deployed ML model for real-time recommendation. Table 1 illustrates the summary of literature review of similar research. Based on this review, several critical research gaps remain unaddressed:

1. Lack of Fully Integrated Systems: Many studies focus exclusively on either monitoring (IoT) or analysis (ML). A significant gap exists in research that demonstrates a complete, operational system that seamlessly integrates real-time IoT data acquisition with ML-driven decision-support and automated control.
2. Limited Real-World Validation: Many proposed ML models are developed and validated on static datasets but are not tested extensively in real-world, dynamic farm environments where live data streams and unpredictable conditions exist.
3. Energy Sustainability: The challenge of powering IoT systems in remote, off-grid locations is often overlooked. There is a lack of research presenting robust, self-powered solutions that ensure continuous operation without relying on existing electrical infrastructure.
4. Actionable Automation: Many systems stop providing recommendations. There is a notable gap in systems that close the loop by integrating with automation systems to perform actions like irrigation or fertilizing based on ML insights without human intervention.

This study aims to address these gaps by presenting a self-powered, context-aware and machine learning enhanced (SPoCAML) intelligent platform that not only integrates IoT and ML but also validates the entire system's performance through real-field deployment, demonstrating a practical path towards sustainable and precision agriculture.

Table 1. Literature review of similar research

Reference	Purpose / Application Area	Method	ML Enabled	IoT Enabled	Accuracy	Limitations / Challenges	Advantages
Gaikwad et al. [15]	Continuously measures variables at the agro-field.	Use IoT platforms to monitor soil health.	No	Yes	Not Mentioned	It must be with wires.	Enables continuous, real-time monitoring of soil health and management through wireless IoT communication.
Shukla et al. [14]	Soil monitoring system	IoT sensors to collect real-time data on soil conditions.	No	Yes	94%	No data analysis techniques are mentioned. lacks details on the algorithms.	Provides real-time data on soil conditions, allowing for immediate actions
Sheeba et al. [16]	Soil Classification based on micronutrients	Classifies soil types based on various soil attributes and environmental conditions.	Yes	No	94%	Restricted to a specific geographical region (Tamil Nadu).	High accuracy (94%) in classifying soil types using environmental and soil attributes.
Thilakarathne et al. [22]	To assist the farmer by developing crop recommendation platform	Recommending suitable crops based on available soil and environmental data.	Yes	No	97.3%	Doesn't apply to real-time data.	Random Forest algorithm to recommend suitable crops based on soil and environmental data, helping optimize agricultural productivity.
Ikram et al. [21]	Correct Selection of Crop	Crop selection based on soil conditions.	Yes	No	97.4%	Limited by the dataset, which includes only two soil types.	crop recommendations based on specific soil conditions, helping improve yields and optimize land use for local farming regions.

MATERIALS AND METHODOLOGY

This section details the architecture, hardware components, software development, and machine learning methodology employed in building the SPoML-CA intelligent platform for precision agriculture.

System Architecture

Figure 1 outlines the core architecture of the system, which is based on a structured three-level IoT framework. This layered approach establishes a continuous pathway from initial data gathering to the generation of actionable intelligence [24-25]. Connectivity between system elements is achieved using hypertext transfer protocol (HTTP) and application program interface (API), allowing for the live visualization and assessment of vital agricultural data points like soil health, weather patterns, and battery usage statistics.

- Perception Layer: This is the physical layer consisting of sensors and actuators that interact directly with the agricultural environment. It includes the JXBS-3001 soil sensor for data collection, and irrigation pumps as actuators. It is powered by solar PV with TP4056 charging module and 18650 lithium-ion battery.
- Edge Layer: This layer comprises the ESP-12E-based microcontroller, which acts as an edge gateway device. It is responsible for collecting raw data from the sensors, performing initial data preprocessing (e.g., unit conversion, simple filtering), and handling network connectivity via Wi-Fi to transmit data to the cloud. It also facilitates basic local decision-making, such as triggering alarms for anomalous readings.
- Cloud/Application Layer: This layer hosts the PHP -based web server and portal, MySQL database, and machine learning model. It is responsible for storing historical data, performing advanced analytics, and executing the trained ML models. The web portal provides the user interface for farmers to visualize data, receive recommendations, integration of OpenWeatherMap API and remotely control the system.
-

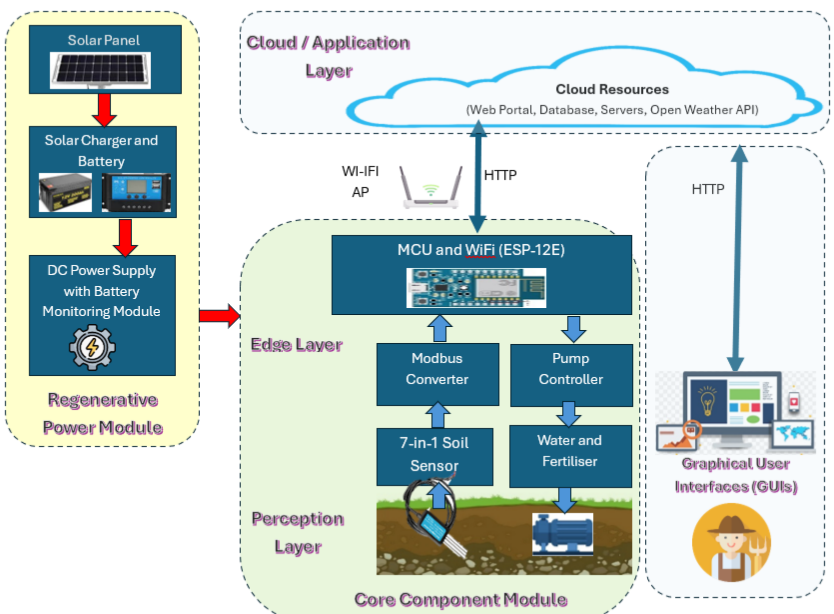


Figure 1. Block diagram of the three-layer IoT system architecture SPoCAML intelligent platform

Hardware Design

The SPoCAML consists of three core components:

1. Regenerative Power Module

A solar-powered DC source independent of the main grid. It includes a solar panel, a 10,000mAh 18650 lithium-ion battery bank, power converters (providing stable +5V and +3.3V DC), and a battery management circuit for monitoring charge/discharge status. This design ensures ease of installation and continuous 24/7 operation in remote locations. Figure 2 shows the block diagram of regenerative power module.

