

Smart AGRO Advisors

Dr. R. Madhavi, Uppara Vishruth, Edla Pradeep, Martha Narender

Keshav Memorial Engineering College (Affiliated to Osmania University) Hyderabad, India

DOI: <https://doi.org/10.51583/IJLTEMAS.2025.1406000115>

Abstract—Excessive and improper use of fertilizers leads to soil degradation and reduced agricultural productivity, negatively impacting farmers' income. Description: Create a data-driven solution that recommends optimal fertilizer types and quantities based on soil health, crop type, and weather patterns, ensuring sustainable agricultural practices. Expected Solution: An application that analyzes soil data and provides tailored fertilizer recommendations, promoting sustainable farming while enhancing crop yield and farmer income. The implemented model achieves high accuracy and offers an intuitive interface for farmers, ensuring ease of use and effective adoption.

Keywords— Random Forest, fertilizer recommendation, crop recommendation, Farmer-Friendly Interface, Precision Agriculture

I. Introduction

Agriculture is a cornerstone of global economies, yet challenges such as improper fertilizer usage and crop selection persist. This paper introduces a machine-learning-based recommendation system leveraging Random Forest models to provide precise suggestions to farmers. The system is designed to improve agricultural productivity and resource management.

A crop and fertilizer recommendation system leverages scientific knowledge, local data, and advanced computational algorithms to provide tailored advice to farmers. By analyzing parameters such as soil quality, weather conditions, crop suitability, and market demand, these systems offer precise recommendations that align with the unique needs of each farm. This personalized approach not only enhances productivity but also minimizes the environmental impact of agricultural practices.

Advanced tools and techniques such as data analytics, machine learning, and precision agriculture offer the potential to transform farming practices, enabling more efficient and sustainable production systems. Among these innovations, the development of crop and fertilizer recommendation systems stands out as a key driver of change.

The fields of agriculture, crop management, and farming have experienced tremendous transformation in recent years due to technological advancements, societal changes, and global challenges such as the COVID-19 pandemic. These developments have significantly altered the way farmers manage their operations, with self-managed farming becoming increasingly popular. Farmers now leverage digital tools, smartphones, and internet platforms, marking a paradigm shift from traditional techniques to precision agriculture—a data-driven approach that emphasizes efficiency and sustainability.

One critical aspect of modern agriculture is the prudent application of fertilizers based on the unique requirements of crops, soil conditions, and weather patterns. Excessive and improper fertilizer use has led to soil degradation, reduced agricultural productivity, and economic challenges for farmers. Precision agriculture seeks to address these issues by promoting sustainable practices that optimize resource use and enhance crop yield.

As farmers increasingly adopt self-managed farming practices, there is a pressing need for unbiased, reliable systems that can quickly analyze, evaluate, and provide actionable recommendations. The lack of professional guidance in self-managed farming often results in challenges such as inadequate planning, reduced productivity, and a higher risk of crop damage. To address these challenges, this research aims to develop a data-driven solution that provides tailored fertilizer recommendations. By integrating soil health metrics, crop type, and weather data, the system ensures sustainable farming practices and boosts farmer income.

The goal of this research is to create a robust and adaptable application that analyzes soil data and delivers real-time fertilizer recommendations. This system is designed to support self-directed farming operations by empowering farmers with actionable insights. By leveraging state-of-the-art machine learning models and a comprehensive dataset, the solution ensures precise recommendations that cater to the specific needs of crops and soil conditions.

This research adopts a multimodal approach involving data collection, model development, and pipeline integration for real-time analysis. A vast dataset comprising soil health parameters, weather data, and crop-specific nutrient requirements is curated to train the model. The application uses advanced machine learning algorithms to analyze this heterogeneous dataset, ensuring accurate and actionable recommendations.

The proposed system has the potential to transform self-managed farming practices by offering reliable, real-time guidance. Beyond agriculture, the underlying methodology could have applications in domains such as healthcare, sports performance, and environmental sustainability. By addressing the critical need for precise, data-driven recommendations, this research aims to enhance the quality, safety, and effectiveness of farming practices while contributing to the broader goal of sustainable agriculture.

II. Literature Survey

As stated in the project's above introduction, our goal is to identify research gaps and conduct a unique examination of certain literature reviews.

