A Review: A Robust Data-Driven Supervised Ensemble Machine Learning Framework for Soil Fertility Classification in Precision Agriculture Using Ensemble Boosting Models

Gautam Palash[1] Dr. Mohit Singh Tomar[2]
[1]Research Scholar [2]HOD (Associate Professor)
[1,2]Department of Computer Science & Engineering
[1,2]SIRT, Bhopal, India

Abstract — Soil fertility is the foundation of sustainable crop production, and its timely, accurate assessment is central to the goals of precision agriculture. Traditional laboratory-based soil testing, while reliable, is slow, costly, and difficult to scale across the spatially and temporally variable conditions found in real farmland. Supervised machine learning, and ensemble boosting methods in particular, have emerged as a practical alternative capable of learning complex, non-linear relationships between soil physicochemical attributes (nitrogen, phosphorus, potassium, pH, organic carbon, electrical conductivity, and micronutrients) and fertility class labels. This review paper surveys recent data-driven approaches to soil fertility classification, with a specific focus on ensemble boosting algorithms including Gradient Boosting, AdaBoost, Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), and Categorical Boosting (CatBoost), alongside bagging-based baselines such as Random Forest and Extra Trees. The paper synthesises reported accuracies, methodological choices (data preprocessing, class-imbalance handling, feature engineering, hyperparameter optimisation, and model interpretability), and evaluation practices across a range of recent studies. Based on this synthesis, the review identifies persistent gaps, including limited dataset diversity, inconsistent benchmarking protocols, weak generalisation across geographies, and limited use of explainable AI, and proposes a conceptual robust ensemble framework intended to address them. The review is intended to serve as a foundation for the design of a data-driven supervised ensemble boosting framework for soil fertility classification in precision agriculture.

Keywords: Soil Fertility Classification; Precision Agriculture; Ensemble Learning; Boosting Algorithms; Xgboost; Lightgbm; Catboost; Random Forest; Supervised Machine Learning; Data-Driven Agriculture

I. Introduction

Agriculture continues to be the backbone of food security and rural livelihoods across much of the world, and the productivity of any cropping system depends fundamentally on the fertility status of the soil in which it is grown. Soil fertility is not a single measurable quantity but an aggregate outcome of several interacting physical, chemical, and biological properties, most notably the primary macronutrients nitrogen (N), phosphorus (P), and potassium (K), together with soil pH, organic carbon content, electrical conductivity, and secondary and micronutrients such as sulphur, zinc, iron, manganese, and copper. Farmers and agronomists have traditionally relied on laboratory soil testing to determine these properties and to classify a given field as low, medium, or high in fertility. While accurate, this process is time-consuming, requires specialised instrumentation, and does not scale well to the fine spatial resolution that precision agriculture demands.

Precision agriculture seeks to move away from uniform, field-level management towards site-specific decision-making, applying inputs such as fertiliser, irrigation, and seed only where and when they are needed. This shift is only possible if fertility status can be estimated quickly, cheaply, and at scale, which has motivated a large and growing body of research into data-driven, machine-learning-based soil fertility prediction and classification. Supervised learning models are trained on soil-test datasets, most commonly Soil Health Card records, Kaggle-hosted agricultural datasets, or locally collected field samples, to learn a mapping between measured soil attributes and a fertility class or nutrient level.

Among the family of supervised learning techniques, ensemble methods, which combine the predictions of multiple base learners to obtain a more accurate and stable estimate than any single model, have consistently outperformed individual classifiers in this domain. Ensemble methods are broadly grouped into bagging approaches, exemplified by Random Forest and Extra Trees, which reduce variance by averaging many decorrelated trees trained on bootstrapped samples, and boosting approaches, exemplified by AdaBoost, Gradient Boosting, XGBoost, LightGBM, and CatBoost, which build an additive sequence of weak learners that iteratively correct the residual errors of their predecessors. Boosting methods, because of their capacity to model complex, non-linear interactions among soil attributes while remaining computationally efficient, have become a particularly active area of research for soil fertility classification and related precision-agriculture prediction tasks.