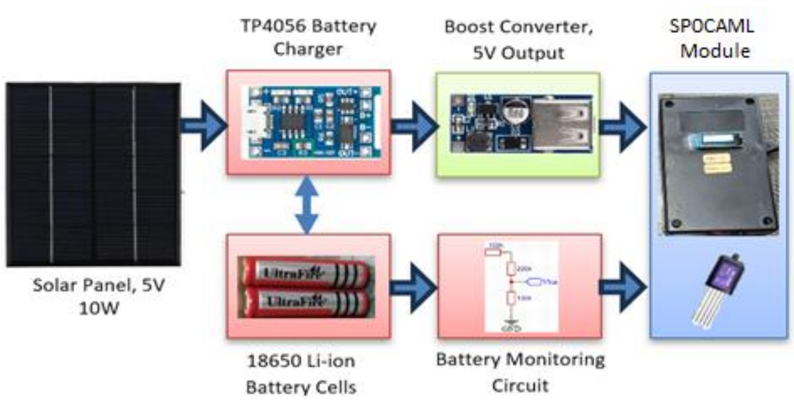


Figure 2. The block diagram of regenerative power module

2. Core Component Module

Functioning as the system's primary data hub, this module handles initial data collection and processing. Its hardware configuration centers on an ESP-12E Wi-Fi-enabled microcontroller, complemented by a JXBS-3001 soil sensor, a Modbus interface module, and a local OLED display. The incorporated JXBS-3001 sensor is an industrial-grade device, pre-calibrated to deliver precise readings for a comprehensive suite of soil conditions: nitrogen, phosphorus, potassium (NPK) concentration, moisture, temperature, electrical conductivity (EC), and pH levels [26]. The specific tolerances for these measurements are tabulated in Table 2.

Table 2. Measurement parameters (model JXBS-3001) [26]

Parameter	Range	Accuracy (Error)	Resolution
Moisture	0-100 %	3%	0.1 %
Temperature	-40~80 ℃	±0.5 ℃ (25 ℃)	0.1 ℃
EC	0~20000 us/cm	±3% F.S	10us/cm
PH	3~9 pH	±0.3 pH	0.1 pH
Nitrogen	1~1999 mg/L	2% F.S	1 mg/L
Phosphorus	1~1999 mg/L	2% F.S	1 mg/L
Potassium	1~1999 mg/L	2% F.S	1 mg/L

Data acquisition is accomplished by using an RS485-to-TTL converter and a TTL logic level shifter (from 5V to 3.3V), connected to JXBS-3001 soil sensor as illustrated in Figure 3. This setup allows the ESP-12E to act as a Modbus master, acquiring data from the JXBS-3001 soil sensor which acts as a slave unit via Modbus communication. Upon data acquisition, the data are streamlined and uploaded to the cloud layer via wireless transmission with HTTP transfer protocol.

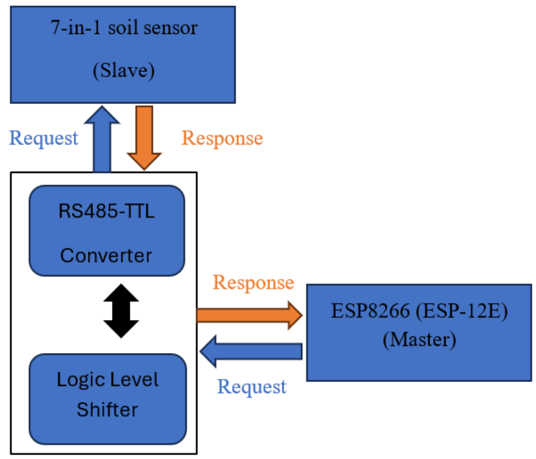


Figure 3. Modbus RS485 based Interfacing Module

3. Cloud Computing & GUIs

This component integrates a web portal, server infrastructure, and external weather data services. Users, primarily farmers, interact with the system through web browsers or mobile applications. The portal itself, a core element of the SPoCAML system, was developed with PHP and HTML and uses a MySQL database for data management. It offers an extensive toolkit for storing, accessing, and visualizing data, alongside enabling remote management of irrigation. It connects to backend systems and external services like OpenWeatherMap and a proprietary cloud server via API calls. Specifically, the portal acquires real-time meteorological information by calling the OpenWeatherMap API with the farm's location data, then parses and archives the resulting JSON data. This weather context is vital for enhancing irrigation plans and developing informed farming recommendations. The architecture of the portal is illustrated in Figure 4.

