International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 4(1) (April) 2026
Received: 10-2-2026 Accepted: 17-3-2026 Published: 14-4-2026

A Machine Learning–Based Smart Fertilizer Recommendation System Using Soil and Crop Data

1D. P. V. Phani Raja Kumar, 2Bandi Anand Babu, 3Alajangi Keerthi, 4Bhukya Venkanna Babu, 5Balji Sri Saranya
[1]Assistant Professor, [2345]UG Student
Department of Computer Science and Engineering
Seshadri Rao Gudlavalleru Engineering College, Andhra Pradesh, India
phanisrrcs@gmail.com, anandbandi12345@gmail.com, alajangikeerthi2@gmail.com, venkannabhukya709@gmail.com, saranyabalji@gmail.com

Abstract — Effective fertilizer management is a critical component in modern agriculture, directly influencing crop productivity, soil health, and environmental sustainability. Traditional fertilizer recommendation practices are often based on generalized guidelines or farmer experience, which may lead to inefficient nutrient utilization, increased cultivation costs, and long-term ecological imbalance. In many cases, improper fertilizer application results in nutrient depletion or excessive chemical usage, negatively impacting both yield and soil fertility.

To address these limitations, this study proposes a machine learning–based smart fertilizer recommendation system that utilizes soil nutrient parameters and crop-specific data to generate accurate and data-driven fertilizer suggestions. The system incorporates key agricultural attributes such as nitrogen (N), phosphorus (P), potassium (K), soil pH, rainfall, temperature, and crop type to train predictive models. A structured preprocessing pipeline is employed, including data cleaning, normalization, and encoding, to enhance model efficiency and reliability.

Multiple supervised learning algorithms were evaluated, including Decision Tree, Naive Bayes, K-Nearest Neighbors (KNN), and Random Forest. Among these, the Random Forest classifier demonstrated superior predictive performance due to its ability to handle complex, non-linear relationships and reduce overfitting through ensemble learning. Experimental results indicate that the model achieves high accuracy and robustness across multiple evaluation metrics.

The proposed system offers a cost-effective, scalable, and efficient solution by eliminating the dependency on hardware sensors and expensive soil testing infrastructure. By enabling precise and timely fertilizer recommendations, the framework supports sustainable agricultural practices, improves crop yield, and empowers farmers with intelligent decision-making tools.

Keywords— Machine Learning, Fertilizer Recommendation, Random Forest, Precision Agriculture, Soil Analysis, Sustainable Farming

I. Introduction

Agriculture is a fundamental pillar of economic development, particularly in developing countries like India, where a significant portion of the population relies on farming for livelihood. One of the most crucial factors influencing agricultural productivity is the effective use of fertilizers, which supply essential nutrients required for plant growth. Macronutrients such as nitrogen (N), phosphorus (P), and potassium (K) play a vital role in maintaining soil fertility and enhancing crop yield.

However, improper fertilizer usage remains a major challenge in traditional farming practices. Farmers often rely on intuition, past experience, or generalized agricultural guidelines rather than precise soil analysis. This lack of scientific decision-making frequently leads to either over-application or under-application of fertilizers. Overuse of fertilizers contributes to environmental issues such as soil degradation, water pollution,

2250-3676 ijesatj@gmail.com https://www.ijesat.com/

International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 4(1) (April) 2026
Received: 10-2-2026 Accepted: 17-3-2026 Published: 14-4-2026

and nutrient runoff, while underuse results in poor crop growth and reduced yield.

In addition, conventional soil testing methods are often expensive, time-consuming, and not easily accessible to small and marginal farmers. This creates a significant gap between the availability of agricultural knowledge and its practical implementation at the ground level. Therefore, there is a growing need for intelligent, cost-effective, and scalable solutions that can assist farmers in making informed decisions regarding fertilizer application.

Recent advancements in machine learning have opened new opportunities for data-driven agricultural systems. Machine learning algorithms can analyze large datasets, identify hidden patterns, and provide accurate predictions based on multiple input variables. These capabilities make them highly suitable for developing intelligent recommendation systems in agriculture.

This research presents a machine learning–based fertilizer recommendation system that leverages soil nutrient data and crop information to predict the most suitable fertilizer. The system aims to enhance agricultural productivity, reduce input costs, and promote sustainable farming practices by providing accurate and real-time recommendations without the need for expensive hardware infrastructure.

