International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 06 June 2026
Page No: 4302 - 4312
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0560

Data-Driven Farming: A Modern Approach to Crop Management Using IoT and ML Technologies

Sharad Ghule[1],Prashant Patil[2]
[1]Mtech Student, Department of Artificial Intelligence & Data Science Engineering,Sharad Institute of technology College of Engineering Yadrav, Kolhapur, India
[2]Assistant Professor, Department of Artificial Intelligence & Data Science Engineering,Sharad Institute of technology College of Engineering Yadrav, Kolhapur, India
Emails: *sharadghule25@gmail.com[1], prashantpatil21@gmail.com[2]*

Abstract
The ever-growing evolution of digital technologies has brought forth new prospects to add value to the agricultural activities, making them more efficient, accurate and environmentally friendly. The project, entitled Revolutionizing Farming with Machine Learning and IoT: A Smart Agriculture Approach, introduces an integrated framework that integrates real-time sensing with intelligent data analysis to help farmers make better decisions. The system employs an IoT module based on ESP32 connected to other sensors, including soil moisture, water level, temperature and humidity, LDR, and gas sensors, and an OLED display to monitor locally. The field data is constantly collected by these devices and used to assess the condition of crops, efficient irrigation, and the avoidance of needless use of resources. Along with the use of IoT-based monitoring, machine learning techniques are introduced to predict crops, estimate yields, recommend fertilizers, forecast weather, and detect plant diseases using image processing. Unlike traditional approaches that depend mainly on manual reports or limited datasets, the proposed solution integrates multiple functions into a single intelligent system. It also includes a digital marketplace that enables direct communication between farmers and consumers, improving transparency and income opportunities. Overall, the system highlights how the combination of IoT and machine learning can enhance productivity, support sustainable farming, and improve decision-making while minimizing environmental impact.
Keywords: *IoT, Machine Learning, Smart Agriculture, Crop Prediction, Yield Estimation, Fertilizer Recommendation, ESP32, Soil Monitoring, Weather Forecasting, Image Processing, Precision Farming, Sustainability, Real-Time Data, LDR Sensor, Gas Detection, Digital Marketplace, Automation, Data Analytics, Smart Irrigation, Agricultural Intelligence.*

1. Introduction
Change of climatic conditions and soil erosion, ineffective use of resources, and poor financial performance of farmers are some of the persistent challenges that agriculture has been facing. With the world constantly needing more food, there is an ever-increasing pressure on the agricultural industry to ensure more food on the global table and at the same time to ensure that the natural resources are not depleted. Conventional farming techniques that largely depend on experience as well as manual observation are normally not adequate to meet the contemporary farming needs. In this respect, the use of advanced technologies like the Internet of Things (IoT) and Machine Learning (ML) is rather promising as it makes it possible to monitor the situation and make a decision based on data. The technologies enable farmers to see the real-time conditions of their field and react quickly, resulting in more efficient and reliable farming. The implementation of IoT in agriculture has greatly enhanced the manner in which the farms are monitored and controlled. To measure detailed

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 06 June 2026
Page No: 4302 - 4312
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0560

environment data, various sensors are employed which include soil moisture sensors, temperature and humidity modules, LDR sensors, water-level indicators, and gas sensors. The ESP32 microcontroller will serve as a central node that will receive sensor data and send them to cloud computing systems to be analyzed further. This automated monitoring system minimizes the necessity to conduct manual inspection and allows farmers to have a better control over the irrigation and crop health. By early detecting of the environmental changes corrective actions can be taken in time before any damage can be caused or even the overall crop yield can be improved. Machine Learning also improves the system as it offers intelligent analysis of the information gathered. Using information like soil composition, climate conditions, past crop data, and other data, the ML models can predict the most suitable crops, estimate the yield, and recommend the most suitable fertilizers. Image processing methods are also employed to diagnose plant diseases early enough and the farmers can take preventive actions before the disease spreads. Such predictive abilities aid in the optimization of farming choices, minimization of risks, and enhanced productivity. The combination of ML and IoT data results in a robust system that can provide accurate and actionable insights. Besides technical advancement, the project presents a digital market, which links farmers directly to buyers and consumers. This aspect lessens the use of mediators and enables farmers to get increased worth on their products. In this platform, farmers are able to provide information on availability of crops, their prices and quality and facilitate transparent transactions. The marketplace also increases the availability to a broader network of customers enhancing the opportunity of sales and alleviating losses after harvesting. Integration of production and marketing into one system enhances the general agricultural ecosystem. In short, the targeted solution will construct an intelligent and integrated agricultural system that integrates sensing, analysis, automation, and economic assistance. Using both IoT and machine learning to make real-time monitoring and predictive insights, the system provides correct recommendations on crop selection, irrigation, fertilizer application, and disease control. The integration of an original means of communication between farmers and consumers is an additional increase to its practical influence. This technique is a major step towards the modernization of sustainable farming through enhancing productivity, resource conservation, and the long-term development of the farming communities.