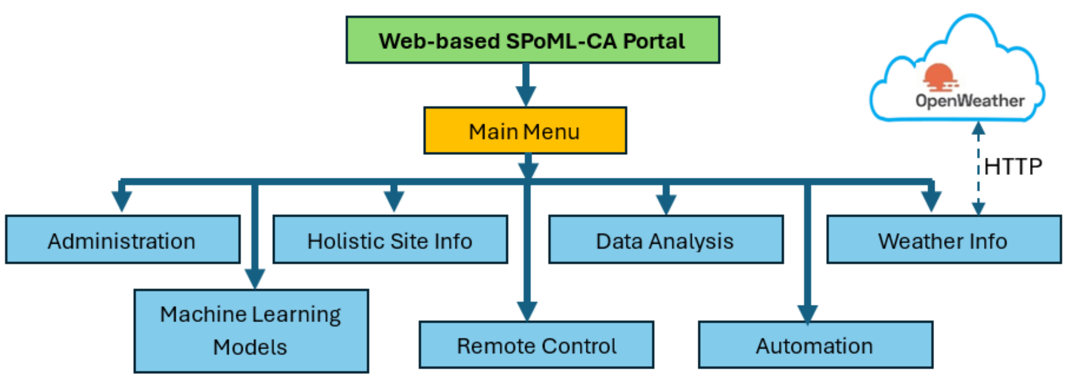


Figure 4. Web-based SPoCAML portal structure

Machine Learning Modelling Procedure

The workflow for developing the machine learning classification models followed a standard supervised learning pipeline, as shown in Figure 5. The process begins with data acquisition, which is followed by a pre-processing phase that purifies the raw sensor data to remove noise, handle missing values, and normalize measurements. Feature extraction for identifying relevant soil parameters for machine learning models to construct, i.e., crop recommendation and soil classification. After the relevant features are extracted from the data, the dataset gets partitioned into the training and testing datasets to enable an unbiased evaluation of the performance of the models. The training dataset gets used to train various machine learning and deep learning models, to enable them to learn typical features for classification models. With the training process complete, the testing dataset is used to evaluate the models and determine their prediction potential.

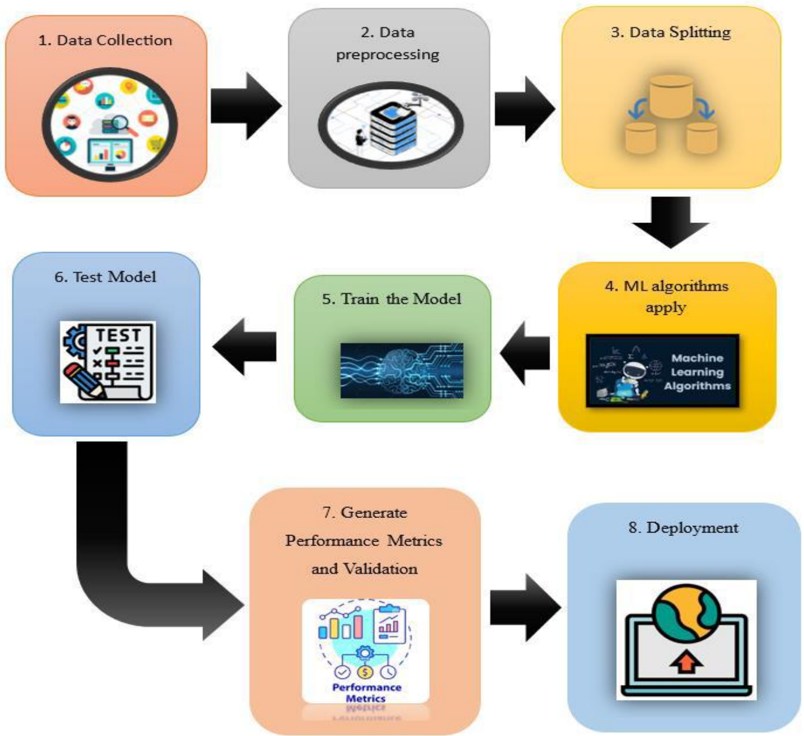


Figure 5. Flowchart of machine learning modelling procedure.

1. Data Collection and Pre-processing

The real-time dataset is acquired from both JXBS-3001 soil sensors and OpenWeatherMap server as described earlier. This dataset consists of those parameters from soil, weather and environment. A comprehensive dataset is essential for training machine learning models. In this work, the ML model was trained on a publicly available crop recommendation dataset [27], containing 2200 instances and several features: macro-nutrient level of nitrogen (N), phosphorus (P), and potassium (K), temperature, humidity, pH, electrical conductivity (EC), and rainfall. The target variable comprised 22 crop classes.

Following data acquisition, the acquired data is undergoing data preprocessing and feature extraction stages. The acquired dataset was imported into Python using the Pandas library. The preprocessing stage involved handling missing values, normalizing the numerical features to a common scale using StandardScaler, and encoding the target appliance labels. All the procedures were programmed in Python script under PyCharm's IDE software. Furthermore, the raw current waveform underwent a feature extraction process. This critical step transforms the high-dimensional raw data into a lower-dimensional, informative set of parameters that uniquely characterize the operational signature of each appliance.

The dataset was partitioned into input features (*X*) and target values (*Y*), and subsequently into 75% training dataset (*X_train, Y_train*) and 25% testing dataset (*X_test, Y_test*), to ensure unbiased evaluation of the model performance.

2. Model Selection and Training

There are numerous options of ML algorithms available for developing multiclass classification modelling in PyCharm IDE. However, the choice of algorithms is made depending on the specific characteristics of the available dataset and the performance requirements. For instance, Neural Networks (NN) can be effective for large datasets due to their capacity to learn complex patterns, while simpler models like Decision Tree (DT) may perform well with smaller datasets. In terms of model performance, there is a trade-off between accuracy and speed. If real-time applications with speed requirement, simpler models like DT may be preferred over more complex models that require more computation.

Five distinct ML classification algorithms were selected for a comparative analysis to identify the best performer for the given dataset and task. The comparison of these ML classification algorithms is tabulated in Table 3.