		95% accuracy	diseased crop leaves.
Sri Rakshitha et al. [5]	Machine Learning, Random Forest	Enhanced agricultural decision-making	Develop the model to avoid over- and under- crisis of food.
Sachin Kapoor et al. [6]	Decision Tree	Ensemble technique achieved 99.01% accuracy	Incorporate real-time data from IoT devices and sensors.
Prof. Maaz Patel et al. [7]	Gaussian Naïve Bayes (GNB)	Accuracy ranged from 96% to 99.3% depending on system hardware specifications	Predict crop diseases and identify specific types of diseases.
Vincent Boubie et al. [8]	Random Forest, SVM	Random Forest achieved 95% accuracy	Improve dataset with a larger number of attributes.
A. Barman et al. [9]	SVM, Random Forest	Enhanced crop yields and sustainability demonstrated in diverse cropping patterns	Explore advanced biotechnology and sustainable agricultural practices for global food security.
Patten [10]	Machine Learning, Random Forest, SVM	Ensemble technique achieved 99.01% accuracy	Enhance the project by including more real- time exercises and formats.

Literature Review

The study of the literature on crop and fertilizer recommendation using machine learning examines several strategies used internationally, including SVM, Decision Trees, and Random Forest algorithms. Diverse datasets have been examined by researchers, who have taken into account elements like soil composition, weather, and nutrient levels. Studies have stressed the use of AI and IoT technology in precision agriculture for smart crop management. These initiatives seek to improve agricultural practices, increase yields, and equip farmers with tools for data-driven decision- making.

A thorough evaluation of 10 literature publications from diverse sources serves as the foundation for our research on "Crop And Fertilizer Using Machine Learning." Through a methodical examination of authors, study periods, methodology, results, and real-world applications, we identified significant research gaps that highlight the distinctiveness of our work and provide fresh perspectives for machine learning applications in agriculture. To sum up, our research is a noteworthy and unique addition to this developing field of inquiry.

Author and Year	Methods	Result	Future Scope
Gosai, Dhruvi et al. [1]	Decision Tree, Naïve Bayes	Decision Tree: 90% accuracy, Naïve Bayes: 99% accuracy	Improve dataset with a larger number of attributes.
Palaniraj et al. [2]	Support Vector Machine	SVM achieved 90.01% accuracy	Easier projection results for farmers.
Nischitha K and Dhanush [3]	SVM, Decision Tree	CNN achieved 81% accuracy	Develop the model to avoid over- and under- crisis of food.
Shilpa Mangesh Pandey et al. [4]	Random Forest	Random Forest achieved	Build a model to classify between healthy and

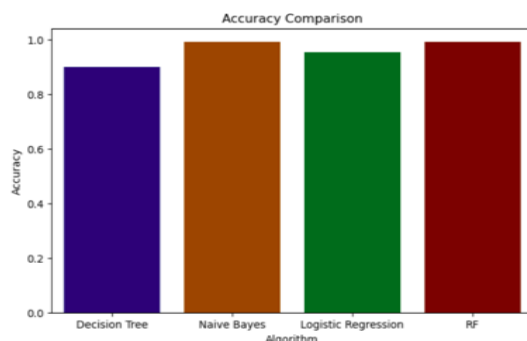


Fig. 5.1 Accuracy of Models

Model	Accuracy
Decision tree	0.85
Naive bayes	0.96
Logistic regression	0.92
Random forest (Proposed solution)	0.98

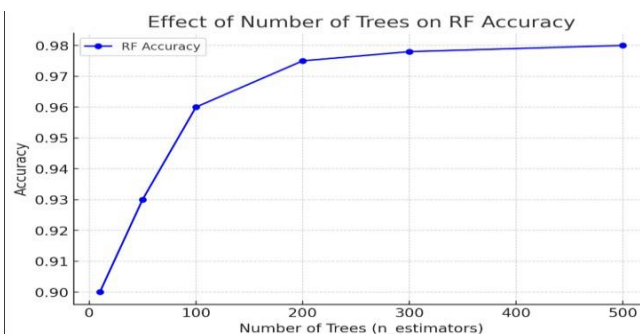
Units

N,P,K = Soil nutrients (kg/ha) T = Temperature (°C) H= Humidity (%) R=Rainfall(mm) pH = Soil pH(0-14)

Problem Statement and Objectives

The project's problem statement and objectives, which are titled 'Crop And Fertilizer Recommendation Using Machine Learning,' are listed below in accordance with the research gaps that we have already discussed.