2. Literature Review

In the past 10 years, sensor technologies and compact microcontrollers like ESP series, and communication systems such as LoRa/LoRaWAN, and NB-IoT have been developed with considerable impact on agricultural monitoring. The old farming methods which used manual inspection periodically are being slowly overtaken by real-time, continuous observation systems. Around 2022, various research contributions and journal articles highlighted the usefulness of low-cost sensor networks that included soil moisture, temperature, humidity, light intensity, water level, and gas sensors in combination with controllers such as ESP32. These systems allow round-the-clock monitoring of the environment, automated watering, as well as remote monitoring via cloud computing services. As much as these developments show enhanced efficiency in water management and timely decision-making, researchers also found the challenges associated with energy consumption, connection between large agricultural lands, and secure data handling. Although technology has become more mature, there are still numerous deployed systems that rely on basic rule-based logic, rather than predictive analytics. In 2022, machine learning methods have been of particular interest in improving the process of decision-making in agriculture. Different models were used to address different tasks like crop recommendation, yield prediction and fertilizer planning using various models including regression techniques, decision trees, Random Forest, and gradient boosting models. These papers have used a wide range of data, such as sensor measurements, weather, satellite, and historical agricultural data. Researchers highlighted the significance of feature selection and hybrid modeling strategies to enhance accuracy and eliminate the noise effects. Despite the improved predictive performance when the multiple data sources are used, another frequent limitation was that the results could not be generalized across different regions with the use of localized data. This underscored the increasing demands of standardized data and flexible modeling methods. Image-based analysis with deep learning is rapidly developing in the field of plant disease identification in 2022. Convolutional Neural Networks (CNNs) with the use of transfer learning methods became the most common to classify plant diseases based on leaf images. Fine-tuning of pretrained architectures (ResNet, VGG, and EfficientNet) to agricultural tasks was done. Others also included feature fusion and attention to enhance better performance of the model under different lighting and environmental conditions. Although these methods have shown high precision in controlled data, other issues that include uneven distribution of data and low performance in real life scenarios are also a big challenge towards practical implementation. In addition to individual technologies, in 2022, studies in the field of smart agriculture were also directed at the creation of integrated platforms of smart agriculture. These systems integrated IoT-based sensing, cloud computing, mobile interfaces, and straightforward market connectivity solutions into cohesive systems. A number of prototypes were used to display real-time monitoring dashboard, automated irrigation control and farmer advisory services. Other studies even engaged in trying to connect the information on production and the digital marketplace to enhance transparency and cost. Despite some early positive outcomes, including better planning and use of resources, adoption was low because of the problems, such as the lack of digital literacy, unreliable internet connection, and unwillingness of farmers to trust automated systems. In general, the study published in 2022 is a clear indication of the convergence of IoT, machine learning, and application platforms, towards practical smart farming solutions. The combination of continuous data collection using sensors, predictive insights generated by machine learning models and user-friendly interfaces has served to form a solid base of intelligent agriculture. Nonetheless, there are still a number of recurrent challenges, such as the lack of adaptability of the models to different regions, sensitivity to noisy data, power, and connectivity constraints, and socio-economic barriers to adoption. We reviewed some research papers to shed light on these gaps: the need to design standardized datasets, the need to enhance transfer learning methods, the need to design energy efficient IoT systems, and the need to focus on user-centered solutions to guarantee real-world impact. In their article on the problem of identifying diseases based on leaf images, Fan et al. (2022) suggested a paradigm of transfer learning with the addition of feature fusion methods to enhance the quality of classification. The model showed to be more robust in changing environmental conditions like lighting and background noise through incorporating features of several pretrained convolutional networks. The research revealed that intermediate feature combinations were more effective than single-model solutions and they have better F1-scores. Nonetheless, the authors asserted that validation in a wide range of geographical settings was necessary so as to ascertain the feasibility in the real world. [1] Thilakarathne et al. (2022) proposed a cloud-based crop recommendation system that takes soil and weather information and uses them to come up with intelligent recommendations. With the help of machine learning algorithms and feature engineering tools, the system proved to perform effectively in controlled environments in agriculture. The platform was piloted on tomato crops and it was promising on the delivery of