Table 3. Comparison of 5 selected ML classification algorithms

ML Algorithm	Characteristics	Strengths	Weaknesses	Use Cases
DT [28]	Sequential decision trees	Handles non-linear and complex relationships, and high accuracy with low bias	Overfitting problem, dense computation, and sensitive to noisy data	Fraud detection, customer churn, image classification
Logistic Regression (LR) [29]	Linear model for binary/probabilistic classification.	Highly interpretable, computationally efficient, stable & reliable	Limited to linear data, may underperform for complexity, fewer expressive capabilities	Medical diagnosis, credit scoring, social sciences
Neural Networks (NN) [30]	Multi-layered, interconnected nodes (neurons) that learn hierarchical representations of data.	High performance, high flexibility which can model extremely complex, non-linear relationships.	Less interpretable, computationally intensive, prone to overfitting, and complex tuning	Image recognition, recommendation, complex data relationships
Random Forest (RF) [31] [32]	Ensemble method that combines multiple decorrelated decision trees.	Robust to noise, handle high-dimensional data, good accuracy with low bias and low variance.	Less interpretable, may require tuning	Classification, regression, feature ranking, diverse domains
XGBoost (Extreme Gradient Boosting) [33]	Ensemble method that builds decision trees sequentially, with each tree correcting the errors of the previous one.	High predictive power, computational efficiency and Handles Missing Data Well	Prone to overfitting, more parameter tuning and more complex of interpretability	Classification, regression, feature ranking, diverse domains

3. Model Evaluation

Model performance was evaluated on the held-out test set using standard metrics such as accuracy, recall rate, precision rate, and F1-score. The calculation formulae are shown in Equation (1)-(4).

- **Accuracy:** The proportion of total correct predictions.

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \quad (1)$$

- **Precision:** The proportion of true positive predictions among all positive predictions.

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

- **Recall:** The proportion of true positives identified correctly among all actual positives.

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

- **F1-Score:** The harmonic mean of precision and recall.

$$F1 - Score = \frac{2\ X\ Precision\ X\ Recall}{Precision + Recall} \tag{4}$$

Where *TP* denotes true positive, *TN* means true negative, *FP* means false positive, and *FN* means false negative.

4. Deployment and Integration

The trained Random Forest model was serialized and saved as a .pkl file using Python's pickle module. It was then integrated into the web application's backend. A PHP script on the server receives sensor data, sends it via a POST request to a Python API endpoint (built using Flask), which loads the model, performs the prediction, and returns the recommended crop or soil class to be displayed on the farmer's dashboard in real-time.

RESULTS AND DISCUSSION

Battery charging and discharging profile

The performance of the Regenerative Power Module, specifically its charge and discharge cycles, was validated through experimentation. An analysis of the battery's behavior was conducted both theoretically and in the field. The foundation for the theoretical runtime calculation is Equation (5):

$$Device\ runtime = \frac{Battery\ Capacity\ (Wh)}{Device\ Power\ Rating\ (W)} \tag{5}$$

Powering the system is a 10000mAh 18650 lithium-ion battery bank, which holds 37.0 Wh of energy. Since the SPoCAML unit consumes an average of 2.55W i.e., 0.51A @ 5.0V, the expected runtime from a full charge is 14.51 hours.

To verify this, a real-world discharging test was performed in the absence of sunlight. The module, connected to a fully charged battery, operated from 6:00 pm on September 10, 2025, to 8:15 am the next day (see Figure 6). The measured runtime of 14 hours and 15 minutes closely matched the calculation, with only a 2.48% error. This successful test proves the 10000mAh 18650 lithium-ion battery bank can achieve the intended self-powered, sustainable design for the SPoCAML system. The operational paradigm uses solar PV for daytime power and battery charging, then switches to the stored battery energy at night. This energy-independent operation confirms the system's viability for maximizing farming efficiency while minimizing its environmental footprint.

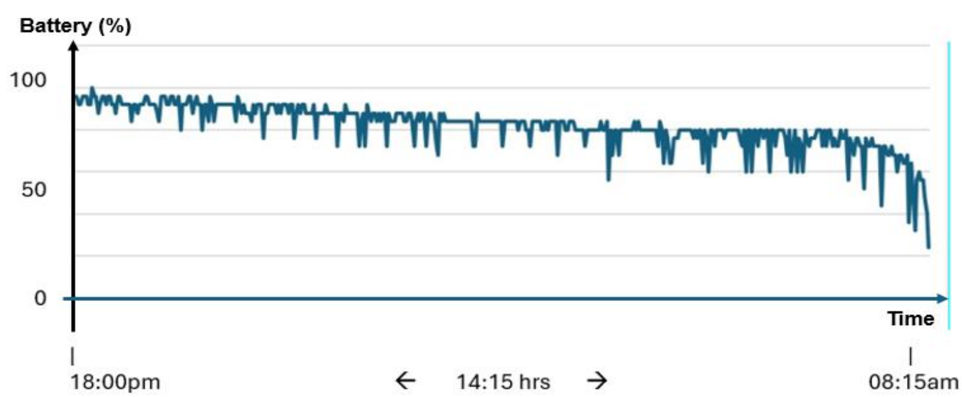


Figure 6. Battery discharging pattern and battery remaining in percentage (%).

