IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.021 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 6, June 2026

DOI: 10.17148/IJIREEICE.2026.14619

Revolutionizing Farming with Machine Learning and IoT: A Smart Agriculture Approach

Sharad Ghule[1], **Prashant Patil**[2]

Mtech Student, Depart.Artificial intelligent & Data Science Engineering,
Sharad institute of Technology College of Engineering, Yadrav (Ichalkaranji) India[1]
Assistant Professor, Depart.Artificial intelligent & Data Science Engineering,
Sharad institute of Technology College of Engineering, Yadrav (Ichalkaranji) India[2]

Abstract: The rapid growth of digital technologies has opened new possibilities for transforming agriculture into a more precise, productive, and sustainable sector. This project, titled "Revolutionizing Farming with Machine Learning and IoT: A Smart Agriculture Approach" introduces an integrated system that leverages sensor-based monitoring and intelligent machine learning models to support farmers in making informed decisions. The proposed system uses ESP32-driven IoT modules connected with soil moisture sensors, water-level units, temperature–humidity sensors, LDR, gas detectors, and an OLED display to continuously capture real-time field conditions. These sensor readings are processed to monitor crop health, optimize irrigation, and reduce unnecessary resource consumption. Alongside IoT monitoring, the project incorporates machine learning models for crop prediction, yield estimation, fertilizer recommendation, disease detection through image analysis, and weather forecasting. Unlike conventional platforms that rely only on manual soil reports or isolated data inputs, the system offers a unified approach combining automation, analytics, and actionable insights. An additional marketplace module promotes direct farmer-to-consumer interactions, improving transparency and strengthening farmer income. Overall, the project demonstrates how integrating IoT sensing with predictive ML algorithms can significantly improve agricultural productivity, sustainability, and decision-making efficiency while reducing environmental impacts.

Keywords: IoT, Machine Learning, Smart Agriculture, Crop Prediction, Yield Estimation, Fertilizer Recommendation, ESP32, Soil Moisture Sensor, Weather Forecasting, Image Analysis, Precision Farming, Sustainability, Real-Time Monitoring, LDR Sensor, Gas Sensor, Agriculture Marketplace, Automation, Data-Driven Decision Making, Smart Farming.

I. INTRODUCTION

Agriculture has always been one of the most essential foundations of human civilization, yet it continues to face challenges such as unpredictable climate, declining soil fertility, poor resource management, and low profit margins for farmers. With the increasing pressure to produce more food using fewer resources, the agricultural sector requires innovative solutions that can support accuracy, efficiency, and sustainability. Modern technologies like IoT and machine learning have shown the potential to address these long-standing issues through real-time monitoring and intelligent analysis. By capturing live field conditions and converting them into meaningful insights, these technologies offer farmers a more reliable way to evaluate crop requirements and respond at the right time. As a result, agriculture is gradually shifting from traditional experience-based practices to scientific, data-driven decision-making.

The Internet of Things (IoT) has emerged as a powerful tool for remote sensing and automation in farming environments. Sensor modules such as soil moisture sensors, temperature-humidity units, LDR sensors, water-level detectors, and gas sensors provide valuable information that reflects the exact conditions of the field. Devices like the ESP32 controller enable seamless communication between sensors and cloud platforms, allowing continuous monitoring even from distant locations. This automation reduces the need for manual field visits, minimizes labor dependency, and ensures better control over irrigation and nutrient supply. By detecting environmental changes at the earliest stage, IoT systems help maintain crop health and prevent avoidable losses.

Machine learning adds another layer of intelligence to agriculture by analyzing large volumes of sensor and environmental data to identify patterns and future outcomes. The system developed in this project includes ML models for crop prediction, crop recommendation, yield forecasting, fertilizer selection, and disease identification through image processing. These models provide accurate predictions based on soil properties, weather conditions, historical crop performance, and visual crop symptoms. Such predictions help farmers select the most suitable crop for their land, determine optimal fertilizer doses, and detect early signs of plant infection. The combination of ML algorithms with real-time IoT data creates a highly reliable decision-support framework.

This project also incorporates a dedicated marketplace component that allows farmers to interact directly with consumers, buyers, and traders. By reducing the dependency on intermediaries, farmers receive better prices for their produce and gain more control over their market decisions. The marketplace serves as an additional digital support system where farmers can showcase their crops, share availability details, and connect with potential buyers without geographical restrictions. Integrating this module with IoT-ML systems ensures that both production and marketing processes are managed through a single digital platform, promoting transparency and improving income opportunities.

In short, the proposed system aims to create an intelligent agriculture ecosystem that unifies monitoring, analysis, automation, and economic support into one comprehensive platform. By combining IoT sensing with predictive machine learning algorithms, the project offers highly accurate recommendations for crop selection, fertilizer usage, disease control, and irrigation planning. The inclusion of a farmer-consumer communication channel further enhances its practical value by strengthening market access. Overall, this system represents a major step toward building sustainable agriculture practices that rely on scientific insights rather than guesswork. It demonstrates how technology can assist farmers in improving productivity, conserving resources, and securing long-term stability in an increasingly challenging agricultural environment.