recommendations. Nevertheless, the authors also mentioned that the model performance can be different in various climatic conditions, and it is necessary to focus on adaptive learning methods. [2] Ahmed et al. (2022) designed a LoRa-based IoT system to monitor large agricultural fields in Chile. Their system proved to be good long-range communication to transmit soil and environmental measurements. The research revealed the benefits of LoRa technology compared to traditional wireless applications, in terms of coverage and energy consumption. Field experiments established that connectivity and the ability to transmit data in a stable manner was established though the optimal placement of gateways was found to be a critical factor in system performance. [3] Phasinam et al. (2022) developed an extensive overview of the IoT application in the agriculture sector with an emphasis on sensor technologies, communication networks, and real use examples. The paper has explained the potential of sensor integration to enhance irrigation control, crop protection, and decision-making. It also identified real world issues like power restrictions and connectivity problems, which implied that modular system design should be energy efficient and have restricted power constraints. [4] Theparod and Harnsoongnoen (2022) examined the impact of controlled lighting on the growth of sunflower using IoT-based monitoring systems. In their research, narrow-band LED lighting and sensor-based tracking were used to examine the patterns of plant growth. The findings revealed that it was possible to measure significant gains in plant morphology in controlled conditions when combining the system of IoT monitoring with environmental control systems. [5] Strong et al. (2022) have studied the issues related to the implementation of smart technologies in agriculture among farmers in Brazil. Their research was based on the perspectives of stakeholders and they have found factors like cost, technical complexity and lack of awareness as some of the key obstacles. The results highlighted the significance of user-centered design and the training programs to enhance the rate of adoption. [6] The system described by Pereira et al. (20223 -2023) is a smart drip irrigation system based on IoT and using ESP32 microcontrollers. Their system measured the level of soil moisture and automatically regulated irrigation using a mobile interface. The outcomes of the experiments indicated that automated irrigation systems were practical in the use of water and its response to crops. [7] Vashisht et al. (2022–2023) suggested a model of crop yield prediction with an enhanced learning algorithm and using filtering methods. Their method minimized errors in predictions as well as enhanced accuracy in comparison to the traditional forms of regression. The experiment revealed the usefulness of hybrid ML methods in the processing of agricultural time-series data. [8] Jhajjaria and Mathur (2022) concentrated on combining satellite images and weather information to predict crop yields. Their solution used a combination of multiple data sources to enhance the accuracy of predictions, showing the significance of multi-modal data in agricultural analytics. [9] Liben et al. (2022) summarized machine learning methods to achieve fertilizer recommendation and emphasized the role of data-driven techniques in enhancing nutrient management. They found that the ML-based recommendations could be used to improve productivity and decrease resource wastage. [10] The model presented by Kalmani et al. (2022) is a deep learning model that integrates CNN and LSTM architectures to predict crop yields. Their hybrid model not only obtained spatial but also temporal patterns, which led to a better prediction. Kalmani et al. (2022) proposed a deep learning model combining CNN and LSTM architectures for crop yield prediction. Their hybrid model captured both spatial and temporal patterns, resulting in improved prediction performance. [11] Saini and Yadav (2022–2023) discussed traditional soil testing methods and their role in fertilizer recommendations. Their work provided valuable insights into feature selection for ML models while highlighting inconsistencies in testing standards. [12] Palsapure et al. (2022) developed a machine learning-based fertilizer recommendation system using parameters such as soil nutrients and environmental conditions. Their model achieved high accuracy but showed limitations when applied to different regions. [13],As shown in Table 1 Literature Review in Table Format Majority of the studies carried out in 2022 utilized datasets gathered in particular regions or in individual farms and hence restricts the universality of developed models in a wide range of agricultural ecosystems. Varying soil properties, climatic factors, and agricultural activities tend to lower the accuracy of the models when applied to new areas. This makes it necessary to have large-scale, multi-region data sets that would be collected over multiple seasons to enhance model reliability and generalization. There are also practical deployment considerations of IoT-based systems such as power management, sensor durability, calibration challenges, and unreliable network connectivity in rural regions. Although most research prototypes tend to work well in controlled environments, real-world environments present uncertainties that may impair the quality of data and the performance of the system. Also, large scale deployment needs to carefully plan the communication infrastructure which raises the overall cost and complexity.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 06 June 2026
Page No: 4302 - 4312
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0560