Web-based SPoCAML portal

The result for the Web Portal of SPoCAML showcases several features, including site information, data lists for analysis and trending, location information, Open Weather Data, remote control of water valves, and automation programming. Figure 7 shows the dashboard of field collected soil information with open weather data

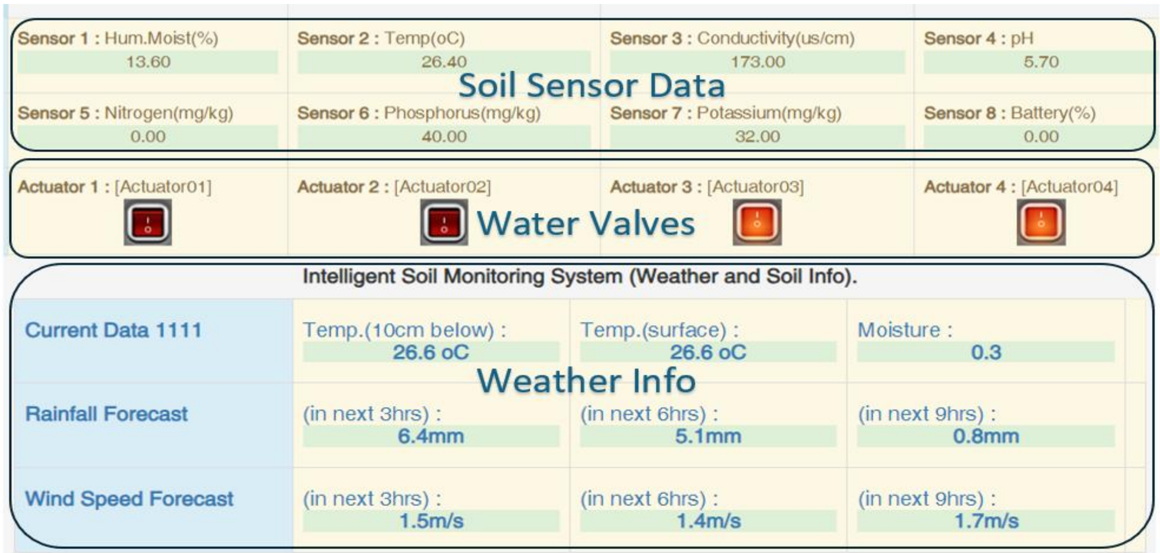


Figure 7. The dashboard of field collected soil information with open weather data

Performance of Machine Learning Models

The core intelligence of the proposed system hinges on the accuracy and reliability of its machine learning models for crop recommendation and soil classification.

1. Crop Recommendation Model

The Random Forest (RF) classifier demonstrated exceptional performance in recommending the most suitable crop based on soil and environmental parameters. The model achieved peak accuracy of 99.63% on the test set. The high F1-score of 99.09% and recall of 99.29% indicate a nearly perfect balance between precision and sensitivity, meaning the model is excellent at both correctly identifying the right crop and rarely missing the correct classification. The confusion matrix visualized in Figure 8 revealed minimal misclassifications. The few errors that occurred, e.g., misclassifying a sample of 'rice' as 'jute', were between botanically or environmentally similar crops, suggesting the model learned nuanced features rather than making random errors.

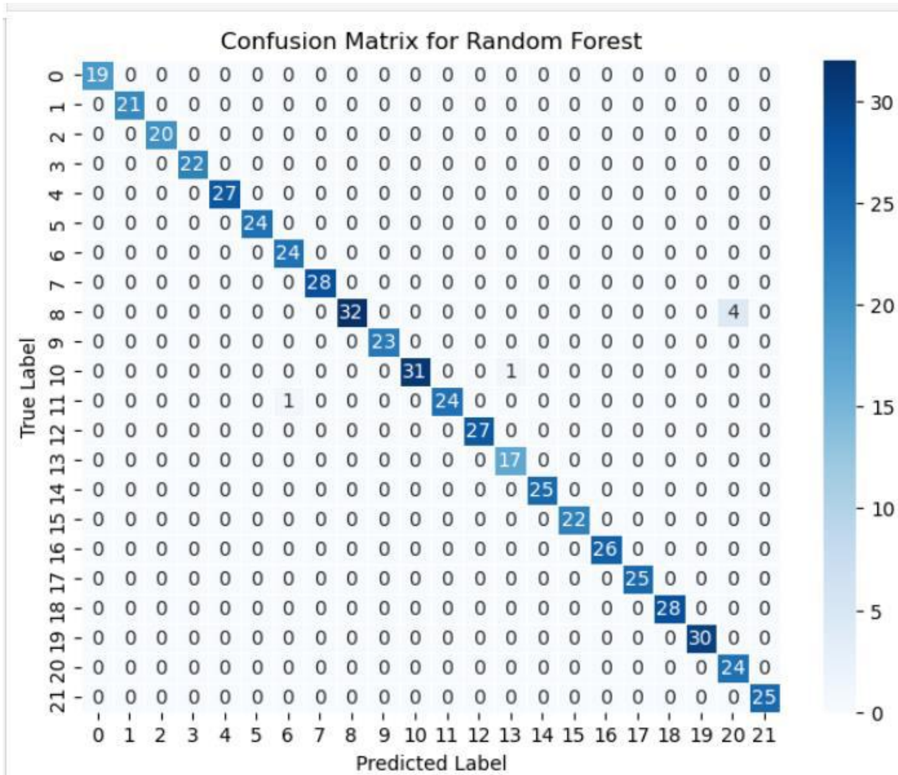


Figure 8. Confusion Matrix for Random Forest for crop

2. Comparative Analysis of ML Algorithms

To validate the choice of the Random Forest algorithm, its performance was benchmarked against three other prominent classifiers: Logistic Regression (LR), an Artificial Neural Network (ANN), and XGBoost (XGB). The results, summarized in Table 4, clearly demonstrate the superiority of the ensemble method.

Table 4. Comparative performance metrics of different machine learning algorithms for crop recommendation

Algorithm	Accuracy (%)	F1-Score (%)	Recall (%)
Logistic Regression (LR)	97.09	97.08	97.17
Neural Network (ANN)	98.54	98.54	98.62
XGBoost (XGB)	98.54	98.55	98.83
Random Forest (RF)	**99.63**	**99.09**	**99.29**

While all models performed respectably, Logistic Regression, as a linear model, struggled with the more complex, non-linear decision boundaries inherent in the agricultural data. The ANN and XGBoost performed very well, nearly matching each other, but fell just short of Random Forest's accuracy. The RF's performance can be attributed to its inherent ability to create a diverse set of decision trees that vote on the outcome, effectively reducing variance and overfitting while capturing complex interactions between features like nitrogen levels and rainfall [19]. This result aligns with findings from Thilakarathne et al. [10], who also reported high performance from tree-based models in agricultural recommendation tasks.