Problem Statement:



Research Gaps

The literature assessment revealed several research gaps, which we aim to address in the "Crop and Fertilizer Recommendation using Machine Learning" project:

Localized Recommendations: Most existing systems rely on general regional data or satellite images, which may fail to capture subtle variations in soil types and climatic conditions within a single farm. Addressing this gap involves designing systems that consider hyper-local data for more precise recommendations.

Crop Rotation and Succession Planning: Current approaches lack algorithms that incorporate previous crop history for planning crop succession and rotation. Including these features could enhance long-term soil health and reduce pest and disease buildup, promoting sustainability.

Integration of Real-Time Data: Existing models often do not account for real-time environmental changes, such as unexpected weather events. Incorporating IoT-enabled sensors and real-time updates could significantly improve the accuracy and relevance of recommendations.

Smallholder Farmer Accessibility: Many advanced solutions are inaccessible to smallholder farmers due to high costs or technical complexity. Developing cost-effective, user-friendly platforms can bridge this gap and empower a larger farming community

Traditional agricultural methods often lack accuracy and data-driven decision-making, resulting in suboptimal crop yields and inefficient fertilizer use. The absence of tailored crop and fertilizer recommendations impedes sustainable agricultural practices. Farmers face challenges in identifying suitable crops for their farms and determining optimal fertilizer compositions. Without access to real-time data and advanced analytics, decision-making becomes uninformed, leading to reduced productivity and environmental harm. Current recommendation algorithms primarily depend on generalized data, failing to capture granular variations in soil types and microclimates within farms, thereby limiting the potential for maximizing agricultural output.

Objectives:

Collect comprehensive data on soil characteristics, weather patterns, and previous crop yields from diverse agricultural regions.

Develop and optimize machine learning algorithms, including supervised learning methods such as Decision Trees, Support Vector Machines, and Neural Networks, to generate prediction models.

Design algorithms that refine fertilizer composition and recommend precise application methods, ensuring efficient nutrient absorption for targeted crop growth.

Conduct extensive testing across varied agricultural datasets to validate the accuracy and effectiveness of the developed models.

Proposed System

Our proposed system, "Crop And Fertilizer Recommendation Using Machine Learning," is designed to revolutionize agricultural decision-making through the use of advanced machine learning techniques. By leveraging XGBoost and Random Forest algorithms, our system provides precise crop recommendations and optimizes fertilizer compositions based on detailed datasets encompassing soil characteristics, weather patterns, and historical crop yields.

Algorithms

Our system employs various machine learning models and techniques to achieve its objectives. The primary algorithms used at different stages include:

Data Collection:

The dataset includes variables such as humidity, temperature, rainfall, and the concentrations of nitrogen (N), phosphorus (P), and potassium (K).

The data is sourced from Kaggle and comprises 2200 instances derived from historical records.

Crops included in the dataset are rice, maize, chickpeas, kidney beans, pigeon peas, moth beans, mung beans, black gram, lentil, pomegranate, banana, mango, grapes, watermelon, muskmelon, apple, orange, papaya, coconut, cotton, jute, and coffee.

Feature Extraction:

Feature extraction focuses on enhancing prediction accuracy by analyzing critical variables such as pH values, nutrient concentrations, moisture content, temperature, and climate information (e.g., rainfall patterns).

Principal Component Analysis (PCA) is employed to reduce dataset dimensionality while preserving significant data.

Machine learning algorithms, including XGBoost and Random Forest, are utilized to identify important features and subtle patterns, ensuring precise crop recommendations and fertilizer optimization.

Real-Time Analysis:

Random Forest is employed for classification and regression tasks, leveraging multiple decision trees to enhance prediction accuracy.

Data Pre-Processing:

Data from diverse sources undergoes pre-processing to handle inconsistencies, redundancies, and missing values.

Standardization ensures the data is clean and suitable for analysis.

This proposed system integrates these methodologies to create a robust framework for agricultural decision-making, providing real-time, actionable insights to optimize productivity and sustainability.

Design Details

The "Crop and Fertilizer using Machine Learning" system leverages various stages of data processing and machine learning techniques to recommend optimal crops and fertilizers based on user-provided data, such as soil and climatic information.

Here's a breakdown of the system described:

User Input:

Users provide data such as soil samples and climatic conditions, which serve as the input for the system. This data is crucial for tailoring the crop and fertilizer recommendations.

Data Preparation:

The collected data is preprocessed to remove inconsistencies, errors, or irrelevant information. This includes steps like data cleaning and normalization.