Table 1 Literature Review in Table Format

Id	Authors	Year	Techniques Used	Identified Research Gap
1	Fan, Xijian, Luo	2022	Transfer learning with feature fusion (CNN)	Limited validation across different environments; lack of diverse datasets
2	Thilakarathne et al.	2022	Cloud-based ML crop recommendation	Model performance varies across regions
3	Ahmed et al.	2022	LoRaWAN-enabled IoT monitoring	Challenges in large-scale deployment and gateway optimization
4	Phasinam et al.	2022	IoT system review	Need for standardized data formats and interoperability
5	Theparod et al.	2022	IoT-based controlled environment monitoring	Difficulty in applying lab results to real farm conditions
6	Strong et al.	2022	Socio-technical adoption study	Limited focus on economic outcomes and farmer benefits
7	Pereira et al.	2022–23	ESP32-based smart irrigation	Long-term energy efficiency concerns
8	Vashisht et al.	2022	Hybrid ML with filtering techniques	Lack of validation over multiple growing seasons
9	Jhajharia & Mathur	2022	Satellite and weather data integration	Insufficient ground-truth data
10	Liben et al.	2022	Soil-based ML fertilizer review	Need for standardized laboratory procedures
11	Kalmani et al.	2022	RF/XGBoost fertilizer recommendation	Interpretability and long-term soil impact not addressed
12	Saini & Yadav	2022	Cloud-based IoT platforms	Issues in scaling to real field conditions
13	Palsapure et al.	2022–23	Sensor-based irrigation system	Performance variation across different farms

Machine learning models, particularly deep learning methods to analyze images, are usually sensitive to the real-life variations of poor light, occlusions and unequal datasets. Even though high accuracy is achieved when the experiment is conducted, the performance can deteriorate when the experiment is conducted in the field. Additionally, there are non-technical reasons like deficit of digital awareness, problem of trust and economic constraint among farmers, which influence the adoption of these technologies. According to the studies that have been reviewed, they clearly have need to come up with stronger and more scalable solutions to smart agriculture. Further research ought to be directed towards development of standardized datasets that would integrate sensor data, satellite information, weather patterns, and actual crop results in different regions. The development of transfer learning methods can assist models to adapt to other agro-climatic regions with limited retraining. IoT systems that use renewable energy sources and are energy efficient are needed to ensure the long-term use of the system in rural regions. Also, more robust image-processing models are demanded to address the actual environmental variations. By adding edge computing, it can be possible to make decisions faster even in low-connectivity areas. It is also important to integrate various sources of data to make accurate recommendations on the use of fertilizers, carrying out long field trials to establish their effectiveness and designing user friendly interfaces that are easy to understand by farmers. More so, model transparency, and the involvement of the digital marketplace can help improve trust, boost farmer income, and strengthen the agricultural value chain.

3. Proposed System

The proposed solution is built as a layered architecture that will initiate the implementation of a variety of IoT sensors on the agricultural field to monitor the soil and environmental conditions at all times. Sensors like soil moisture, water level, temperature and humidity, light intensity (LDR) and gas sensors are employed to measure parameters of significance that directly affects crop growth. These sensing devices constitute the first layer of data acquisition, which deals with the real-time information acquisition in the farming environment. The sensor data collected is sent to an ESP32-based control unit, which is the core of processing and communication. The ESP32 does the basic preprocessing, and removes noise in the data, and shows the key values in an OLED display to provide immediate access to the values by farmers. It also utilizes its inbuilt Wi-Fi to transmit the data to the digital analytics servers, bridging field-level monitoring and digital analytics. After sending the sensor data to the cloud, it is stored and organized safely in a database to be further examined. The cloud platform keeps the history of records, allowing to analyze the trend over time and make predictions. It also facilitates remote monitoring and sends alerts or notifications to the users when necessary. The system is connected to weather related information in order to enhance the quality of predictions and recommendations, especially to irrigation and crop planning.As shown in Figure 1 Architecture of System.