3. Soil Classification Model

The same ML modelling procedure was applied to classify soil types based on sensor readings. This model also performed robustly, achieving an accuracy of 95.65%. While slightly lower than the crop recommendation accuracy, this is still a highly effective result for a multi-class classification problem for soil classification task. The high performance confirms that the selected sensor parameters (N, P, K, moisture, pH, EC, temperature) contain sufficient information to accurately discern between different soil types, a finding that supports the approach taken by Sheeba et al. [20], though with a different algorithmic approach.

Field Testing and Validation

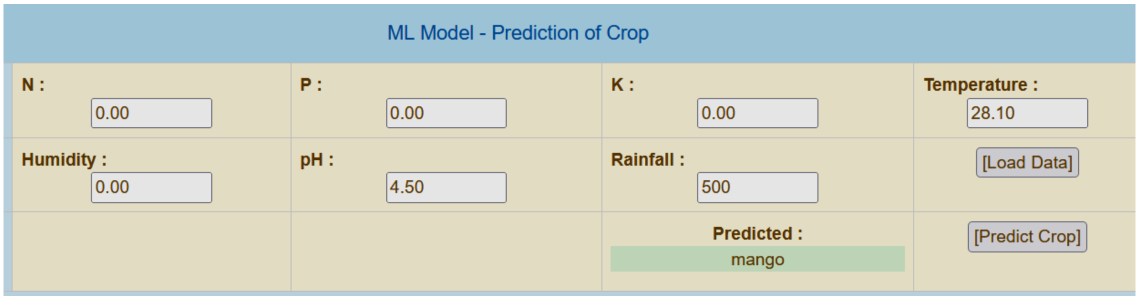
The functionality and operational efficiency of the fully integrated SPoCAML system were assessed through real-world deployment on smart farms. A key validation site was a 100.0m x 50.0m pineapple plantation in Sibu, Sarawak, Malaysia, along Jalan Teku, which acted as the experiment plot (see Figure 9a).

(a) (b) (c)

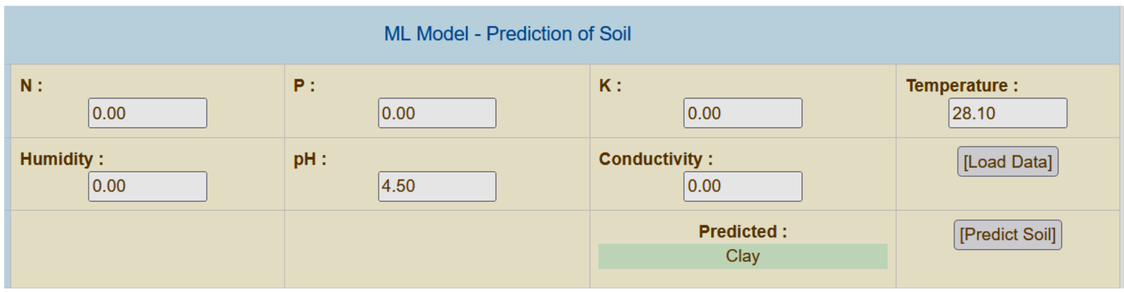
Figure 9. SPoCAML deployment for validation, (a) site location, (b) system deployment, and (c) placement of sensor.

For protection, the SPoCAML unit was installed in a compact, weatherproof (IP65) enclosure measuring 15.0cm on each side. This case contained the core electronics, while the soil moisture sensor and solar PV panels remained external. Wires for these external components were routed through a dedicated hole drilled into the enclosure. Once assembled, the unit was fixed to a PVC pipe. Figures 9(b) and 9(c) depict the final setup of the SPoCAML system within its plant box.

The web portal provides the user interface for farmers to visualize data, receive recommendations, integration of OpenWeatherMap API and remotely control the system. In the dashboard of the web portal, a new visualization was made to demonstrate feasibility of the deployed ML models, i.e., crop recommendation and soil classification, to predict the best crops to be grown and classify the farming soil type based on the current soil and environment parameters. Specifically, the PHP-scripted dashboard asks for the deployed ML prediction services by calling its application program interface (API) request, and in return, the prediction result is sent back to the dashboard for visualization purpose. Figure 10 (a) shows the crop recommendation dashboard for real time predictions of the most suitable crops. Similarly, the soil classification dashboard is depicted in Figure 10 (b) to predict the soil type based on the detected soil and environment parameters.



(a)



(b)

Figure 10. Dashboard incorporated with the deployed ML models, (a) crop recommendation and (b) soil classification.

CONCLUSION

This study successfully designed, developed, and validated a self-powered context-aware and machine learning enhanced solution for precision agriculture. The project's primary objective was to bridge the critical gap between data collection and actionable intelligence by seamlessly integrating IoT-based sensing with robust machine learning analytics, all powered by a sustainable energy source. The findings demonstrate that this integrated approach presents a viable and highly effective solution for modernizing agricultural practices and promoting sustainability.

The core achievement of this work is the development and deployment of a highly accurate machine learning models for decision support. Random Forest proved superior, delivering outstanding performance with 99.63% accuracy in recommending crops and 95.65% in classifying soil types. Such high reliability not only offers farmers a dependable tool but also fosters crucial confidence in the system's guidance. A comparative study validated that Random Forest exceeded the performance of other well-known models like Logistic Regression, Neural Networks, and XGBoost for these particular tasks.