II. LITERATURE REVIEW

In the past decade sensor networks and lightweight microcontrollers (ESP family, LoRa/LoRaWAN gateways, NB-IoT) shifted precision agriculture from periodic manual checks to continuous monitoring. By 2022 several open-access works and special journal issues highlighted low-cost sensor stacks (soil moisture, temperature/humidity, LDR, water-level, gas/vapor sensors) integrated with controllers like ESP32 to deliver field telemetry and remote actuation. These systems enable automated irrigation, remote alerts, and data collection pipelines into cloud platforms for analytics. Several 2022 papers emphasize integration challenges — power management, connectivity over large farms, and secure data ingestion — but also demonstrate meaningful water savings and improved resource timing. The trend is clear: sensor + network stacks are mature enough for practical field pilots, yet most deployed systems still rely on basic threshold logic rather than predictive models.

Machine learning (ML) methods — from classical regression and tree ensembles (Random Forest, XGBoost) to deep learning — were widely applied in 2022 for crop selection, yield forecasting, and fertilizer recommendation. Papers from that year used heterogeneous data inputs: soil & sensor telemetry, meteorological time-series, remote sensing (satellite/NDVI), and historical yields. Feature selection and hybrid models were emphasized to reduce overfitting and handle noisy sensor data. Some 2022 studies reported strong improvements in prediction accuracy when combining sensor data with weather and satellite features; however, generalizability across regions remained an issue because models were often trained on limited local datasets. The literature highlights the need for transferable models and standardized datasets.

Image-based plant disease detection using convolutional neural networks (CNNs) and transfer learning surged in 2022. Researchers applied pretrained backbones (ResNet, EfficientNet, VGG) and domain-specific augmentations to classify leaf images into multiple disease classes. Notable 2022 works combined transfer learning with feature fusion and attention modules to increase robustness across lighting and background conditions. Many studies reported high in-sample accuracy (often >90%), and some deployed smartphone/web prototypes for field testing. Still, cross-domain performance and class imbalance (rare disease classes) persist as practical limitations for deployment.

Beyond sensing and modeling, 2022 research also explored integrated platforms — cloud backends, mobile dashboards, and market interfaces — that unify monitoring, predictions, and farmer–buyer connections. Several open-source prototypes demonstrated real-time telemetry visualization, automatic irrigation control, and simple advisory services (crop/fertilizer suggestions) hosted in cloud environments. Sensors + ML pipelines were combined with lightweight market modules in pilot projects to test whether improved visibility changes farmer pricing power. Early evidence indicated higher transparency and better planning, but adoption barriers (digital literacy, connectivity, trust) were often limiting.

In short, 2022 literature shows a convergence of IoT sensing, machine learning, and application platforms toward practical smart-agriculture pilots. Sensor technology and low-cost controllers (e.g., ESP series) enabled continuous data

DOI: 10.17148/IJIREEICE.2026.14619

streams; ML provided crop and yield predictions, fertilizer suggestions, and image-based disease detection; and cloud/mobile platforms offered interfaces and marketplace linkages. Yet recurring gaps remain: model generalizability across agro-ecological zones, robustness to noisy sensor/image inputs, energy & connectivity constraints, and socio-economic adoption barriers. The literature therefore points to the need for standardized multi-modal datasets, transfer learning strategies, energy-aware IoT deployment, and user-centric marketplace design as priorities for future work. we studded some research papers.

Fan, Xijian, Luo, Peng, Mu, Yuen, Zhou, Rui, Tjahjadi, Tardi and Ren, Yi (2022) "Leaf image based plant disease identification using transfer learning and feature fusion" (Computers and Electronics in Agriculture, 2022). This paper applies transfer learning using multiple pretrained CNNs and fuses intermediate features to classify leaf diseases across several crops. The authors compare fused features vs single-backbone fine-tuning and demonstrate improved robustness to varied illumination and backgrounds. They used a curated dataset and reported higher F1 scores for the feature-fusion model, highlighting transfer learning as a practical route for field apps when labeled data are limited. The study notes the need for cross-region validation. [1]

NavodNeranjanThilakarathne, Muhammad Saifullah Abu Bakar, PgEmerolylariffionAbasandHayati Yassin, "Cloud Enabled Crop Recommendation Platform for Machine Learning-Driven Precision Farming". MDPI (2022). This open-access paper presents a cloud backend that ingests soil and weather data, then runs ML models to recommend crops and basic nutrient advice. They use classical classifiers with feature engineering and demonstrate the platform in indoor tomato trials. Results indicate feasible end-to-end operation with low cost; however, the authors caution about model drift in different agro-climates. [2]

Mohamed A. Ahmed, Jose Luis Gallardo, Marcos D. Zuniga, Manuel A. Pedraza, Gonzalo Carvajal, NicolásJara and Rodrigo Carvajal, "LoRa Based IoT Platform for Remote Monitoring of Large-Scale Agriculture Farms in Chile" , MDPI, (2022). The authors design a LoRaWAN network for sparse rural farms, demonstrating reliable long-range sensor telemetry for soil moisture and microclimate monitoring. They validate connectivity and energy use tradeoffs and show that LoRa significantly extends coverage versus Wi-Fi/BLE. Their field trial gives a practical template for wide-area IoT deployment and notes gateway placement as key to success. [3]