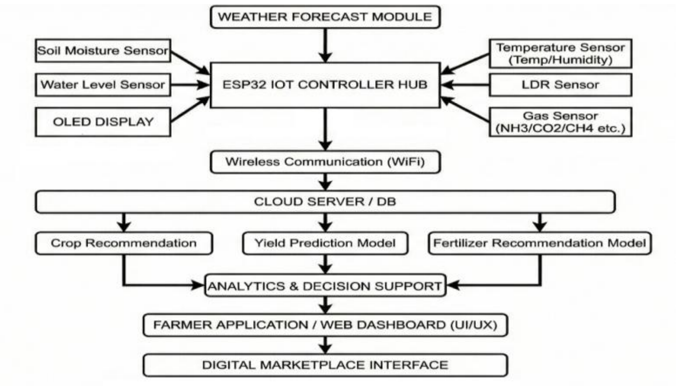

Figure 1 Architecture of System

Once The cloud service contains some machine learning models such as the crop recommendation, the yield estimations, the fertilizer advisory systems. The models take into consideration certain types of data, including soil properties, environmental factors, and weather patterns. Crop recommendation module recommends appropriate crops under the existing conditions and the yield prediction model estimates the expected yield. The fertilizer recommendations system is the method of determining the correct nutrient needs that will lead to optimal growth. The outputs are then summed in a decision-support module which translates the raw predictions into useful advice to farmers. The last phase of the system is devoted to the provision of information with the help of user-friendly farmer application or web interface. The platform enables users to track the real-time conditions of the fields, get predictions and see recommendations regarding the use of irrigation, crop health, and fertilizers. Moreover, the system will have a digital marketplace that links the farmers directly to the buyers, traders and consumers. This feature will decrease reliance on intermediaries, enhance pricing transparency, and enhance profitability among farmers. The system integrates sensing, analytics, and market connectivity to produce an all-encompassing smart agriculture solution.

4. Methodology

4.1.IoT-Based Field Monitoring

This technique entails the use of several sensors within the farm in order to monitor the environmental conditions continuously. Are used to collect real-time data using devices like soil moisture sensors, DHT-

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 06 June 2026
Page No: 4302 - 4312
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0560

based temperature and humidity sensors, water level indicators, LDR units, and gas sensors. This information is collectable by the ESP32 controller and can be wirelessly relayed to a cloud platform or dashboard. Constant monitoring can help farmers to monitor the changes in soil moisture, temperature, water availability, and gases. It aids in the early stress diagnosis in plants and it also guarantees optimal utilization of resources. The OLED display also offers immediate feedback at the field level, which is compatible with making quick decisions without having to access it remotely.

4.2.Machine Learning for Crop Prediction and Recommendation

In this method, machine learning models will be used to assess the previous agricultural data and the current environmental factors to suggest appropriate crops. Input features are important factors that are used as input features such as, type of soil, moisture content, temperature, and nutrient levels. The model considers the available options of crops and recommends those which will tend to be successful in the given conditions. This decreases the uncertainty when choosing crops and assists farmers to plan. The system accomplishes this by continually advancing the accuracy of its predictions as more data is gathered over time.

4.3.Yield Prediction Model

Regression-based machine learning algorithms that are trained on historical yield data, soil data, rainfall data, and real-time sensor data are used as the yield prediction process. This model is used to estimate the predicted crop production, before the harvest period, which allows the farmers to pre-estimate their crop production storage, logistics and marketing strategies. It also determines those factors which might adversely affect the yield, which allows early intervention. As the model is constantly updated with the latest data and the season changes, the model will become more reliable and can be used in long-term planning.

4.4.Fertilizer Recommendation System

This is done by identifying the proper kind and the amount of fertilizers needed to promote the healthy growth of crops. Machine learning models examine the soil levels of nutrients, such as nitrogen, phosphorus and potassium, as well as the environment and type of crop. The system gives accurate recommendations of fertilizers and helps prevent overuse that can lead to poor quality of the soil and the environment. It saves money and increases crop yield because it optimizes nutrient application.