This review paper examines the current state of research on ensemble and boosting-based machine learning for soil fertility classification. It organises the literature around the modelling pipeline, dataset characteristics and preprocessing, algorithmic choices, evaluation metrics, and reported performance, and it uses this synthesis to motivate a robust data-driven ensemble boosting framework. The remainder of the paper is structured as follows: Section II introduces soil fertility and its role in precision agriculture; Section III reviews the use of machine learning for soil fertility assessment; Section IV details ensemble and boosting algorithms; Section V presents a structured literature review; Section VI identifies research gaps; Section VII proposes a conceptual framework; Section VIII discusses evaluation metrics; Section IX outlines challenges and future directions; and Section X concludes the paper.

II. SOIL FERTILITY AND PRECISION AGRICULTURE

Soil fertility refers to the capacity of soil to sustain plant growth by supplying essential nutrients in adequate quantity and balance, together with favourable physical structure and biological activity. Its assessment conventionally rests on measuring the primary nutrients N, P, and K, secondary nutrients such as calcium, magnesium, and sulphur, micronutrients including zinc, iron, manganese, copper, and boron, and supporting indicators such as pH, electrical conductivity, and organic carbon. Government-run initiatives such as India's Soil Health Card scheme have generated very large volumes of geo-referenced soil-nutrient records; one recent review notes that more than 23.5 crore soil health cards had been issued under this scheme by 2023, although the resulting data remains difficult to use directly for machine learning because of inconsistencies in standardisation and accessibility.

Precision agriculture uses this kind of spatially resolved soil information, together with remote sensing, IoT sensor networks, and weather data, to tailor agronomic decisions such as fertiliser dosage, crop selection, and irrigation scheduling to the specific needs of small management zones within a field rather than treating the field as uniform. The economic and environmental stakes of getting fertility estimation right are considerable: under-fertilisation depresses yield, while over-fertilisation wastes input costs and contributes to nutrient runoff and groundwater contamination. This trade-off is precisely why fast, accurate, and low-cost data-driven fertility classification has become an important research problem, motivating the shift from purely laboratory-based assessment toward machine-learning-based approaches that can be embedded into decision-support tools, mobile applications, and low-cost sensor platforms usable directly by farmers and extension workers.

III. MACHINE LEARNING IN SOIL FERTILITY ASSESSMENT

Machine learning has been applied to soil science across several related tasks: digital soil mapping, prediction of continuous nutrient values (regression), classification of fertility into discrete categories (e.g., low/medium/high), soil-moisture estimation, and crop recommendation driven by soil parameters. A wide spectrum of algorithms has been evaluated for these tasks, spanning linear models such as Logistic Regression, instance-based methods such as k-Nearest Neighbours, kernel methods such as Support Vector Machines, single tree models such as Decision Trees, ensemble tree models such as Random Forest and Extra Trees, boosting models such as Gradient Boosting, AdaBoost, XGBoost, LightGBM, and CatBoost, and neural approaches such as the Multi-Layer Perceptron.

A recurring empirical finding across this literature is that tree-based ensemble and boosting models outperform single classifiers and often outperform deep learning approaches on the relatively small, tabular datasets typical of soil-fertility studies. For instance, a comparative evaluation of sixteen classifiers on a soil fertility dataset reported that Random Forest achieved the highest overall accuracy, close to 91 percent, along with the highest precision among all models tested, indicating relatively few false-positive fertility predictions, while Extra Trees, CatBoost, and XGBoost also performed strongly, each exceeding roughly 88 percent accuracy. In a separate grading study built on an 880-sample soil dataset that incorporated SMOTE-based class-balancing, MinMax scaling, and engineered interaction features, an XGBoost classifier reached an accuracy of roughly 94 percent, and SHAP-based interpretability analysis confirmed that nitrogen, phosphorus, potassium, and pH remained the dominant predictive features, while also pointing to interaction effects among soil attributes that added further insight into nutrient dynamics.

Beyond direct fertility-class prediction, boosting methods have also proven effective for related regression tasks. In a study predicting soil organic carbon from remote-sensing-derived topographic and climatic covariates, an XGBoost model achieved a training R^2 of approximately 0.95 with a very low error rate, clearly outperforming a comparable Random Forest model, which reached a training R^2 of about 0.86 with a somewhat higher error rate on the same task. However, the same study cautioned that test-set performance degraded substantially relative to training performance, underscoring the risk of overfitting in boosting models trained on limited soil samples.