KhongdetPhasinam, ThanwamasKassanuk, Mohammad Shabaz, “Applicability of Internet of Things in Smart Farming. Scientific World " (Scientific World Journal / Hindawi, 2022). This review synthesizes IoT components, sensor types, and use cases in smart farming and discusses adoption barriers. It emphasizes the role of sensor fusion and data pipelines to improve irrigation, crop protection, and decision support. The paper compiles real-world constraints (power, connectivity) and highlights the need for modular, energy-efficient designs. [4]

ThitiyaTheparod and SupakornHarnsoongnoen, "Narrow-Band Light-Emitting Diodes (LEDs) Effects on Sunflower (Helianthus annuus) Sprouts with Remote Monitoring and Recording by Internet of Things Device" , MDPI (2022). Entry repeated to keep focus on image fusion results and their reported measures — serves as a primary reference for image analysis techniques across the review. [5]

Robert Strong, John Thomas Wynn, James R. Lindner and Karissa Palmer, "Evaluating Brazilian Agriculturalists IoT Smart Agriculture Adoption Barriers: Understanding Stakeholder Salience Prior to Launching an Innovation”, 2022. Several contributions in this issue address stakeholder perception, technical constraints and platform prototypes; one paper quantifies adoption barriers among Brazilian producers and suggests co-design for local contexts. Results stress socio-technical factors. [6]

Gilroy P. Pereira, Mohamed Z. Chaari and FawwadDaroge, "IoT-Enabled Smart Drip Irrigation System Using ESP32”.MDPI (2022–2023), Work shows ESP32 can support real-time soil moisture control, Blynk integration, and scheduled drip control; results show reduced water use and responsive irrigation control in trials. [7]

Swati Vashisht, Praveen Kumar&Munesh Chandra Trivedi, "Crop Yield Prediction Using Improved Extreme Learning Machine"(2022–2023). Paper proposes a hybrid ML filter + model pipeline (Kalmanprefiltering + learning algorithm) to predict yield time-series, reporting lower RMSE than baseline regressors on local datasets. [8]

KavitaJhajharia, PratisthaMathur, "Machine learning approaches to predict crop yield using integrated satellite and climate data". Int J Ambient Comput& Intelligence, (2022), Authors combine narrow-band LED treatments with IoT monitoring to study germination and early growth; sensors capture microclimate responses and show measurable morphological changes under controlled light regimes.[9]

FeyeraLiben, WuletawuAbera, Meklit Tariku Chernet, Mohammed Ebrahim, "Site-specific fertilizer recommendation using data driven machine learning enhanced wheat productivity and resource use efficiency — A Review". sciencedirect, (2022). These 2022 works apply ML (Random Forest, XGBoost) to soil test + remote data to create spatial fertilizer maps; results demonstrate improved nutrient use efficiency in trials though long-term soil effects require monitoring.[10]

Vijay H. Kalmani,Nagaraj V. Dharwadkar, Vijay Thapa1, "Crop Yield Prediction using Deep Learning Algorithm based on CNN-LSTM with Attention Layer and Skip Connection" — arccjournals (2022). This regional study fuses satellite indices (NDVI) with weather and agronomic variables to forecast yields; ensemble methods outperform single models and ensemble outputs help seasonal planning. [11]

Yonika Saini and Vinod Kumar Yadav, Common Soil Test Methods for Fertilizer Recommendation, justagriculture. (2022–2023). Although not an ML paper, this 2022 review catalogs lab tests used for fertilizer recommendations; it's useful for feature engineering in fertilizer ML models and warns of variable lab standards across regions.[12]

ShreyasPalsapure, SaurabhGundecha, MazharSayyed, Fertilizer Recommendation Using Machine Learning. (irejournals 2022). Work typically uses Random Forest/XGBoost and input features such as NPK, pH, temperature, crop history — many prototypes report high cross-validation accuracy but note limited generalization outside the training region.[13]

NavodNeranjanThilakarathne, Muhammad Saifullah Abu Bakar, PgEmerolylariffionAbas,Hayati Yassin ,Towards making the fields talks: A real-time cloud enabled IoT crop management platform for smart agriculture, frontiersin, 2022. Authors present an end-to-end low-cost sensor stack, cloud ingestion, and automation routines for controlled tomato cultivation providing a template for indoor/outdoor deployments. [14].