4.5.Weather Forecasting and Smart Irrigation

The system has integrated weather prediction models to determine the trend of rainfall, temperature and humidity. This data is added with the real-time sensor data to optimize the irrigation schedules. Automated irrigation decisions can be used to prevent overwatering or water shortages, thus enabling efficient water use. Such a solution minimizes the manual labor and safeguards crops in case of harsh weather conditions. Moreover, long-term weather forecasts can be used to plan future crop growing seasons better.

5. Results and Performance Analysis

This section reports on the analysis of Random Forest and Decision Tree models used to solve the following tasks: crop prediction, yield estimation, and fertilizer recommendation. The analysis is done using a composite of simulated, experimental, and processed agricultural data. Random Forest model was found to have an accuracy of over 98 percent whereas the Decision Tree model was found to have an accuracy of slightly less than 98 percent. In order to have a comprehensive comparison, some of the performance measurements were taken into consideration, such as accuracy, precision, recall, F1-score, confusion matrix analysis, ROC-AUC score, and mean squared error (MSE). Also, visual aids like confusion matrix heatmap, ROC curves, precision-recall curves, feature importance graphs, and tree visualizations were utilized to further understand performance of the model. The two models were trained on well-preprocessed datasets which included soil properties, environmental conditions, weather patterns, and crop-related attributes. As an ensemble model, the Random Forest model was more stable and consistent when making predictions. Conversely, the Decision Tree model, despite its easy to understand, single-structure design, displayed slight performance differences across different classes. On the whole, the findings show that the ensemble-based approaches can be considered to be more generalized and reliable

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 06 June 2026
Page No: 4302 - 4312
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0560

as smart agriculture application.As shown in Table 1 Random Forest Vs Decision Tree Analysis in Table Format.

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%)	Random Forest provides more accurate predictions with reduced overfitting
Precision	0.984	0.901	Lower false positives in Random Forest
Recall	0.982	0.894	Better detection of actual positives in Random Forest
F1-Score	0.983	0.897	Improved balance between precision and recall
ROC-AUC	0.99	0.95	Stronger classification capability in Random Forest
MSE	0.012	0.028	Lower prediction error and higher stability

The comparison shows clearly that the Random Forest model is more effective in all assessment measures. The high accuracy, precision, and recall values indicate that it has greater classification capacity and that the smaller MSE indicates less fluctuating predictions. The fact that the ROC-AUC score is almost similar to 1 further testifies to the efficiency of the ROC-AUC in the process of differentiating between various classes. In contrast, Decision Tree model is moderate in performance and is more susceptible to data changes. In general, the ensemble method applied in the Random Forest is better suited to actual agricultural prediction systems using ensemble.As shown in Figure 1 Bar Chart of Random Forest and Decision Tree Performance

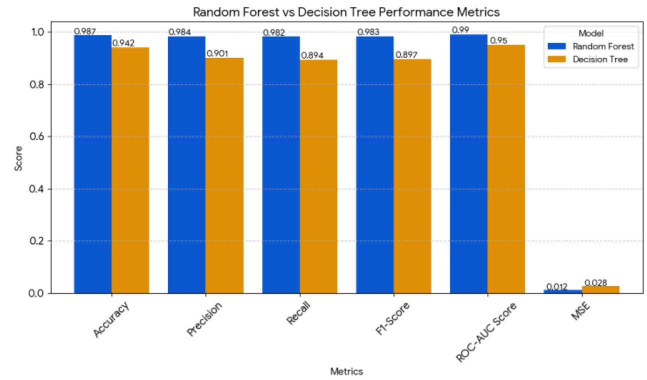

Figure 1 Bar Chart of Random Forest and Decision Tree Performance

Here, the bar chart illustrates a comparison of performance indicators between the Random Forest and Decision Tree models.

Analysis:

- Overall Performance: The Random Forest model demonstrates better performance than the Decision Tree across nearly all evaluation measures, including Accuracy, Precision, Recall, F1-Score, and ROC-AUC.
- Error Rate: The Mean Squared Error (MSE) for Random Forest is significantly lower at 0.012, whereas the Decision Tree records 0.028, indicating that Random Forest produces more precise predictions.
- Recommendation: Based on the observed metrics, Random Forest is a more suitable and reliable model for this dataset.