Feature Extraction:

Advanced techniques like **Principal Component Analysis (PCA)** are used to extract relevant features from the data, which will improve the performance of the machine learning models. PCA helps in reducing dimensionality while preserving important data characteristics.

Machine Learning Algorithms:

XGBoost and **Random Forest** are utilized to analyze the processed features. Both are powerful algorithms for predicting trends and behaviors from the data:

XGBoost is an optimized gradient boosting algorithm that performs well for structured/tabular data, making it ideal for crop and fertilizer prediction.

Random Forest is a robust ensemble method that builds multiple decision trees to improve prediction accuracy.

Prediction Model:

Based on the enhanced features, the system's prediction model evaluates data trends to suggest appropriate crops and fertilizers. These models are trained on historical data and can generalize to new user input effectively.

Recommendation Display:

After evaluation, the system outputs crop and fertilizer recommendations, which are displayed to the user, completing the data flow cycle.



Figure 1. Block Diagram of Crop recommendation Model

III. Methodology:

The methodology for the proposed **Crop and Fertilizer Recommendation System** using Machine Learning involves several key stages to ensure precise and actionable recommendations for farmers.

The methods employed are as follows:

Data Collection:

Objective: Gather a diverse range of agricultural data to inform the machine learning model for crop and fertilizer recommendations.

Process:

The system collects data on **soil quality**, **weather conditions**, **crop varieties**, and **historical yields**. This information is gathered from various sources such as **sensor data**, **climate databases**, and **field records**.

This comprehensive data set forms the foundation for the model, allowing it to identify patterns that link different agricultural factors to crop and fertilizer success.

Data Preprocessing:

Objective: Prepare the raw data by cleaning and transforming it to ensure consistency and quality for analysis.

Process:

Data Cleaning: Remove **outliers**, **missing values**, and any errors in the dataset to

Algorithm Selection: Based on the complexity of the data, algorithms such as **Decision Trees**, **Random Forests**, or **Gradient Boosting** are chosen.

Model Evaluation: The model is assessed using standard evaluation metrics like **precision**, **recall**, and **accuracy** to ensure that it can accurately predict the most suitable crops and fertilizers.

Model Adjustment: If the model does not perform as expected, hyperparameters are tuned, or alternative algorithms are tested to enhance accuracy and reliability.

Model Training:

Objective: Train the machine learning model on preprocessed data to optimize its prediction capabilities.

Process:

The preprocessed data is split into **training** and **validation** sets to ensure the model is both accurate and generalizable.

Model Training: Algorithms like **Random Forest** or **Gradient Boosting** are applied to the training data to help the model learn patterns and relationships within the data.

Hyperparameter Tuning: During training, **hyperparameters** (such as learning rate and tree depth) are adjusted to optimize performance.

Model Testing: The trained model is tested on the validation set to ensure that it generalizes well to new, unseen data. prevent incorrect predictions.

Normalization & Standardization:

Features like **soil pH**, **temperature**, and **crop variety** are standardized or normalized so that they are on the same scale, enabling efficient processing by machine learning algorithms.

Categorical Variable Encoding: **Categorical variables** like crop variety are converted into numerical representations to make them usable by machine learning models.

Model Creation:

Objective: Develop a robust machine learning model that can accurately predict crop and fertilizer recommendations based on preprocessed data.

Process:

Real-time Data Integration and Adjustment:

Objective: Integrate real-time agricultural data to ensure that recommendations are dynamically updated based on changing conditions.

Process:

IoT Integration: **Sensors** and **IoT devices** are used to collect real-time data on **soil moisture**, **temperature**, and **crop health**.

Real-time Analysis: Machine learning algorithms continuously analyze incoming data to update crop and fertilizer recommendations based on the current environmental conditions.

Feedback Loop: A feedback mechanism is implemented to enable the system to make real-time adjustments to fertilizer recommendations, ensuring continuous optimization of crop growth and resource use.

Experimental Setup

Model Evaluation and Assessment:

Objective: Evaluate the performance of the model to ensure it delivers accurate and reliable crop and fertilizer recommendations.