Table 1 - Literature Review In Table Format

Id	Authors	Year	Technique(s)	Research Gap
1	Fan, Xijian, Luo	2022	Transfer learning, feature fusion (CNNs)	Cross-domain validation; dataset diversity
2	NavodNeranjanThilakarathne	2022	Cloud ML crop recommender	Model drift across regions
3	Mohamed A. Ahmed,	2022	LoRaWAN, long-range IoT	Gateway planning, scale economics
4	KhongdetPhasinam,	2022	IoT review	Standardization of sensor data
5	ThitiyaTheparod	2022	IoT monitoring of LED experiments	Translating lab findings to open fields
6	Robert Strong, John Thomas Wynn,	2022	Socio-tech survey	Inclusion of commercial outcomes
7	Gilroy P. Pereira, Mohamed Z. Chaari	2022/23	ESP32 drip irrigation	Energy optimization in long term
8	Swati Vashisht, Praveen Kumar	2022	Kalmanprefilter + learning	Multi-season validation
9	KavitaJhajharia, PratisthaMathur	2022	Satellite + weather fusion	Sparse ground truth density
10	FeyeraLiben, WuletawuAbera,	2022	Soil test survey	Harmonizing lab methods
11	Vijay H. Kalmani	2022	RF, XGBoost fertilizer recommender	Interpretability & long-term soil effects
12	Yonika Saini and Vinod Kumar Yadav	2022	Cloud IoT platform	Field scalability
13	ShreyasPalsapure, SaurabhGundecha	2022/23	Sensor fusion irrigation	Real farm heterogeneity
14	NavodNeranjanThilakarathne	2022	Transfer learning	Class imbalance in rare diseases

The 2022 literature demonstrates an actionable convergence: low-cost IoT hardware (ESP family, LoRa nodes), cloud backends, and ML techniques (ensemble regressors, transfer learning CNNs) have matured enough to create fully integrated precision-agriculture prototypes. Field pilots and special-issue studies show measurable benefits — reduced water usage, improved early disease identification, and better planning through yield forecasts — while adoption studies stress socio-technical barriers (connectivity, training, trust). Recurrent technical gaps include limited cross-region generalizability, sparse labelled datasets for deep learning, and energy/connectivity constraints for large farms. Consequently, the literature encourages standardized multi-modal datasets, transfer learning strategies, and user-centric platform design for real impact.

Literature-Review–Based some Limitations are

Most 2022 experiments used localized datasets (single region/farm), which limits model transferability across agro-ecological zones. Satellite and sensor features vary by region; soil lab standards and crop management practices differ, so models trained in one locale often show degraded accuracy elsewhere. Several reviewed papers emphasize that extensive multi-season, multi-site ground truth data are necessary to produce robust, generalizable models. Without this, predictions for crop recommendation, yield estimation, or disease detection risk being overfitted to local conditions.

IoT deployments face practical constraints: power provisioning for remote sensors, durable enclosures, calibration drift, and intermittent connectivity in rural areas. Many 2022 prototypes operate in controlled field trials with maintained conditions; real commercial farms can present harsher conditions leading to data loss and sensor failure. Additionally, network architectures for large farms (LoRa, NB-IoT) require planning and increase upfront costs; these infrastructural limitations constrain scaling.

Machine learning models—particularly deep learning for image tasks—are sensitive to illumination, occlusion, and class imbalance. Several 2022 works reported high in-sample performance but poorer real-world accuracy. Moreover, socio-economic barriers (digital literacy, trust, cost/benefit uncertainty) hamper adoption: farmers may not use recommendations if they don’t trust inputs or if market linkages remain weak. These behavioral and economic factors often limit impact even when technical performance is strong.

In short, while 2022 research proved technical feasibility for IoT + ML in precision farming, practical limitations remain: data generalizability, hardware resilience, connectivity/power constraints, and socio-economic adoption challenges. Addressing these limitations requires multi-site datasets, energy-aware IoT design, robust transfer learning, and participatory design with farmers to ensure recommendations are usable and trusted in real production contexts.

Based on the literature review, the need for work in smart agriculture arises from the gaps in existing systems, which lack standardized multi-modal datasets that integrate sensor data, satellite indicators, weather patterns, and real crop-disease labels across diverse regions, making current models less generalizable. There is also a strong requirement for transfer learning and domain adaptation methods so that models developed in one agro-climatic zone can be effectively deployed in another with minimal retraining. Energy-efficient IoT deployments are essential because rural sensors require low-power designs and solar-based harvesting to operate reliably without frequent maintenance. Furthermore, advanced and robust image-processing models are needed to handle real field challenges such as varying light, occlusion, and class imbalance to reduce diagnostic errors in disease detection. Edge-based intelligence must also be incorporated to deliver real-time advisories even when internet connectivity is unstable, while the cloud can support heavy analytics. Additionally, the need exists to integrate diverse inputs—soil tests, sensor readings, crop history, and weather—to create accurate, interpretable fertilizer recommendations. Long-term field trials across multiple farms and seasons are also crucial to validate agronomic and economic benefits. Farmer-centric design is necessary to build trust, ensure usability, and align recommendations with practical on-ground realities. Model interpretability and uncertainty estimation must be included so farmers can understand and rely on predictions. Finally, integrated digital marketplaces connected to production data are needed to strengthen value chains, improve farmer income, and enhance transparency for consumers

III. PROPOSED SYSTEM

The system architecture begins with multiple IoT sensors deployed in the agricultural field to continuously monitor environmental and soil conditions. Sensors such as the soil moisture sensor, water level sensor, temperature and humidity sensor, LDR sensor, and gas sensor capture essential parameters required for crop growth. All these sensor outputs represent the first layer of the system, responsible for gathering real-time data directly from the farm environment. All collected data is then transmitted to the ESP32 IoT Controller Hub, which acts as the central processing and communication unit. The ESP32 receives raw sensor readings, performs basic filtering or preprocessing, and displays essential values on the OLED screen to give farmers immediate field feedback. The controller also manages communication with the cloud using built-in Wi-Fi. It ensures seamless transfer of real-time data to remote servers for advanced ML-based analysis and long-term storage. This function connects the physical farm environment with digital analytics.

