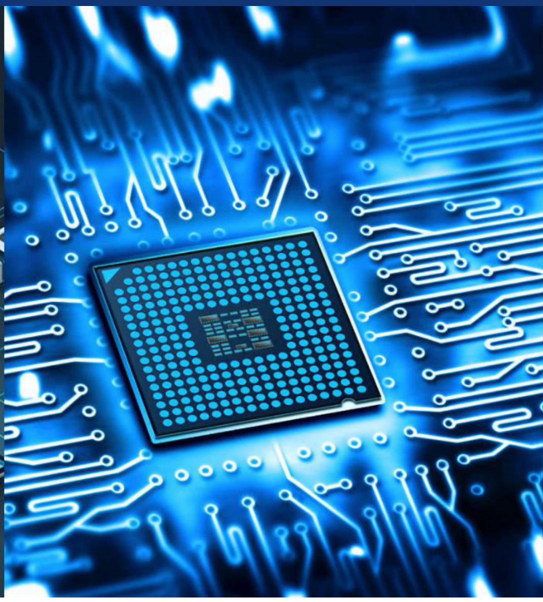
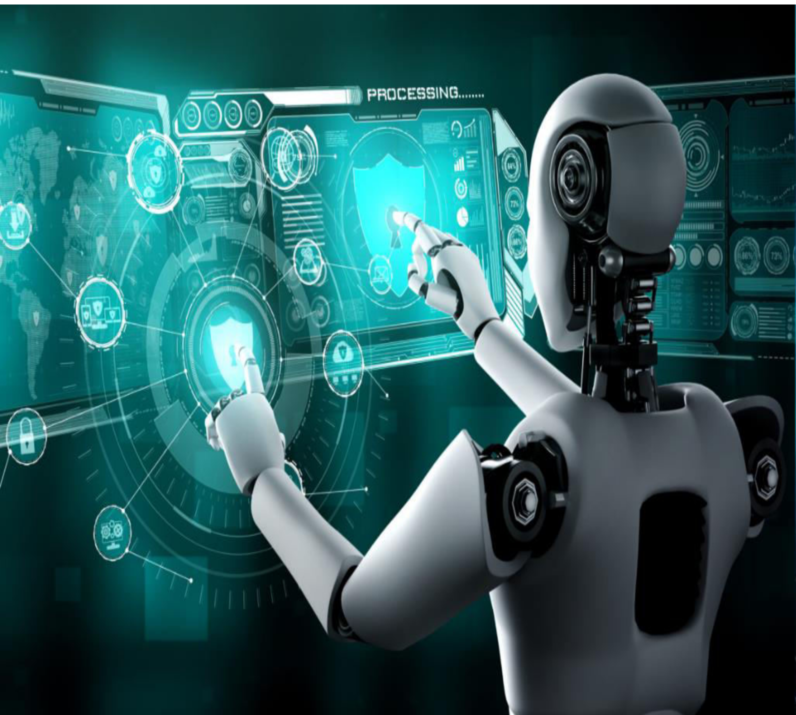
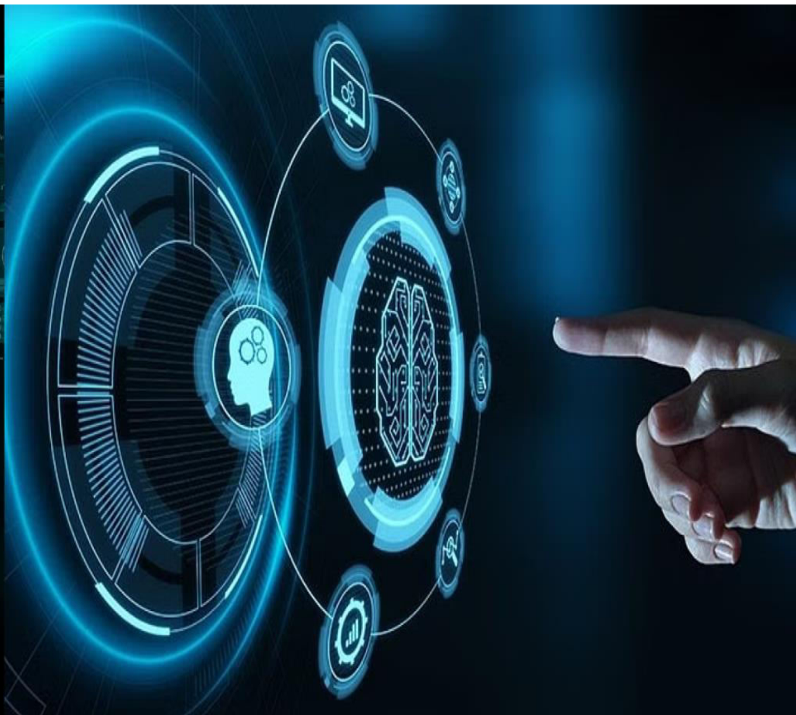
ISSN(O): 2320-9801
ISSN(P): 2320-9798


# International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)





Impact Factor: 8.625

Volume 13, Issue 1, January 2025

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

# Smart Agriculture Solution for Crop Prediction and Disease Detection

**Priyanshu Priyadarshi*, Kishan P Tikare*, Riddhi A Jain*, Alina Shibu*,**

**Dr. Vijayakumar Adaickalam#**

UG Student, School of Computer Science Engineering Presidency University, Bangalore, India *

Professor, School of Computer Science Engineering Presidency University, Bangalore, India#