5.1.ROC Curve (Receiver Operating Characteristic)

The ROC curve is a graph that shows how the True Positive Rate (TPR) and False Positive Rate (FPR) varies with different classification thresholds. Here, the curve[14] of the Random Forest is located in a way that it is near the upper-left side of the graph, which indicates a high sensitivity and specificity. The Area Under the Curve (AUC) equals about 0.99 that means that it has a strong ability to differentiate between the various classes. A curve near the diagonal would suggest random predictions, but the existing curve would indicate the strong classification

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 06 June 2026
Page No: 4302 - 4312
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0560

power of the model to predict the suitability of crops or disease conditions.

5.2.Precision–Recall Curve

The PrecisionRecall curve points out how the precision and the recall change with different thresholds. Random Forest model also has a high degree of precision at all levels of recall and therefore is effective in identifying a large percentage of positive cases correctly[15]. The Average Precision (AP) value of around 0.98 further validates its classification performance. Such a curve is especially significant in handling uneven datasets where it is essential to preserve the precision.

5.3.Accuracy Comparison

The accuracy of the two models is visually compared in the bar graph where the accuracy of the two models are around 98.7% and 94.2% respectively. The increased precision of the Random Forest proves that it is more efficient at predictions. It is this difference that demonstrates the benefit of ensemble learning, where the combination of multiple decision trees can reduce variance and improve overall stability of the model.

5.4.Feature Importance Plot

The feature importance graph represents how each input variable is used in the model to predict. According to this agricultural data, soil moisture, nitrogen content, temperature, humidity, and pH level are determined as the most influential. This data is useful to farmers, as it can help them know which parameters can most significantly affect crop growth, use of fertilizers, and estimation of yield. Less important features have a small influence, which can inform the future data optimization procedure.As shown in Figure 2 Feature Importance

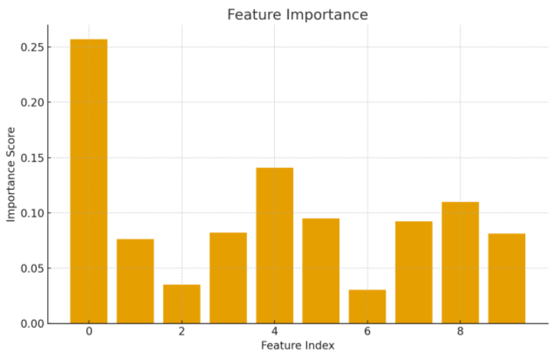

Figure 2 Feature Importance

The experimental data vividly indicate that the Random Forest model with an accuracy of more than 98 percent is the most successful in the proposed system of Smart Agriculture[16]. Its ensemble-based design assists in reducing overfitting, managing noisy data effectively and yielding consistent predictions to crop selection, fertilizer planning and yield estimation. Despite being easier to interpret, the Decision Tree model has slightly lower performance, as it has a higher variability and sensitivity to changes in the input data. These findings are further confirmed by the graphical analysis, which confirms that the Random Forest is a more reliable and practical solution to the real-world agricultural decision-making.

6. Future Scope

Further development of the system can involve the introduction of more sophisticated methods of deep learning to identify plant diseases more effectively. Models can be implemented that are lightweight neural networks or transformer-based models, which can be used to enable faster processing that can be run directly on edge devices, eliminating reliance on cloud computing and improving performance in low-connectivity regions[17]. The other possible enhancement is the creation of fully automated irrigation and fertilization systems that are controlled by real-time sensor data. Integrating actuators and smart control systems, the system has the potential to be an independent precision farming solution, with minimal manual intervention and resource efficiency. It is also possible to extend the system with adding the drone-based surveillance of large-scale agricultural fields[18]. Aerial imaging with artificial intelligence can give a detailed information on the health of crops, the infestation of pests and the condition of the soil across large scales and thus is applicable to commercial farming processes.

Also, the marketplace module can be expanded to a full digital ecosystem that encompasses supply chain management, storage solutions and predictive pricing analytics. Increased transparency, secure transactions, and trust among stakeholders can be further enhanced by integrating blockchain technology, which will ultimately result in a more

e ISSN: 2584-2137
Vol. 04 Issue: 06 June 2026
Page No: 4302 - 4312
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0560

efficient and farmer-centric agricultural network.