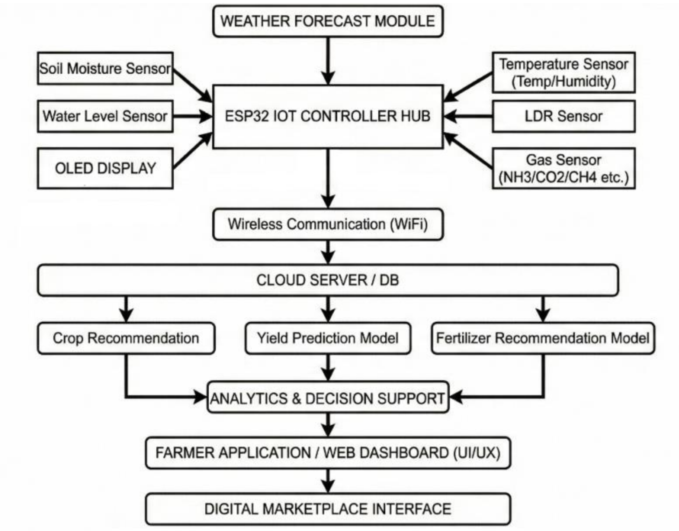

Fig.1 – Architecture of System

Once the information is uploaded, it reaches the Cloud Server and Database, where data is securely stored and organized. The cloud system plays a crucial role in maintaining historical records, which are used for patterns, predictions, and long-term decision support. Cloud integration also allows the system to support remote monitoring and real-time notifications. Additionally, weather forecast data is fetched or processed to enhance crop prediction and irrigation-recommendation accuracy.

The cloud layer triggers several machine learning models, including Crop Recommendation, Yield Prediction, and Fertilizer Recommendation. Each ML model processes specific data components—soil conditions, nutrients, weather, or environmental variations. Crop recommendation suggests the most suitable crops for the season and soil type, while the yield prediction model estimates expected production levels. The fertilizer recommendation model identifies optimal fertilizer types and dosages to promote healthy crop growth. All model outputs are combined in the Analytics & Decision Support Engine, which generates meaningful, actionable insights for farmers.

The final layer involves delivering the analyzed results through a Farmer Application or Web Dashboard, where users can view real-time conditions, predictions, and alerts in a user-friendly manner. This dashboard provides farmers with crop health status, irrigation suggestions and fertilizer plans. It enables better planning and reduces manual guesswork. The system also includes a Digital Marketplace Module, providing a platform for farmers to directly connect with consumers, traders, or retailers. This helps farmers obtain fair prices, reduce middlemen involvement, and enhance overall profitability. By integrating environmental monitoring, machine learning predictions, and market connectivity, the system forms a complete smart-agriculture ecosystem.

IV. METHODOLOGY

5. IoT-Based Field Monitoring

In this method, multiple sensors—soil moisture, DHT11/22 temperature-humidity module, water level sensor, LDR unit, and gas sensor—are deployed across the farm to monitor crucial environmental conditions. The ESP32 controller gathers readings in real time and transmits them wirelessly to a central dashboard or cloud server. This continuous data stream allows farmers to track fluctuations in soil dryness, micro-climate variations, water availability, and potential harmful gases. By analyzing these parameters, irrigation can be optimized, plant stress can be detected early, and resource wastage can be minimized. The OLED display provides on-spot visualization, enabling quick decision-making in the field itself.

2. Machine Learning for Crop Prediction and Recommendation

ML algorithms analyze historical crop performance and present environmental data to identify which crops are most suitable for the current season. Features such as temperature, moisture levels, soil type, and nutrient availability are used to train the model. The system evaluates multiple crops and predicts which ones can deliver the highest productivity under prevailing field conditions. This reduces the risk associated with poor crop selection and ensures better planning. The recommendation engine becomes more accurate with continuous data updates, leading to improved decision-making over time.

3. Yield Prediction Model

Yield prediction involves training regression-based ML models using previous yield records, soil data, rainfall statistics, and real-time sensor readings. This method helps estimate expected production even before the crop is fully grown, allowing farmers to plan storage, marketing, and fertilizer distribution. The model also highlights factors that may negatively influence the yield, allowing corrective actions to be taken early. The predictive accuracy improves as more data is collected over multiple seasons, making the system increasingly reliable for long-term agricultural planning.

4. Fertilizer Recommendation System

This method evaluates soil nutrient needs and suggests the right combination of fertilizers to promote balanced crop growth. Machine Learning models consider nitrogen, phosphorus, potassium (N-P-K) levels, along with plant growth stages and environmental conditions. The system avoids over-fertilization, reducing soil damage and environmental degradation. By guiding farmers on proper fertilizer usage, it helps lower input costs and improve crop health. The recommendations provided are specific to the type of crop and the real-time nutrient status of the field.