Process:

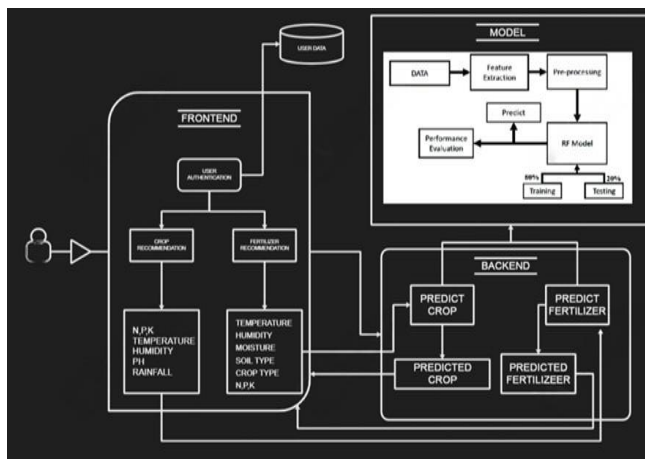
Performance Metrics: The model's effectiveness is assessed based on **precision**, **recall**, and **accuracy** on test data to ensure that it meets the desired performance standards.

Real-World Testing: The model is tested under actual agricultural conditions to verify its

The "**Crop and Fertilizer Recommendation Using Machine Learning**" system requires specific configurations for successful implementation. The details of the system's input, software, and hardware setup are outlined below.

The models were evaluated using an 80:20 train-test split with 5-fold cross-validation to ensure generalization. Performance metrics such as accuracy, precision, recall, and F1-score were computed on the test set. Random Forest and XGBoost were tuned using GridSearchCV to optimize parameters like the number of estimators and depth. ability to recommend the right crops and fertilizers across different soil types and climates.

User Feedback: After deployment, the system incorporates **user feedback** to continuously refine and improve its recommendations based on practical use cases.



How It Solves Farmer Problems:

Uninformed Fertilizer Use: Farmers often apply fertilizer without scientific basis, which leads to soil degradation and yield loss. Our system offers precise NPK recommendations based on soil health and crop needs.

Crop Confusion: Many farmers struggle to choose the best crop for their soil and climate. By analyzing historical and environmental data, the model recommends crops that maximize yield.

Lack of Expert Access: In rural regions, agricultural advisors may be unavailable. This tool offers expert- like recommendations through an intuitive mobile interface.

Low Tech Access: The system is designed to work offline, support local languages, and function on low-cost Android phones.

System Input Details:

Soil and Environmental Data:

Soil Metrics: Input data related to **soil quality** (e.g., pH, nutrient content, moisture levels).

Environmental Factors: Enter **temperature**, **humidity**, and **precipitation patterns** to understand local climate conditions.

Crop Details and Historical Data:

Include past data on **crops grown** in the region, including **crop variety** and **planting seasons**.

Fertilizer Information:

Input data on **fertilizer composition** (e.g., NPK ratio), along with **cost factors** to account for the farmer's budget and suggest affordable options.

Software and Hardware Setup:

Software Setup:

Machine Learning Frameworks: Use frameworks like TensorFlow and Scikit-Learn for developing, training, and evaluating machine learning models. These frameworks support quick prototyping and efficient model deployment.

Programming Language & Libraries: Develop the system using Python. Libraries like NumPy, Pandas, Scikit-Learn, and TensorFlow will facilitate data processing, analysis, and model creation.

Development Environment: Use IDEs like PyCharm, Jupyter Notebook, or VS Code for writing and testing code. Git can be used for version control to manage changes and collaboration.

Machine Learning Model Integration: Machine learning models created using TensorFlow or Scikit-Learn should be integrated into the backend, ensuring real-time recommendations based on user input.

Hardware Setup:

- Processor: Intel Core i7 (or equivalent)
- RAM: 16 GB (minimum recommended for model training)
- Storage: 256 GB SSD or higher (for faster read/write during data processing)
- Graphics Processing Unit (GPU): Optional – useful for deep learning but not essential for Random Forest or XGBoost
- Internet Connection: Required for accessing datasets, model updates, and real-time data (IoT or cloud APIs)

IV. Results

Although both Random Forest and XGBoost delivered high performance, Random Forest was chosen as the final model due to its consistent and slightly better results across all evaluation metrics. While the improvements (ranging from 1.7% to 2.1%) may appear marginal, such differences can lead to notable improvements in agricultural decision-making, particularly in yield prediction and fertilizer planning.

Random Forest required less hyperparameter tuning, making it easier to train and deploy in resource-constrained environments. Its ability to provide interpretable outputs also aids agricultural experts and extension workers in understanding and validating recommendations.