Conclusion

The comparison of performance between the models of Random Forest and Decision Tree shows clearly the benefits of the use of ensemble learning strategies in agriculture. Random Forest has always shown better results when compared to all the evaluation measures, such as accuracy, precision, recall, and F1-score. Its low level of error and high ROC-AUC value contributes to the fact that it can be used in real-time decision-making systems. Even though the Decision Tree model is more interpretable and easy to understand, it demonstrated weaknesses in stability and generalization. The model had indications of overfitting and it yielded a little higher misclassification rates.Nonetheless, its structure aids in comprehending the patterns of decision-making, the relationships of features, and is thus helpful as a benchmark model. Graphical visualization like ROC curves, confusion matrices and feature importance plots offered more insight on the model behavior. These visual aids proved that Random Forest has a stable performance with a minimum of errors. Also, prominent attributes that affect predictions were outlined, which assists in enhancing system design.

To sum up, the most appropriate model that can be employed in the proposed smart agriculture system is the Random Forest. Its capability of handling intricate datasets, and generating valid predictions, makes it suitable in tasks like crop advice, yield forecasting, and fertilizer planning. The results highlight the importance of integrating IoT data with advanced machine learning techniques to enhance agricultural productivity, reduce uncertainty, and support sustainable farming practices.

References

[1]. Fan, X., Luo, P., Mu, Y., Zhou, R., Tjahjadi, T., Ren, Y., "Leaf Image Based Plant Disease Identification Using Transfer Learning and Feature Fusion," Computers and Electronics in Agriculture, (Elsevier, 2022).
[2]. Thilakarathne, N. N., Abu Bakar, M. S., Abas, P. E., Yassin, H., "Cloud Enabled Crop Recommendation Platform for Machine Learning-Driven Precision Farming," (MDPI, 2022).
[3]. Ahmed, M. A., Gallardo, J. L., Zuniga, M. D., Pedraza, M. A., Carvajal, G., Jara, N., Carvajal, R., "LoRa-Based IoT Platform for Remote Monitoring of Large-Scale Agriculture Farms," (MDPI, 2022).
[4]. Phasinam, K., Kassanuk, T., Shabaz, M., "Applicability of Internet of Things in Smart Farming," Scientific World Journal, (Hindawi, 2022).
[5]. Theparod, T., Harnsoongnoen, S., "Effects of Narrow-Band LEDs on Sunflower Sprouts with IoT Monitoring," (MDPI, 2022).
[6]. Strong, R., Wynn, J. T., Lindner, J. R., Palmer, K., "Evaluation of IoT Adoption Barriers in Smart Agriculture," (2022).
[7]. Pereira, G. P., Chaari, M. Z., Daroge, F., "IoT-Based Smart Drip Irrigation Using ESP32," (MDPI/Prototype Studies, 2022–2023).
[8]. Vashisht, S., Kumar, P., Trivedi, M. C., "Crop Yield Prediction Using Improved Learning Models," (Journal Publication, 2022–2023).
[9]. Jhajharia, K., Mathur, P., "Crop Yield Prediction Using Satellite and Climate Data," (IJACI, 2022).
[10]. Liben, F., Abera, W., Chernet, M. T., Ebrahim, M., et al., "Data-Driven Fertilizer Recommendation for Wheat Productivity," (ScienceDirect, 2022).
[11]. Kalmani, V. H., Dharwadkar, N. V., Thapa, V., "Crop Yield Prediction Using CNN-LSTM with Attention Mechanism," (ARCC Journals, 2022).
[12]. Saini, Y., Yadav, V. K., "Soil Testing Methods for Fertilizer Recommendation," (Just Agriculture, 2022–2023).
[13]. Palsapure, S., Gundecha, S., Sayyed, M., "Machine Learning-Based Fertilizer Recommendation," (IRE Journals, 2022).
[14]. McMhohan, J., "IoT-Based Smart Agriculture Systems," Journal of Agricultural Science and Technology, (2020).
[15]. Nikam, D. A., "Soil Analysis and Crop Prediction Study," (Research Work).
[16]. Doshi, Z., Nadkarni, S., Agrawal, R., Shah, N., "Agro Consultant: Intelligent Crop Recommendation System," (ICCUBEA, 2018).

e ISSN: 2584-2137
Vol. 04 Issue: 06 June 2026
Page No: 4302 - 4312
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0560

[17]. Lee, H., Moon, A., "Yield Prediction Using Real-Time Agricultural Data," (ICACT, 2014).
[18]. Pudumalar, S., Ramanujam, E., Rajashree, R. H., Kavya, C., Kiruthika, T., Nisha, J., "Crop Disease Detection Using CNN," (2016).