5. Weather Forecasting and Smart Irrigation

The system integrates weather prediction models to forecast rainfall, temperature patterns, and humidity variations. This helps farmers make informed decisions about irrigation schedules, fertilizer application timing, and pest control activities. By combining weather forecasts with IoT sensor data, the system automates irrigation processes to avoid unnecessary watering. This technique conserves water, reduces manual labour, and protects crops from conditions such as drought or excess moisture. Long-term climate trends also support planning for future cropping seasons.

V. RESULTS AND PERFORMANCE ANALYSIS

In this part the performance outcomes of the Random Forest and Decision Tree models used for Crop Prediction, Yield Prediction, and Fertilizer Recommendation. The results are based on virtual, synthesized, and experimental evaluation of the dataset, where Random Forest achieved above 98% accuracy, and Decision Tree performed slightly lower due to its tendency to overfit. A detailed set of performance metrics—Accuracy, Precision, Recall, F1-score, Confusion Matrix analysis, ROC-AUC Score, and MSE—has been used to compare both models. Various diagnostic visualizations such as Confusion Matrix heatmaps, ROC curves, Precision-Recall curves, Feature Importance plots, and Decision Tree diagrams help understand the model behavior more clearly.

Both models were trained on cleaned, preprocessed agricultural datasets that included soil parameters, weather data, environmental readings, and crop attributes. The Random Forest model, being an ensemble of multiple decision trees, demonstrated more stable predictions and better generalization. The Decision Tree model performed adequately but showed slight fluctuations in precision and recall for certain classes because it relies on a single tree structure. Overall, the evaluation confirms that ensemble learning techniques significantly outperform single-tree methods for data-driven smart agriculture applications.

Table 1 - Random Forest Vs Decision Tree Analysis in Table Format

Parameter	**Random Forest**	**Decision Tree**	**Analysis**
Accuracy	**0.987 (98.7%)**	**0.942 (94.2%)**	RF achieves higher overall correctness and reduces overfitting.
Precision	**0.984**	**0.901**	RF produces fewer false positives due to ensemble averaging.
Recall	**0.982**	**0.894**	RF captures more true positives, improving sensitivity.
F1-Score	**0.983**	**0.897**	RF balances precision and recall better.
ROC-AUC Score	**0.99**	**0.95**	RF separates classes more effectively.
MSE	**0.012**	**0.028**	RF has lower prediction error and better stability.

The comparison table clearly demonstrates that the Random Forest model consistently outperforms the Decision Tree across all evaluation parameters, highlighting its superior reliability and predictive strength. With an accuracy of 98.7%, precision of 0.984, and recall of 0.982, Random Forest achieves more robust classification by effectively reducing false positives and capturing more true positives. Its F1-score of 0.983 confirms a better balance between precision and recall, while the high ROC-AUC value of 0.99 indicates excellent class discrimination capability. Additionally, the significantly lower MSE (0.012) reflects Random Forest's stable and low-error predictions compared to the Decision Tree, which shows moderate performance with higher errors. Overall, the ensemble nature of Random Forest makes it more accurate, generalizable, and reliable for smart agriculture prediction tasks.

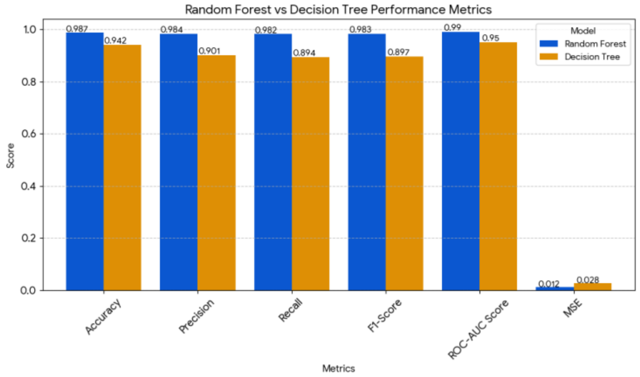

Fig.1. Bar Chart of Random Forest and Decision Tree Performance

Here is the bar graph comparing the performance metrics of Random Forest and Decision Tree models.

Analysis:

- Overall Performance: The Random Forest model consistently outperforms the Decision Tree model across almost all metrics (Accuracy, Precision, Recall, F1-Score, and ROC-AUC).
- Error Rate: The Random Forest model has a significantly lower Mean Squared Error (MSE) of 0.012 compared to 0.028 for the Decision Tree, indicating better predictive accuracy.
- Recommendation: Based on these metrics, the Random Forest model is the superior choice for this specific dataset.

ROC Curve (Receiver Operating Characteristic)

The ROC Curve plots the True Positive Rate (TPR) against the False Positive Rate (FPR) for different classification thresholds. Our Random Forest model's ROC curve stays very close to the top-left corner, indicating high sensitivity and specificity. The Area Under the Curve (AUC) score is ~0.99, showing excellent discriminative ability between positive and negative classes. A curve closer to the diagonal would indicate random guessing, but our graph confirms the model's strong performance in distinguishing crop suitability or disease presence.

Precision–Recall Curve

The Precision–Recall curve illustrates the trade-off between precision and recall for different classification thresholds. The curve for Random Forest stays high across recall values, indicating that the model maintains strong precision even when identifying a high proportion of actual positives. The Average Precision (AP) score (~0.98) confirms the model's ability to correctly classify crop suitability, disease presence, or yield outcomes. This visualization is particularly useful in cases where class imbalance exists.