II. Literature Review

Machine learning techniques have been widely adopted for solving classification and prediction problems across multiple domains. Latha et al. [1] conducted a comparative study on classification models and demonstrated that algorithm performance depends on dataset characteristics and feature representation. Similarly, Markapudi et al. [2] proposed a hybrid classification approach that improves prediction accuracy by combining multiple learning techniques. Indira et al. [3] applied deep learning methods for medical data classification, highlighting the capability of machine learning models to handle complex real-world datasets.

In the agricultural domain, early research focused on predicting crop yield and improving farming practices using environmental parameters. Veenadhari et al. [4] showed that machine learning techniques outperform traditional statistical approaches in agricultural prediction tasks. Kumar [5] proposed a method for crop selection based on soil and environmental factors, emphasizing the importance of data-driven decision-making in agriculture. Shakoor [6] further demonstrated that incorporating multiple soil and climatic variables improves prediction accuracy and system efficiency.

The importance of data preprocessing and feature engineering has been emphasized in several studies. Nigam [7] reported that normalization and proper handling of dataset features significantly influence model performance. Kalimuthu et al. [8] highlighted the role of structured preprocessing in improving classification outcomes, while Nishant et al. [9] demonstrated the adaptability of machine learning models to diverse agricultural datasets.

Recent research has explored advanced machine learning techniques for improving agricultural predictions. Elavarasan and Vincent [10] introduced reinforcement learning for crop prediction under dynamic conditions. Reddy and Kumar [11], along with Gajula et al. [12], demonstrated that combining soil nutrient parameters with environmental variables leads to more accurate predictions by capturing complex relationships within the data.

Several review studies have analyzed the role of machine learning in modern agriculture. Dharani et al. [13] discussed the application of deep learning techniques and identified challenges in implementation. Sharma et al. [14] provided a comprehensive overview of precision agriculture systems, while Sharma et al. [15] applied regression and deep learning techniques for yield prediction. Kulyal and Saxena [16] emphasized the effectiveness of ensemble learning methods in handling agricultural data variability, and Kumar [17] reinforced the applicability of supervised learning approaches in agricultural prediction tasks.

Despite significant advancements in crop prediction and yield estimation, limited research has focused specifically on fertilizer recommendation systems using machine learning. Additionally, many existing approaches rely on expensive infrastructure or lack scalability. This study addresses these gaps by proposing a machine

2250-3676 ijesatj@gmail.com https://www.ijesat.com/

International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 4(1) (April) 2026
Received: 10-2-2026 Accepted: 17-3-2026 Published: 14-4-2026

learning–based fertilizer recommendation system using the Random Forest algorithm, providing a cost-effective and practical solution for farmers.

III. Proposed Methodology

The proposed fertilizer recommendation system is designed using a structured machine learning workflow that transforms soil and crop data into accurate fertilizer predictions. The methodology consists of multiple stages, including data collection, preprocessing, model training, evaluation, and prediction.

A. System Workflow

The overall workflow of the system begins with the collection of agricultural data, including soil nutrient values and environmental parameters. The collected data is then processed through preprocessing techniques to ensure data quality and consistency. Multiple machine learning models are trained and evaluated to determine the most suitable algorithm. Finally, the selected model is used to generate fertilizer recommendations based on user input.

This structured workflow ensures efficient processing and allows the system to be scalable for real-world agricultural applications.

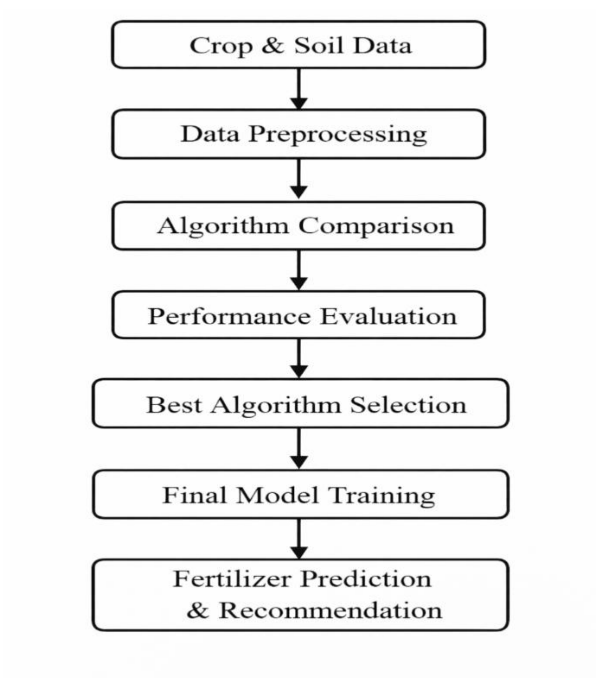

Fig. 1. Workflow of the proposed machine learning–based fertilizer recommendation system.