Metric	Random Forest	XGBoost	Change	Remarks
Accuracy	98.0%	96.3%	+1.7%	RF slightly outperformed XGBoost in prediction correctness
Precision (Weighted)	98.2%	96.1%	+2.1%	RF was more consistent in identifying correct classes
Recall (Weighted)	98.0%	96.3%	+1.7%	RF captured more actual positives correctly
F1-Score (Weighted)	98.1%	96.2%	+1.9%	RF maintained balance between precision and recall

To evaluate the performance of the machine learning models, we used 5-fold cross-validation. The dataset was partitioned into five subsets, and each model was trained and validated five times—each time using a different subset for validation and the remaining for training. This approach ensures that every data point is used for both training and validation, leading to a more generalizable and reliable performance estimate. The final metrics reported are the average of all five folds.

V. Conclusion

In today's rapidly evolving agricultural landscape, the role of technology in fostering sustainable and productive farming practices cannot be overstated. This research highlights the transformative potential of an intelligent crop and fertilizer recommendation system that is tailored to specific local conditions.

Such a system serves as a guiding tool for farmers, helping them achieve better harvests, improved financial outcomes, and a positive impact on environmental sustainability. At its core, this transformation is driven by the power of data-driven insights and advanced algorithms. By analyzing factors such as soil quality, climatic conditions, and region-specific crop types, we can provide highly personalized recommendations that are not only beneficial but also necessary in the diverse global agricultural context.

Future Scope

While the current system provides reliable crop and fertilizer recommendations, there is ample scope to enhance its capabilities and broaden its impact in the agricultural ecosystem.

In future iterations, we aim to integrate real-time data from IoT-based soil and weather sensors, enabling the system to offer dynamic and context-aware recommendations. This will allow farmers to respond more effectively to sudden weather changes or nutrient fluctuations in the soil.

Another key direction involves expanding the dataset to include more geographical regions, crop varieties, and soil profiles. This would improve the model's generalizability and make it usable across different agro-climatic zones.

We also envision adding market intelligence to the recommendation engine—helping farmers not just choose what can grow well,

but also what can sell well. This will empower them to make economically beneficial decisions, not just agronomically sound ones.

In the long term, the system can evolve into a complete digital farming assistant, with features such as pest and disease prediction, irrigation scheduling, and personalized yield forecasts. These enhancements would help smallholder farmers adopt smart, sustainable, and profitable farming practices with ease.

Ultimately, the goal is to develop a robust, intelligent platform that can scale nationwide, reduce dependency on manual expertise, and bridge the gap between traditional farming methods and modern data-driven agriculture.

VI. Acknowledgment

We would like to express our sincere gratitude and appreciation to our Project Guide, R.Madhavi , and our Project Mentor, Mrs.T.Gayathri , for their valuable guidance, insightful suggestions, and constant encouragement throughout the preparation of this project synopsis. Their support has been instrumental in the successful completion of this task.

Our heartfelt thanks also go to all the faculty members of the Department of CSE (AI & ML), Keshav Memorial Engineering College especially our esteemed Head of Department, Prof. Mr. P.Naresh Kumar, and our respected Principal, Prof. P.V.N Prasad. Their guidance and cooperation were pivotal in completing this project work.

We are truly grateful to everyone who contributed to this project, as their cooperation, inspiration, continuous support, and constructive feedback made it possible to finish this task and synopsis report within the given time frame. Without their efforts, this accomplishment would not have been possible.

References

1. G Sharma, V., Kumar, N. and Rao, L. (2022) 'Big data and ML applications in farming', International Journal of Agricultural Data Science, 16(4), pp. 233– 249.
2. Thompson, P., Roberts, G. and Singh, T. (2021) 'AI- driven decision support systems for farmers', Agricultural Informatics Journal, 10(2), pp. 145–165.
3. Wilson, R., George, M. and Patel, S. (2020) 'Comparing traditional and ML-based crop prediction methods', AI & Agriculture Review, 9(1), pp. 101– 119.
4. Das, P., Mehta, K. and Roy, S. (2023) 'Smart irrigation management using ML models', Journal of Precision Agriculture, 20(3), pp. 289–307.
5. Miller, J., Wang, T. and Chang, H. (2022) 'Automating yield prediction with AI techniques', Computational Agriculture & AI, 19(4), pp. 199–216.
6. Fernando, R., Li, J. and Chen, T. (2021) 'Remote sensing and ML for crop monitoring', IEEE Sensors Journal, 33(5), pp. 90-104.