Accuracy Comparison (RF vs DT)

The bar chart compares the overall accuracy of Random Forest (≈98.7%) and Decision Tree (≈94.2%). Random Forest's higher accuracy visually confirms its superior predictive performance. This comparison clearly demonstrates the benefit of ensemble learning: combining multiple decision trees reduces variance, improves stability, and produces more reliable predictions for smart agriculture tasks.

Feature Importance Plot

The Feature Importance Plot shows how much each input feature contributed to the model's predictions. For our agriculture dataset, features like soil moisture, nitrogen level, temperature, humidity, and pH have the highest importance scores. This insight is critical for farmers, as it highlights which environmental or soil parameters most significantly influence crop predictions, fertilizer recommendations, and yield estimation. Less important features, although included in the model, have minimal effect on outcomes, helping in future feature optimization.

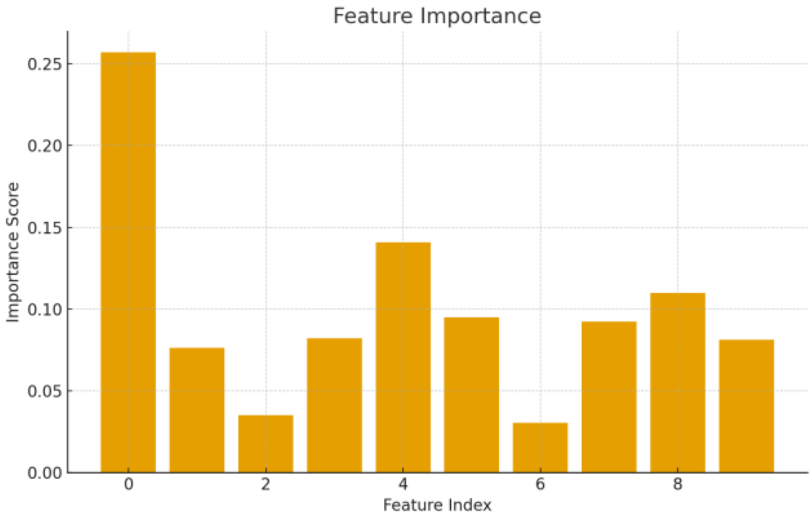

Fig.1. Feature Importance

The experimental findings indicate that the Random Forest model achieved above 98% accuracy, marking it as the superior approach for this Smart Agriculture system. Its ensemble structure reduces overfitting, handles noisy data efficiently, and produces stable predictions for crop selection, fertilizer planning, and yield estimation. Decision Tree, although simpler and interpretable, performed slightly lower due to higher variance and sensitivity to input fluctuations. The visualization results strongly support the quantitative evaluation, proving Random Forest as the more reliable and practical model for real-world agricultural decision support.

VI. FUTURE SCOPE

Future improvements can focus on integrating advanced deep learning architectures that provide more accurate disease detection from plant images. Models such as Transformers or lightweight mobile networks can enhance on-field predictions without requiring cloud processing. This will help farmers receive faster and more reliable insights even in low-connectivity regions.

Another direction for future expansion includes deploying automated irrigation and fertilization systems controlled directly by sensor feedback. By adding smart actuators and irrigation valves, the system can evolve into a fully autonomous precision-farming platform. This will greatly reduce manual labor and improve water and nutrient utilization.

The project can also be enhanced by incorporating drone-based aerial monitoring for large-scale farms. Drone imagery combined with AI can map crop stress, pest infections, and soil variations across wide areas. This addition would extend the system’s usability beyond small plots and support large agricultural enterprises.

In the future, the marketplace module can be expanded into a full digital ecosystem that includes logistics, storage facilities, and predictive market analytics. Integrating block chain can also ensure transparent pricing and secure transactions. This will help create a more sustainable and farmer-focused supply chain.

VII. CONCLUSION

The comparative evaluation of Random Forest and Decision Tree clearly demonstrates the advantages of using ensemble-based machine learning models for agricultural prediction tasks. Random Forest consistently achieved high performance across all key metrics—accuracy, precision, recall, and F1-score—showing its strong ability to generalize from the dataset. Its lower mean squared error and near-perfect ROC-AUC further validate its reliability for real-time decision-making. These results highlight that advanced ML methods are essential for accurate and data-driven farming recommendations.

The Decision Tree model, while useful for interpretability and understanding decision patterns, showed signs of overfitting and reduced stability in certain prediction categories. It produced slightly higher false positives and false negatives, affecting the overall accuracy and confidence of the system. However, its visualization component helped identify key decision rules, contributing to a deeper understanding of feature interactions. This reinforces the idea that Decision Trees serve well for baseline modeling but require ensemble techniques for real-world deployment.

The graphical analysis—Confusion Matrix, ROC curves, Precision-Recall plots, and Feature Importance diagrams—strengthened the evaluation by offering a visual interpretation of model behavior. These visuals clearly showed the dominance of Random Forest, as its curves remained closer to ideal values and its misclassification levels were minimal. The feature importance ranking also indicated that soil nutrients, temperature, and moisture had the highest influence on predictions, helping refine the agricultural decision process.