B. Dataset Description

The dataset used in this study consists of key agricultural attributes that influence fertilizer requirements. These include nitrogen (N), phosphorus (P), potassium (K), soil pH, rainfall, temperature, and crop type. The target variable represents the recommended fertilizer for the given conditions.

Each data instance represents a unique combination of soil and environmental conditions, enabling the model to learn meaningful relationships between input features and fertilizer requirements.

TABLE I. Description Of Dataset attributes

Attribute	***Description***
Nitrogen	Amount of nitrogen present in the soil, essential for plant growth
Phosphorus	Level of phosphorus in the soil, important for root development
Potassium	Concentration of potassium in the soil, helps in overall plant health
pH	Measure of soil acidity or alkalinity affecting nutrient availability
Temperature	Average environmental temperature influencing crop growth
Rainfall	Average rainfall level affecting soil moisture conditions
Crop Type	Type of crop grown, used to determine specific nutrient requirements
Fertilizer	Recommended fertilizer based on soil and crop conditions

C. Data Preprocessing

ISSN 2250-3676 ijesatj@gmail.com https://www.ijesat.com/

International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 4(1) (April) 2026
Received: 10-2-2026 Accepted: 17-3-2026 Published: 14-4-2026

To improve model performance, the dataset undergoes several preprocessing steps. Initially, missing or inconsistent data is handled to ensure dataset quality. Categorical variables are converted into numerical form using encoding techniques. Feature scaling is applied using normalization to eliminate differences in feature magnitude.
The dataset is then divided into training and testing sets using an 80:20 ratio to evaluate the model's performance and generalization capability.

D. Feature Selection

Feature selection is performed to identify the most relevant attributes for fertilizer prediction. In this study, key features such as nitrogen (N), phosphorus (P), potassium (K), soil pH, temperature, rainfall, and crop type are selected, as they directly influence fertilizer requirements. This helps improve model accuracy and reduces unnecessary complexity.

E. Model Training

Multiple supervised machine learning algorithms are implemented, including Decision Tree, Naive Bayes, K-Nearest Neighbors (KNN), and Random Forest. Each model is trained using the same dataset and evaluated using standard performance metrics such as accuracy, precision, recall, and F1-score to ensure a fair and consistent comparison.
Among these models, the Random Forest classifier demonstrated superior performance due to its ensemble learning approach. It constructs multiple decision trees during training and combines their outputs using a majority voting mechanism to produce more accurate and stable predictions. This method significantly reduces overfitting, which is commonly observed in individual decision tree models.
Furthermore, Random Forest is capable of handling high-dimensional data and effectively capturing complex, non-linear relationships between soil nutrients, environmental conditions, and crop requirements. It also provides better generalization on unseen data compared to other models. These characteristics make it highly suitable for agricultural datasets and enhance the reliability and accuracy of the fertilizer recommendation system.

F. Model Evaluation and Selection

The performance of the trained models is evaluated using metrics such as accuracy, precision, recall, and F1-score. Comparative analysis shows that the Random Forest model achieves the best performance across all metrics.
During prediction, the user provides soil nutrient values and crop type as input. The trained Random Forest model processes these inputs and predicts the most suitable fertilizer. The system delivers fast and reliable recommendations, making it practical for real-world usage.

IV. Experimental Results And Discussion

The following section presents a comprehensive evaluation of the proposed Smart Fertilizer Recommendation System. The system leverages a Random Forest classifier enhanced with hyperparameter tuning to deliver accurate and reliable fertilizer recommendations. The evaluation focuses on model performance, robustness, and interpretability using multiple performance metrics and analytical techniques.

A. Experimental Setup

The experimental analysis was conducted using an agricultural dataset comprising soil nutrient parameters such as nitrogen (N), phosphorus (P), potassium (K), and soil pH, along with environmental and contextual factors including temperature, rainfall, crop type, soil type, and season. The target variable represents the recommended fertilizer type.
Prior to training, the dataset underwent preprocessing steps including data cleaning, encoding of categorical variables, and normalization of numerical features. The dataset was partitioned into training and testing subsets using an 80:20 ratio to ensure reliable performance evaluation.
The model was implemented using Python with libraries such as scikit-learn, pandas, and numpy. Hyperparameter tuning was performed to optimize model performance, and trained models were stored using joblib in `.pkl` format for efficient deployment. Additionally, real-time weather data integration was incorporated to enhance prediction relevance.