IV. ENSEMBLE LEARNING AND BOOSTING ALGORITHMS

A. Bagging versus Boosting

Ensemble methods improve on single-model predictions by combining multiple base learners, but the two dominant families do so in fundamentally different ways. Bagging (bootstrap aggregating) trains many base learners, typically decision trees, independently and in parallel on different bootstrapped subsets of the training data, then averages their predictions to reduce variance; Random Forest and Extra Trees are the most widely used bagging ensembles in the soil-fertility literature. Boosting, in contrast, trains base learners sequentially, with each new learner focused on correcting the errors made by the ensemble so far, which reduces both bias and variance but requires more careful regularisation to avoid overfitting.

B. AdaBoost and Gradient Boosting

AdaBoost (Adaptive Boosting) was among the earliest practical boosting algorithms; it works by iteratively re-weighting misclassified training samples so that subsequent weak learners, usually shallow decision stumps, focus more heavily on the hardest cases, and it combines all learners through a weighted vote. Gradient Boosting generalises this idea by fitting each successive learner to the negative gradient (residual error) of a chosen loss function, allowing it to be applied flexibly to both classification and regression. Recent work on IoT-enabled crop-prediction systems has emphasised that the choice of loss function, the type and depth of base learners, the learning rate, the number of boosting rounds, and regularisation strategies such as constraining tree size, applying shrinkage, and subsampling the training data are the key levers that jointly determine how well a gradient boosting model balances fit against generalisation, and that gradient boosting, while powerful and

flexible, can overfit quickly if these settings are not carefully tuned.

C. XGBoost, LightGBM, and CatBoost

The three most widely used modern boosting frameworks in the soil-fertility and precision-agriculture literature are Extreme Gradient Boosting (XGBoost), Light Gradient Boosting Machine (LightGBM), and Categorical Boosting (CatBoost). XGBoost extends gradient boosting with a regularised objective function, second-order gradient approximations, and efficient handling of sparse data, which together give it strong accuracy and relatively fast training on tabular soil datasets. LightGBM uses a histogram-based, leaf-wise tree-growth strategy that makes it considerably faster on large datasets while retaining competitive accuracy. CatBoost is designed to handle categorical variables natively and uses ordered boosting to reduce prediction bias, which is described as making it especially effective in applications that mix categorical and numerical features. Comparative studies bear this out: in one crop-recommendation benchmark built on soil-nutrient and environmental data, both XGBoost and LightGBM reached accuracies of roughly 99.8 percent, well ahead of a Random Forest baseline near 96 percent and several deep-learning models, while in a separate fertiliser-recommendation task built on nitrogen, phosphorus, potassium, and moisture data, CatBoost was reported as the strongest of several ensemble models evaluated, achieving close to 95 percent accuracy.

D. Stacking and Hybrid Ensembles

A further refinement combines heterogeneous base learners, often mixing bagging and boosting models, through a meta-learner in a stacking architecture. Stacking has been reported to reduce overfitting relative to single boosting models and to squeeze out additional accuracy gains; a large-scale crop-recommendation study combining eighteen base classifiers through feature fusion and stacking reported accuracy and precision exceeding comparable individual ensemble techniques while also reducing overfitting, particularly when paired with SMOTE-based class-balancing for imbalanced datasets. Similar stacking-based strategies have been applied directly to nitrogen-management decisions in precision agriculture, where an innovative stacking ensemble combining multiple data sources was used to recommend site-specific nitrogen application rates across a large agricultural region.

V. LITERATURE REVIEW

Table 1 summarises representative recent studies on ensemble and boosting-based machine learning for soil fertility and related precision-agriculture prediction tasks, highlighting the algorithms compared, the dataset used, and the key reported outcome.