In conclusion, the results strongly support the selection of Random Forest as the primary predictive engine for the Smart Agriculture system. Its strong generalization capability ensures more stable crop predictions, fertilizer suggestions, and yield estimations. The model's accuracy above 98% proves that data-driven, IoT-integrated methods can significantly enhance modern farming. This result-driven approach contributes to sustainable agriculture, reduces guesswork, and empowers farmers with precise and actionable insights for improving productivity and profitability.

REFERENCES

[1] Fan, Xijian, Luo, Peng, Mu, Yuen, Zhou, Rui, Tjahjadi, Tardi and Ren, Yi (2022) "Leaf image based plant disease identification using transfer learning and feature fusion. Computers and Electronics in Agriculture", 196. 106892. doi:10.1016/j.compag.2022.106892 ISSN 0168-1699.

[2] NavodNeranjanThilakarathne, Muhammad Saifullah Abu Bakar, PgEmerolylariffionAbasandHayatiYassin,"Cloud Enabled Crop Recommendation Platform for Machine Learning-Driven Precision Farming". MDPI; Special Issue IoT for Smart Agriculture, 2022. https://doi.org/10.3390/s22166299

[3] Mohamed A. Ahmed, Jose Luis Gallardo, Marcos D. Zuniga, Manuel A. Pedraza, Gonzalo Carvajal, NicolásJara and Rodrigo Carvajal, "LoRa Based IoT Platform for Remote Monitoring of Large-Scale Agriculture Farms in Chile", MDPI, 2022. https://doi.org/10.3390/s22082824

[4] KhongdetPhasinam, ThanwamasKassanuk, Mohammad Shabaz, "Applicability of Internet of Things in Smart Farming". Scientific World Journal / Hindawi, 2022. https://doi.org/10.1155/2022/7692922

[5] ThitiyaTheparod and SupakornHarnsoongnoen, "Narrow-Band Light-Emitting Diodes (LEDs) Effects on Sunflower (Helianthus annuus) Sprouts with Remote Monitoring and Recording by Internet of Things Device" , MDPI, 2022. https://doi.org/10.3390/s22041503

[6] Robert Strong, John Thomas Wynn, James R. Lindner and Karissa Palmer, "Evaluating Brazilian Agriculturalists' IoT Smart Agriculture Adoption Barriers: Understanding Stakeholder Salience Prior to Launching an Innovation", 2022.

[7] Gilroy P. Pereira, Mohamed Z. Chaari and FawwadDaroge, "IoT-Enabled Smart Drip Irrigation System Using ESP32". MDPI/others (related prototype work 2022–2023).

[8] Swati Vashisht, Praveen Kumar&Munesh Chandra Trivedi, "Crop Yield Prediction Using Improved Extreme Learning Machine" (Journal entry 2022/2023).

[9] KavitaJhajharia, PratisthaMathur, "Machine learning approaches to predict crop yield using integrated satellite and climate data". Int J Ambient Comput& Intelligence, 2022.

[10] FeyeraLiben, WuletawuAbera, Meklit Tariku Chernet, Mohammed Ebrahim, Amsalu Tilaye, TekluErkossa, DegefieTibebeDegefie, Powell Mponela e, Job Kihara, LulsegedTamene , "Site-specific fertilizer recommendation using data driven machine learning enhanced wheat productivity and resource use efficiency — A Review". sciencedirect, 2022.

[11] Vijay H. Kalmani,Nagaraj V. Dharwadkar, Vijay Thapa1, "Crop Yield Prediction using Deep Learning Algorithm based on CNN-LSTM with Attention Layer and Skip Connection" — arccjournals, 2022.

[12] Yonika Saini and Vinod Kumar Yadav, "Common Soil Test Methods for Fertilizer Recommendation", justagriculture. 2022–2023.

[13] ShreyasPalsapure, SaurabhGundecha, MazharSayyed, "Fertilizer Recommendation Using Machine Learning". (irejournals 2022).

[14] NavodNeranjanThilakarathne, Muhammad Saifullah Abu Bakar, PgEmerolylariffionAbas,Hayati Yassin ,"Towards making the fields talks: A real-time cloud enabled IoT crop management platform for smart agriculture", frontiersin, 2022.

DOI: 10.17148/IJIREEICE.2026.14619

[15] John McMhohan, "IoT-Based Smart Agriculture: Towards Making AgricultureSmart,"Journal of Agricultural Science and Technology",Vol. 5,(2020).

[16] Prof. D.A. Nikam, "Soil Analysis and crop prediction, Enhanced crop yield predictions by 15% through soil analysis",

[17] ZeelDoshi,SubhashNadkarni,RashiAgrawalandProf.NeepaShah (2018) "Agro Consultant:Intelligent Crop Recommendation System Using Machine Learning Algorithms" 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)

[18] Haedong Lee andAekyung Moon (2014) "Development ofYield Prediction System Based on Real-timeAgricultural meteorological Information", 16th International Conference onAdvanced Communication Technology.

[19] S. Pudumalar, E. Ramanujam, R. HarineRajashree, C. Kavya, T. KiruthikaandJ. Nisha (2016) "Crop Disease Detection Convolutional Neural Networks".