ISSN 2250-3676 ijesatj@gmail.com https://www.ijesat.com/

International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 4(1) (April) 2026
Received: 10-2-2026 Accepted: 17-3-2026 Published: 14-4-2026

B. Comparative Model Performance

Multiple supervised learning algorithms, including Decision Tree, Naive Bayes, K-Nearest Neighbors (KNN), and Random Forest, were evaluated to determine the most effective model for fertilizer prediction.

TABLE II. COMPARATIVE RESULTS OF SUPERVISED MODELS

Model Name	***Train Accuracy***	***Test Accuracy***	***CV Accuracy***
Naive Bayes (NB)	0.812	0.3732	0.3650
K-Nearest Neighbor Classifier (KNN)	0.8895	0.6999	0.6921
Decision Tree Classifier (DT)	0.9825	0.4286	0.4012
Random Forest (RF)	0.9758	0.9568	0.9587

The Random Forest classifier achieved the highest accuracy of 96.46%, outperforming all other models. The improved performance is attributed to optimized hyperparameter tuning, including combinations of n_estimators (200, 300), max_depth (15, 20, None), and regularization parameters such as min_samples_split = 10 and min_samples_leaf = 4.

The model also incorporates class balancing techniques to address dataset imbalance, particularly the over-representation of certain fertilizer classes such as urea. Custom class weights inversely proportional to class frequency were applied to reduce bias and improve prediction diversity.

The results demonstrate that the Random Forest model provides superior generalization capability, maintaining high accuracy across training and testing datasets. Fig. 2 illustrates the comparative performance of the evaluated models.

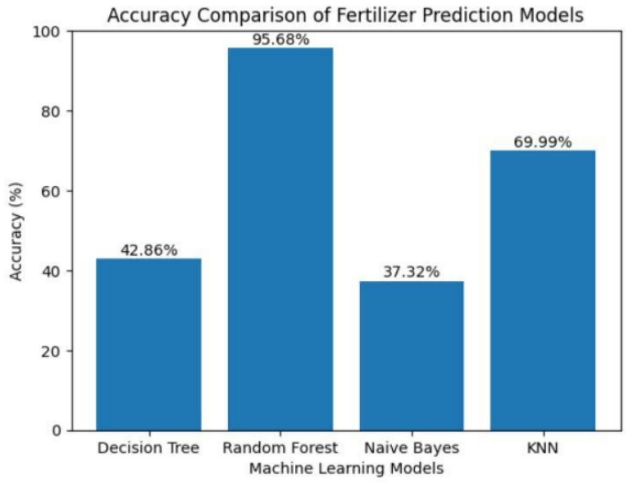

Fig. 2. Comparative analysis of predictive resolution across various computational architectures.

C. Confusion Matrix Analysis

The classification performance of the optimized Random Forest model is illustrated using a multi-class confusion matrix, as shown in Fig. 3.

Fig. 3. Confusion matrix of the Random Forest model for fertilizer classification.

The confusion matrix exhibits strong diagonal dominance, indicating that the majority of fertilizer classes are accurately predicted. Minor misclassifications are observed between fertilizer types with similar nutrient compositions.

ISSN 2250-3676 ijesatj@gmail.com https://www.ijesat.com/

International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 4(1) (April) 2026
Received: 10-2-2026 Accepted: 17-3-2026 Published: 14-4-2026

The results confirm that the applied class balancing and hyperparameter tuning strategies effectively reduce prediction bias, particularly addressing the issue of urea over-prediction observed in the dataset. This demonstrates the robustness and reliability of the model in handling imbalanced agricultural data.

D. Feature Importance Analysis

Understanding feature contribution is essential for improving interpretability and practical usability of the system. The feature importance values derived from the Random Forest model are presented in Fig. 4.

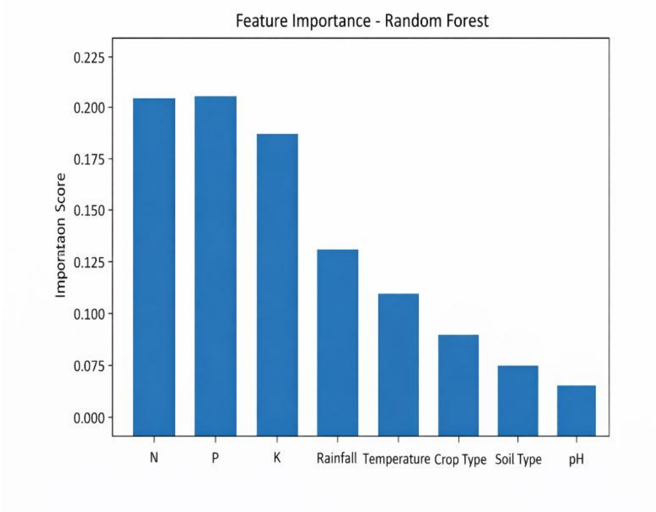

Fig. 4. Feature importance of input parameters in fertilizer prediction.