Furthermore, a key practical challenge was resolved: operational sustainability in resource-limited settings. The incorporation of a regenerative power module facilitated uninterrupted system operation exceeding 14 hours, thereby verifying its feasibility for remote, off-grid farming scenarios. This capability for self-sufficient energy operation constitutes a significant improvement over systems requiring connection to established electrical grids, dramatically improving the prospects for widespread adoption.

In conclusion, this research contributes a comprehensive, practical, and sustainable blueprint for intelligent precision agriculture. It effectively demonstrates that the convergence of IoT, machine learning, and renewable energy can create systems that are not only technologically advanced but also accessible and beneficial for the global agricultural community.

ACKNOWLEDGEMENT

This work is fully funded by the UTS Research Grant (Project ID: UTS/RESEARCH/<3/2023/08>(01)) of the University of Technology Sarawak. The authors would like to thank the Centre of Research and Development (CRD) of UTS for its support.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] T. Daszkiewicz, "Food Production in the Context of Global Developmental Challenges," *Agriculture 2022, Vol. 12, Page 832*, vol. 12, no. 6, p. 832, Jun. 2022, doi: 10.3390/AGRICULTURE12060832.

[2] V. P. Kour and S. Arora, "Recent Developments of the Internet of Things in Agriculture: A Survey," *IEEE Access*, vol. 8, pp. 129924–129957, 2020, doi: 10.1109/ACCESS.2020.3009298.

[3] A. Chehri, H. Chaibi, R. Saadane, N. Hakem, and M. Wahbi, "A Framework of Optimizing the Deployment of IoT for Precision Agriculture Industry," *Procedia Comput. Sci.*, vol. 176, pp. 2414–2422, Jan. 2020, doi: 10.1016/J.PROCS.2020.09.312.

[4] G. S. Malhi, M. Kaur, and P. Kaushik, "Impact of Climate Change on Agriculture and Its Mitigation Strategies: A Review," *Sustainability 2021, Vol. 13, Page 1318*, vol. 13, no. 3, p. 1318, Jan. 2021, doi: 10.3390/SU13031318.

[5] T. Paepae, P. N. Bokoro, and K. Kyamakya, "From Fully Physical to Virtual Sensing for Water Quality Assessment: A Comprehensive Review of the Relevant State-of-the-Art," *Sensors 2021, Vol. 21, Page 6971*, vol. 21, no. 21, p. 6971, Oct. 2021, doi: 10.3390/S21216971.

[6] L. Delgado and J. J. Stoorvogel, "Role of soil perception and soil variability by smallholder farmers in the low adoption rates of extension packages in Central America," *J. Rural Stud.*, vol. 93, pp. 92–103, Jul. 2022, doi: 10.1016/J.JRURSTUD.2022.05.009.

[7] M. Dhanaraju, P. Chenniappan, K. Ramalingam, S. Pazhanivelan, and R. Kaliaperumal, "Smart Farming: Internet of Things (IoT)-Based Sustainable Agriculture," *Agriculture 2022, Vol. 12, Page 1745*, vol. 12, no. 10, p. 1745, Oct. 2022, doi: 10.3390/AGRICULTURE12101745.

[8] M. Sanyaolu and A. Sadowski, "The Role of Precision Agriculture Technologies in Enhancing Sustainable Agriculture," *Sustainability 2024, Vol. 16, Page 6668*, vol. 16, no. 15, p. 6668, Aug. 2024, doi: 10.3390/SU16156668.

[9] G. R. Babu, M. Gokuldhev, and P. S. Brahmanandam, "Integrating IoT for Soil Monitoring and Hybrid Machine Learning in Predicting Tomato Crop Disease in a Typical South India Station," *Sensors*, vol. 24, no. 19, p. 6177, Oct. 2024, doi: 10.3390/S24196177/S1.

[10] D. Dunsin, M. C. Ghanem, K. Ouazzane, and V. Vassilev, "A comprehensive analysis of the role of artificial intelligence and machine learning in modern digital forensics and incident response," *Forensic Science International: Digital Investigation*, vol. 48, p. 301675, Mar. 2024, doi: 10.1016/J.FSIDI.2023.301675.

[11] T. van Klompenburg, A. Kassahun, and C. Catal, "Crop yield prediction using machine learning: A systematic literature review," *Comput. Electron. Agric.*, vol. 177, p. 105709, Oct. 2020, doi: 10.1016/J.COMPAG.2020.105709.

[12] A. Dolatabadian, T. X. Neik, M. F. Danilevicz, S. R. Upadhyaya, J. Batley, and D. Edwards, "Image-based crop disease detection using machine learning," *Plant Pathol.*, vol. 74, no. 1, pp. 18–38, Jan. 2025, doi: 10.1111/PPA.14006.

[13] C. Musanase, A. Vodacek, D. Hanyurwimfura, A. Uwitonze, and I. Kabandana, "Data-Driven Analysis and Machine Learning-Based Crop and Fertilizer Recommendation System for Revolutionizing Farming Practices," *Agriculture 2023, Vol. 13, Page 2141*, vol. 13, no. 11, p. 2141, Nov. 2023, doi: 10.3390/AGRICULTURE13112141.

[14] V. Shukla, S. Dixit, and P. Dixit, "An IoT Based User Authenticated Soil Monitoring System," *Ad-Hoc and Sensor Wireless Networks*, vol. 53, no. 3–4, pp. 269–283, 2022, doi: 10.32908/AHSWN.V53.9453.

[15] S. V. Gaikwad, A. D. Vibhute, K. V. Kale, and S. C. Mehrotra, "An innovative IoT based system for precision farming," *Comput. Electron. Agric.*, vol. 187, p. 106291, Aug. 2021, doi: 10.1016/J.COMPAG.2021.106291.