#	Focus / Task	Algorithms Compared	Dataset	Key Reported Result
1	Soil fertility classification benchmark	16 classifiers incl. RF, Extra Trees, XGBoost, LightGBM, CatBoost, AdaBoost	Soil fertility dataset	Random Forest highest accuracy (90.9%); Extra Trees, CatBoost, XGBoost > 88%
2	Soil fertility grading with feature engineering	XGBoost (with SHAP)	880 soil samples (SMOTE-balanced)	XGBoost accuracy ≈ 94.3%; N, P, K, pH most influential features
3	Soil organic carbon prediction	XGBoost, Random Forest, Stacking	Satellite + soil health card data	XGBoost R^2 = 0.95 (train); RF R^2 = 0.86; stacking reduced overfitting
4	Real-time soil fertility & crop prediction	Random Forest, Extra Trees, MLP, LSTM	Sensor + lab soil data	Random Forest ≈ 92% accuracy; ensemble methods outperform single models
5	Soil nutrient enhancement review	RF, AdaBoost, Gradient Boosting, XGBoost + GA/PSO/Optuna tuning	2,000 surface soil samples (Johannesburg)	Optuna-optimised models ≥ 13% more precise than GA/PSO-tuned models
6	Crop recommendation from soil & environment data	XGBoost, LightGBM, Random Forest, GRU, LSTM	Kaggle crop recommendation (2,200 records)	XGBoost & LightGBM ≈ 99.8% accuracy; NPK ≈ 60% of feature importance
7	Stacking ensemble crop recommendation	18-classifier stacking ensemble incl. CatBoost	Kaggle + 28,242-record dataset	CatBoost 99.5% accuracy; stacking reduced overfitting
8	Soil nutrient regression comparison	CatBoost, LightGBM	Kaggle crop recommender + soil nutrients	CatBoost R^2 = 94.7% (RMSE 1.85); LightGBM R^2 = 94.1% (RMSE 1.92)
9	Precision nitrogen management for corn	Stacking ensemble (XGBoost, CatBoost, SVR, MLP)	Multi-source U.S. Corn Belt data	Stacking ensemble improved site-specific N-rate estimation over single models
10	Fertiliser type recommendation	CatBoost and other ensembles	N, P, K, moisture data (Andhra Pradesh)	CatBoost highest accuracy ≈ 94.8%

Table 1: Summary Of Representative Studies

Across these studies, several consistent patterns emerge. First, tree-based ensembles, whether bagging (Random Forest, Extra Trees) or boosting (XGBoost, LightGBM, CatBoost, Gradient Boosting), consistently outperform linear models, single decision trees, and, in most tabular soil-data settings, even deep learning architectures such as MLP and LSTM. Second, boosting models, and XGBoost in particular, are reported most frequently as the single best-performing algorithm, though the specific winner varies with dataset size, feature composition, and preprocessing choices. Third, data preprocessing steps, especially SMOTE-based handling of class imbalance, feature engineering of derived ratios and indices, and hyperparameter optimisation using genetic algorithms, particle swarm optimisation, or Optuna, have a measurable and sometimes decisive effect on final accuracy. Fourth, stacking and other hybrid ensemble architectures that combine multiple boosting and bagging learners through a meta-learner tend to reduce overfitting relative to any single ensemble model, at the cost of additional computational and design complexity. Fifth, interpretability tools such as SHAP are increasingly used alongside boosting models to confirm that predictions align with established agronomic understanding, most consistently identifying nitrogen, phosphorus, potassium, and pH as the dominant predictive features.

VI. Research Gaps

- Dataset diversity and scale: many studies rely on a small number of publicly available datasets (e.g., Kaggle crop-recommendation data or a few hundred local soil samples), which limits the geographic and soil-type generalisability of reported results.
- Inconsistent benchmarking protocols: differences in train/test splits, cross-validation strategy, and class-imbalance handling make direct accuracy comparisons across studies unreliable.
- Limited external and cross-region validation: models are rarely tested on soil data from a different agro-climatic zone than the one they were trained on, leaving real-world robustness largely unverified.
- Under-use of explainable AI: while SHAP has begun to appear in recent studies, systematic interpretability analysis is still the exception rather than the norm, which limits trust and adoption by agronomists and farmers.
- Class-imbalance and label-quality issues: fertility classes (e.g., low/medium/high) are frequently imbalanced, and label quality depends on the accuracy of the underlying laboratory tests used to construct ground truth.
- Limited integration with real-time sensing: most reported models are trained and validated offline on static datasets rather than being deployed and evaluated on live IoT or remote-sensing data streams.
- Sparse comparison of ensemble diversity strategies: relatively few studies systematically compare bagging, boosting, and stacking within a single, controlled experimental design, which makes it difficult to isolate the source of reported performance gains.
- Deployment and computational-cost considerations are seldom reported, despite their importance for low-resource, smallholder farming contexts where lightweight, edge-deployable models are needed.

VII. Proposed Conceptual Framework

Building on the patterns and gaps identified above, this review proposes a conceptual robust data-driven supervised ensemble boosting framework for soil fertility classification, structured as five stages.

1) Data Acquisition and Integration

Soil-nutrient records are collected from multiple sources, such as Soil Health Card data, field-collected laboratory samples, and, where available, IoT sensor and remote-sensing covariates, to improve dataset diversity and reduce dependence on any single data source.