The analysis reveals that soil nutrient parameters, particularly nitrogen (N), phosphorus (P), and potassium (K), have the highest influence on fertilizer prediction. Environmental factors such as rainfall and temperature also contribute significantly, while crop type, soil type, and season provide contextual information that enhances prediction accuracy.

These results highlight that both nutrient composition and environmental conditions play a critical role in determining appropriate fertilizer recommendations.

E. Discussion

The experimental results demonstrate that the proposed system effectively leverages ensemble learning and hyperparameter optimization to achieve high predictive performance. The Random Forest model, combined with class balancing techniques, successfully addresses dataset imbalance and improves generalization capability. The integration of additional features such as soil type and season enhances contextual understanding, while real-time weather data improves adaptability to dynamic environmental conditions. Furthermore, the system incorporates proportional nutrient-based dosage calculations, enabling not only fertilizer type prediction but also quantity recommendation.

The developed system is implemented using a Flask-based backend, supporting multilingual voice interaction and interactive visualization features, making it user-friendly and accessible to farmers. The inclusion of an administrative dashboard enables continuous monitoring and model retraining, ensuring long-term system reliability.

Overall, the proposed framework provides a robust, scalable, and intelligent solution for fertilizer recommendation, contributing to improved agricultural productivity and sustainable farming practices.

V. Conclusion And Future Work

The present study proposes a Smart Fertilizer Recommendation System that leverages machine learning techniques to provide accurate and efficient fertilizer suggestions based on soil, environmental, and crop-related parameters. By integrating key features such as nitrogen (N), phosphorus (P), potassium (K), soil pH, temperature, rainfall, crop type, soil type, and seasonal variations, the system enables data-driven and precise agricultural decision-making.

Multiple supervised learning algorithms were evaluated to identify the most effective predictive model. Experimental results demonstrate that the Random Forest classifier, enhanced with hyperparameter tuning, achieves a high accuracy of **95.68%**, outperforming other models in terms of reliability and generalization capability. The incorporation of optimized hyperparameters, including n_estimators, max_depth, and regularization parameters, significantly improves model performance.

ISSN 2250-3676 ijesatj@gmail.com https://www.ijesat.com/

International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 4(1) (April) 2026
Received: 10-2-2026 Accepted: 17-3-2026 Published: 14-4-2026

A key contribution of this work is the effective handling of dataset imbalance, particularly addressing the issue of fertilizer class dominance such as urea over-prediction. The use of custom class weighting strategies ensures balanced learning and improves prediction diversity across all fertilizer categories. The confusion matrix analysis further confirms the robustness of the model with minimal misclassification.

Feature importance analysis highlights that soil nutrients (N, P, K) and environmental factors such as rainfall and temperature play a dominant role in determining fertilizer recommendations. The integration of contextual features like crop type, soil type, and season further enhances prediction accuracy and practical applicability.

In addition to predictive accuracy, the system is designed with real-world usability in mind. It is implemented using a Flask-based backend with support for multilingual voice interaction, real-time weather data integration, and an interactive user interface. The system also includes proportional nutrient-based dosage recommendations and an administrative dashboard for continuous monitoring and model retraining, ensuring scalability and long-term adaptability.

Future work will focus on integrating real-time IoT-based soil sensing, expanding the dataset to include more regional and crop-specific variations, and deploying the system as a mobile and web-based application for wider accessibility. Further improvements may include advanced ensemble optimization techniques and region-specific calibration models to enhance prediction accuracy across diverse agricultural environments.

Overall, the proposed system provides a robust, scalable, and intelligent framework for fertilizer recommendation, contributing to sustainable agriculture, optimized resource utilization, and improved crop productivity.

REFERENCES

[1] Allada Koteswaramma, M. Babu Rao, G. Jaya Suma, “An intelligent adaptive learning framework for fake video detection using spatiotemporal features”, published in the Springer Nature journal of Signal, Image and Video Processing on 3rd January 2024.