[16] T. Blesslin Sheeba *et al.*, "Machine Learning Algorithm for Soil Analysis and Classification of Micronutrients in IoT-Enabled Automated Farms," *J. Nanomater.*, vol. 2022, 2022, doi: 10.1155/2022/5343965.

[17] P. S. Kiran, G. Abhinaya, S. Sruti, and N. Padhy, “A Machine Learning-Enabled System for Crop Recommendation,” *Engineering Proceedings 2024, Vol. 67, Page 51*, vol. 67, no. 1, p. 51, Sep. 2024, doi: 10.3390/ENGPROC2024067051.

[18] S. Shastri, S. Kumar, V. Mansotra, and R. Salgotra, “Advancing crop recommendation system with supervised machine learning and explainable artificial intelligence,” *Sci. Rep.*, vol. 15, no. 1, pp. 1–16, Dec. 2025, doi: 10.1038/S41598-025-07003-8;SUBJMETA.

[19] S. Srisruthi, N. Swarna, G. M. Susmitha Ros, and E. Elizabeth, “Sustainable agriculture using eco-friendly and energy efficient sensor technology,” *2016 IEEE International Conference on Recent Trends in Electronics, Information and Communication Technology, RTEICT 2016 - Proceedings*, pp. 1442–1446, Jan. 2017, doi: 10.1109/RTEICT.2016.7808070.

[20] T. Blesslin Sheeba *et al.*, “Machine Learning Algorithm for Soil Analysis and Classification of Micronutrients in IoT-Enabled Automated Farms,” *J. Nanomater.*, vol. 2022, 2022, doi: 10.1155/2022/5343965.

[21] A. Ikram *et al.*, “Crop Yield Maximization Using an IoT-Based Smart Decision,” *J. Sens.*, vol. 2022, no. 1, p. 2022923, Jan. 2022, doi: 10.1155/2022/2022923.

[22] 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,” *Sensors 2022, Vol. 22, Page 6299*, vol. 22, no. 16, p. 6299, Aug. 2022, doi: 10.3390/S22166299.

[23] S. Srisruthi, N. Swarna, G. M. Susmitha Ros, and E. Elizabeth, “Sustainable agriculture using eco-friendly and energy efficient sensor technology,” *IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology*, pp. 1442–1446, Jan. 2016, doi: 10.1109/RTEICT.2016.7808070.

[24] G. Mokhtari, A. Anvari-Moghaddam, and Q. Zhang, “A New Layered Architecture for Future Big Data-Driven Smart Homes,” *IEEE Access*, vol. 7, pp. 19002–19012, 2019, doi: 10.1109/ACCESS.2019.2896403.

[25] K. K. Kee, S. L. B. Yew, Y. S. Lim, Y. P. Ting, and R. Rashidi, “Universal cyber physical system, a prototype for predictive maintenance,” *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 1, pp. 42–49, Feb. 2022, doi: 10.11591/EEI.V11I1.3216.

[26] “JXCT-IoT Soil Monitoring | Moisture, Temperature, EC, PH, NPK.” Accessed: May 14, 2024. [Online]. Available: http://www.jxct-iot.com/product/product.php?class2=118&gclid=Cj0KCQjw0IGnBhDUARIsAMwFDLnU5u1MxjN_VpTTMVtla2mYzmzKPVEzeZbyOKBTapapr70ECnfe2HUaAn46EALw_wcB

[27] “Crop Recommendation Dataset.” Accessed: Sep. 29, 2025. [Online]. Available: https://www.kaggle.com/datasets/atharvaingle/crop-recommendation-dataset

[28] R. Joshi, “Accuracy, Precision, Recall & F1 Score: Interpretation of Performance Measures - Exsilio Blog,” 2016-09-09. [Online]. Available: https://blog.exsilio.com/all/accuracy-precision-recall-f1-score-interpretation-of-performance-measures/

[29] D. V. Bhargav and R. Dhanalakshmi, “Performance Analysis of Logistic Regression Algorithm and Random Forest Algorithm for Predicting Product Review Analysis,” *Proceedings of 9th International Conference on Science, Technology, Engineering and Mathematics: The Role of Emerging Technologies in Digital Transformation, ICONSTEM 2024*, 2024, doi: 10.1109/ICONSTEM60960.2024.10568774.

[30] Y.-H. Lin and Y.-C. Hu, “Electrical Energy Management Based on a Hybrid Artificial Neural Network-Particle Swarm Optimization-Integrated Two-Stage Non-Intrusive Load Monitoring Process in Smart Homes,” *Processes*, vol. 6, no. 12, p. 236, Nov. 2018, doi: 10.3390/pr6120236.

[31] Z. Ling *et al.*, “A Nonintrusive Load Monitoring Method for Office Buildings Based on Random Forest,” *Buildings 2021, Vol. 11, Page 449*, vol. 11, no. 10, p. 449, Sep. 2021, doi: 10.3390/BUILDINGS11100449.

[32] H. Zhang, J. Zhou, D. J. Armaghani, M. M. Tahir, B. T. Pham, and V. Van Huynh, “A Combination of Feature Selection and Random Forest Techniques to Solve a Problem Related to Blast-Induced Ground Vibration,” *Applied Sciences 2020, Vol. 10, Page 869*, vol. 10, no. 3, p. 869, Jan. 2020, doi: 10.3390/APP10030869.

[33] A. Ibrahem Ahmed Osman, A. Najah Ahmed, M. F. Chow, Y. Feng Huang, and A. El-Shafie, “Extreme gradient boosting (Xgboost) model to predict the groundwater levels in Selangor Malaysia,” *Ain Shams Engineering Journal*, vol. 12, no. 2, pp. 1545–1556, Jun. 2021, doi: 10.1016/J.ASEJ.2020.11.011.