2) Preprocessing and Feature Engineering

Missing values are imputed, features are scaled (e.g., MinMax or standard scaling), class imbalance is addressed using techniques such as SMOTE, and domain-informed derived features (e.g., nutrient ratios, soil acidity indices) are engineered to strengthen the signal available to downstream models.

3) Ensemble Model Training

Multiple ensemble learners, spanning bagging models (Random Forest, Extra Trees) and boosting models (AdaBoost, Gradient Boosting, XGBoost, LightGBM, CatBoost), are trained under a consistent, cross-validated protocol, with hyperparameters tuned using a systematic search strategy such as Bayesian optimisation or Optuna rather than manual tuning.

4) Stacked Meta-Ensemble and Interpretability

The best-performing base learners are combined through a stacking meta-learner to improve robustness and reduce overfitting, and SHAP-based feature-importance analysis is applied to the final model to verify agronomic plausibility and support interpretability for end users.

5) Validation and Deployment

The framework is validated using held-out data from a different region or season than the training data to test generalisation, and computational cost is profiled to determine suitability for deployment on low-cost, edge, or mobile platforms used directly by farmers and extension workers.

VIII. Evaluation Metrics

Consistent, multi-metric evaluation is essential for meaningfully comparing ensemble models on soil fertility classification, particularly given the class imbalance common in fertility datasets. The metrics most frequently used across the reviewed literature are summarised in Table 2.

Metric	Purpose
Accuracy	Overall proportion of correctly classified fertility instances; can be misleading under class imbalance.
Precision	Proportion of predicted-positive fertility class instances that are truly positive; reflects false-positive control.

Recall (Sensitivity)	Proportion of actual-positive instances correctly identified; reflects false-negative control.
F1-score	Harmonic mean of precision and recall; useful summary metric under class imbalance.
RMSE / MAE	Used for regression variants (e.g., nutrient-value prediction) to quantify average prediction error magnitude.
R^2 (Coefficient of Determination)	Proportion of variance in the target nutrient value explained by the model, used for regression tasks.
Confusion Matrix	Class-wise breakdown of correct and incorrect predictions, useful for diagnosing systematic misclassification.
SHAP / Feature Importance	Not a performance metric per se, but widely used to validate that model predictions rely on agronomically meaningful features.

Table 2: Common Evaluation Metrics for Soil Fertility Classification

IX. Challenges and Future Directions

Several practical and methodological challenges remain before ensemble boosting models can be reliably deployed at scale in precision agriculture. Data quality and availability remain the most fundamental constraint: soil-testing infrastructure is unevenly distributed, and many regions, particularly those served by smallholder farmers, lack the dense, high-quality labelled datasets needed to train and validate robust models. Related to this, class imbalance across fertility categories can bias models toward the majority class unless explicitly corrected, and the choice of balancing technique itself can materially affect reported accuracy, complicating cross-study comparison.

Model generalisation across soil types, crops, and agro-climatic zones is another open challenge; a model trained on one region's soil-health-card data may not transfer reliably to a different geography with different parent material, climate, and cropping history, yet cross-region validation is rarely performed in the current literature. Computational and deployment constraints are also under-addressed: boosting ensembles, and stacking architectures in particular, can be computationally heavier than simpler models, which matters for deployment on low-cost edge devices or smartphone applications intended for use directly by farmers, as highlighted in recent work on affordable, TinyML-oriented precision agriculture.

Looking forward, promising directions include tighter integration of ensemble boosting models with real-time IoT sensor networks and remote-sensing data streams to move beyond static, one-off soil testing toward continuous fertility monitoring; wider adoption of explainable AI techniques such as SHAP and LIME to build agronomist and farmer trust in model recommendations; systematic, standardised benchmarking protocols that would allow fair comparison of bagging, boosting, and stacking architectures across studies; and greater attention to lightweight, resource-efficient model variants suitable for deployment in low-connectivity, low-resource farming environments.