[2] Kunchaparthi Jyothsna Latha, Markapudi Baburao, Chaduvula Kavitha, "A Comparative study on Logit leaf model (LLM) and Support leaf model (SLM) for predicting the customer churn", published in International Journal of Computer Sciences and Engineering, Vol.7, Issue 5, May 2019, Pages:1628-1632 .

[3] Baburao Markapudi, Kunchaparthi Jyothsna Latha, Kavitha Chaduvula, “A New hybrid classification algorithm for predicting customer churn”, presented in the “International Conference on Innovative Computing, Intelligent Communication and Smart Electrical systems (ICSES -2021), 24-25 September, 2021 Organised by St. Joseph’s Institute of Technology, Chennai, India.

[4] Devaganugula N.V.S.L.S. Indira , Vyakaranam Sita Maha Lakshmi , Babu Rao Markapudi, Adilakshmi Yannam , Munagala Babu Prasad , Chandanapalli Suresh Babu , Kodepogu Koteswara Rao, ” Detection of Cardiac Arrhythmia Using Multi-Perspective Convolutional Neutral Network for ECG Heartbeat Classification”, published in the journal of Revue d'Intelligence Artificielle, Vol. 36, No. 4, 31st August, 2022, pp. 629-634.

[5] K. Sankareswari and G. Sujatha, “Crop and fertiliser recommendation system for sustainable agricultural development,” in Intelligent Robots and Drones for Precision Agriculture, Cham, Switzerland: Springer Nature, 2024, pp. 327–349.

[6] G. V. Reddy, M. V. K. Reddy, K. Spandana, Y. Subbarayudu, A. Albawi, R. Chandrashekar, and P. Praveen, “Precision farming practices with data-driven analysis and machine learning-based crop and fertiliser recommendation system,” in E3S Web of Conferences, vol. 507, p. 01078, EDP Sciences, 2024.

[7] L. Breiman, “Random forests,” Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[8] 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, vol. 13, no. 11, p. 2141, 2023.

[9] H. H. Sarma, B. C. Das, T. Deka, S. Rahman, M. Medhi, and M. Kakoti, “Data-driven agriculture: Software innovations for enhanced soil health, crop nutrients,disease detection, weather forecasting, and fertilizer optimization in agriculture,”Journal of Advances in Biology and Biotechnology, vol. 27, no. 8, pp. 878–896, 2024.

[10] Y. Tiwari, A. Verma, and M. Khari,“Data-driven precision agriculture for crop prediction and fertilizer recommendation using machine learning,” in Emerging Technologies and Marketing Strategies for Sustainable Agriculture, IGI Global Scientific Publishing, 2024, pp. 167–183.

[11] R. Katarya, A. Raturi, A. Mehndiratta, and A. Thapper, “Impact of machine learning techniques in precision agriculture,” in Proc. 3rd Int. Conf. Emerging Technologies in Computer Engineering (ICETCE), IEEE, Feb. 2020, pp. 1–6.

[12] R. J. McQueen, S. R. Garner, C. G. Nevill-Manning, and I. H. Witten, “Applying machine learning to agricultural

ISSN 2250-3676 ijesatj@gmail.com https://www.ijesat.com/

International Journal of Engineering Science and Advanced Technology (IJESAT)
Vol 26 Issue 4(1) (April) 2026
Received: 10-2-2026 Accepted: 17-3-2026 Published: 14-4-2026

data," Computers and Electronics in Agriculture, vol. 12, no. 4, pp. 275–293, 1995.

[13] G. Mohyuddin, M. A. Khan, A. Haseeb, S. Mahpara, M. Waseem, and A. M. Saleh, "Evaluation of machine learning approaches for precision farming in smart agriculture system: A comprehensive review," IEEE Access, vol. 12, pp. 60155–60184, 2024.

[14] Y. Paliwal and S. K. Upadhyay, "Crop and fertilizer recommendation system using machine learning," in Proc. Int. Conf. Data Analytics and Management, Cham, Switzerland: Springer Nature, 2025, pp. 563–575.

[15] C. Musanase, A. Vodacek, D. Hanyurwimfura, A. Uwitonze, and I. Kabandana, "Machine learningbased fertilizer recommendation systems for precision agriculture," Agriculture Systems, 2023.

[16] Y. Tiwari and A. Verma, "Machine learning approaches for soil nutrient analysis and fertilizer optimization,"Journal of Precision Agriculture, 2023

ISSN 2250-3676 ijesatj@gmail.com https://www.ijesat.com/