X. Conclusion

This review has surveyed the growing body of research applying supervised ensemble machine learning, and boosting algorithms in particular, to the problem of soil fertility classification in precision agriculture. The evidence across the reviewed studies consistently favours tree-based ensembles, whether bagging methods such as Random Forest or boosting methods such as XGBoost, LightGBM, CatBoost, and Gradient Boosting, over simpler linear and single-tree models, and increasingly over deep learning architectures as well, particularly on the small-to-medium tabular datasets characteristic of this domain. At the same time, the literature reveals meaningful gaps in dataset diversity, benchmarking consistency, cross-region validation, and interpretability that constrain confident real-world deployment. The conceptual robust data-driven ensemble boosting framework proposed in this review, spanning multi-source data integration, careful preprocessing, systematically tuned base learners, stacked meta-ensembling, and interpretability-aware validation, is intended to address these gaps and to provide a foundation for developing a soil fertility classification system that is both accurate and practically deployable in precision agriculture settings.

References

[1] An In-Depth Comparative Analysis of Machine Learning Models for Soil Fertility Prediction. MDPI (2026). Available at: https://www.mdpi.com/2673-4591/124/1/116
[2] Soil Analysis and Crop Fertility Prediction using Machine Learning. Available at: https://www.academia.edu/46014132/Soil_Analysis_and_Crop_Fertility_Prediction_using_Machine_Learning
[3] Random Forest Algorithm for Soil Fertility Prediction and Grading Using Machine Learning. Int. J. Innov. Technol. Explor. Eng., 9(1), 1301–1304. Available at: https://www.researchgate.net/publication/337418079
[4] Predicting soil organic carbon with ensemble learning techniques by using satellite images for precision farming. Scientific Reports (2025). Available at: https://www.nature.com/articles/s41598-025-09804-3
[5] A multi-task learning framework for soil fertility assessment and nutrient prediction using machine learning. Modeling Earth Systems and Environment, Springer Nature (2025). Available at: https://link.springer.com/article/10.1007/s40808-025-02647-x
[6] A data driven comparison of hybrid machine learning techniques for soil moisture modeling using remote sensing imagery. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC12678758/
[7] Real-time soil fertility analysis, crop prediction, and insights using machine learning and deep learning

algorithms (2025). Available at: https://www.researchgate.net/publication/397123963
[8] Machine learning-based approaches to enhance the soil fertility — A review. ScienceDirect (2023). Available at: https://www.sciencedirect.com/science/article/abs/pii/S0957417423030592
[9] Machine-Learning Classification of Soil Bulk Density in Salt Marsh Environments. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8271383/
[10] Sorn-In, K., Chansanam, W., & Netayawijit, P. (2026). Smart Crop Recommendation for Precision Agriculture: A Comparative Analysis of Ensemble and Deep Learning Models Using Soil and Environmental Data. Journal of Advances in Information Technology, 17(1), 190–202. https://doi.org/10.12720/jait.17.1.190-202
[11] High-precision crop recommendation system with stacking ensemble classifiers for optimizing agricultural productivity. Scientific Reports, 15, 43377 (2025). Available at: https://www.nature.com/articles/s41598-025-09640-5
[12] An IoT-enabled AI system for real-time crop prediction using soil and weather data in precision agriculture. ScienceDirect (2025). Available at: https://www.sciencedirect.com/science/article/pii/S2772375525004940
[13] Estimation of NPK from soil data using a novel stacked ensemble model. Multidisciplinary Science Journal (2025). Available at: https://malque.pub/ojs/index.php/msj/article/view/8184
[14] Developing an innovative stacking ensemble machine learning and multi-source data fusion-based precision nitrogen management strategy for corn. Precision Agriculture, Springer Nature (2026). Available at: https://link.springer.com/article/10.1007/s11119-026-10344-7
[15] Soil Nutrient Prediction for Precision Agriculture with ML Algorithms. The Bioscan (2024). Available at: https://thebioscan.com/index.php/pub/article/view/2996
[16] Affordable Precision Agriculture: A Deployment-Oriented Review of Low-Cost, Low-Power Edge AI and TinyML for Resource-Constrained Farming Systems. arXiv:2603.15085. Available at: https://arxiv.org/pdf/2603.15085
[17] GeaGrow: a mobile tool for soil nutrient prediction and crop recommendation. Frontiers in Sustainable Food Systems (2025). Available at: https://www.frontiersin.org/journals/sustainable-food-systems/articles/10.3389/fsufs.2025.1533423/pdf
[18] Gimba, A. M., & Mishra, P. K. (2025). Leveraging Ensemble Learning Technique for Efficient Fertilizer Recommendation. JITSI: Jurnal Ilmiah Teknologi Sistem Informasi, 6(4), 363–371. https://doi.org/10.62527/jitsi.6.4.509