**ABSTRACT**: Many nations, like India, struggle with problems like plant disease outbreaks and climate change while relying mostly on agriculture to sustain the livelihoods of millions of people. To solve these issues, a web application has been developed that provides real-time crop selection recommendations by analyzing data such as temperature, humidity, soil nutrient levels, pH, and rainfall. This solution revolutionizes plant disease management and crop yield prediction by leveraging the capabilities of artificial intelligence (AI) and machine learning (ML) to enable precise, data-driven agricultural decision-making. Seven machine learning models—Decision Tree, Naive Bayes, Support Vector Machine (SVM), Logistic Regression, Random Forest, XGBoost, and kNearest Neighbors (KNN)—were used for training and evaluation. Furthermore, the project includes a Convolutional Neural Network (CNN)-based Plant Disease Identification system that analyzes leaf photos to accurately identify and classify plant illnesses. By enabling farmers to act promptly to avert possible crop losses, this feature further improves agricultural efficiency.

**KEYWORDS:** CNN, Crop Recommendation, Machine Learning, Plant Disease Identification, Random Forest

## I. INTRODUCTION

Agriculture is one of the industries that has benefited most from the profound transformation brought about by the combination of artificial intelligence (AI) and machine learning (ML) technology. Smart crop recommendation systems and plant disease detection are two noteworthy developments among the many uses of these technology. One example of how machine learning and data science have transformed agricultural practices is the creation of systems that use image-based classification techniques to identify plant diseases and offer farmers customized crop suggestions.

This proposed system presents a robust Crop Recommendation System implemented using Python, leveraging multiple machine learning algorithms for classification tasks. It is intended to help farmers make data-driven choices regarding crop choice by examining factors such as soil nutrients, weather patterns, and geographic characteristics. Important characteristics that users can enter include rainfall information, the location, and the levels of potassium, phosphorus, and nitrogen in the soil. Through the utilization of an extensive dataset and the application of many categorization algorithms, the system forecasts which crops are best suited for production in particular areas. With a lower chance of crop failure, this can assist farmers in selecting the best crop quickly and simply.

Using image classification algorithms, the system also has a novel plant disease detection capability. The system makes use of convolutional neural networks (CNNs), a type of deep learning suitable for image processing and analysis tasks.In order to prevent crop loss, farmers can upload photographs of crop leaves exhibiting disease symptoms. The system correctly detects the plant and any potential diseases by comparing the uploaded image to many images in the collection.It has a user-friendly web interface created using Streamlit, the Crop Recommendation System makes it simple for farmers and other agricultural professionals to access and engage with it. Users may efficiently enter pertinent data, get crop advice, and gain insights into crop health with this interface.

This paper's organizational structure is as follows: The relevant literature is reviewed in Section 2, and Section 3 represents the proposed system, Section 4 discusses the results, and Section 5 concludes with the study's findings and processes behind the efficient crop recommendation. It also encompasses disease prediction and future directions.

© 2025 IJIRCCE | Volume 13, Issue 1, January 2025| DOI: 10.15680/IJIRCCE.2025.1301018

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

## II. RELATED WORK

The introduction of smart farming technologies has revolutionized agriculture by enhancing crop production, optimizing resource use, and efficiently addressing plant diseases. This document provides an in-depth analysis of smart crop recommendation systems and plant disease detection. Early identification of diseases helps farmers take timely action to prevent spread and minimize crop losses [11]. S.M. Pande et al. (2021) highlighted that machine learning (ML) has advanced crop prediction and management. Algorithms like Decision Trees and K-Nearest Neighbors (KNN) can forecast crop yields and suggest optimal crops based on geographic and environmental factors, supporting yield maximization and profitability [10]. Decision Trees: These models classify and predict crop performance by analyzing relationships between environmental factors like temperature, moisture, and pH. Their ability to visualize data patterns aids in decision-making and crop suitability analysis [3][10]. Random Forests: By combining multiple decision trees, random forests reduce overfitting and increase accuracy. This approach is effective in analyzing complex agricultural data and has been used in crop yield forecasting, providing reliable predictions and improving resource planning [6]. Convolutional Neural Networks (CNNs): CNNs, inspired by the brain's visual system, excel in plant disease detection by analyzing image data. They detect disease symptoms such as discoloration and pest damage, enabling timely diagnoses using on-site images or UAV-based remote sensing [2][4][7]. Combining deep learning with AI has improved plant disease detection, allowing for timely management. Deep Learning Models: CNNs have transformed disease detection by analyzing large datasets of annotated crop images. Systems like DeepCrop enable farmers to upload images of diseased crops and receive real-time diagnoses [2][5]. UAV-Based Monitoring: Bouguettaya et al. (2023) explored drones with high-resolution cameras for real-time aerial monitoring. Deep learning models such as YOLO and Faster R-CNN analyze aerial images to detect diseases, though challenges like data resolution and operational costs remain [7]. Web-based crop recommendation systems are vital for providing actionable insights to farmers. Responsive Design: Ensuring systems are accessible across devices (PCs, tablets, smartphones) improves usability. DeepCrop's user-friendly interface allows non-technical users to effectively diagnose diseases [2]. Integration of Real-Time Data: Incorporating real-time data, such as weather forecasts, enhances recommendation systems by providing up-to-date information for informed decision-making [10]. Optimized Performance: Lightweight models, such as those in CropCast, improve system speed and usability in rural areas with limited internet connectivity, enhancing user engagement and satisfaction [11].

## III. PROPOSED ALGORITHM

A crucial part of the framework for plant disease identification and smart crop recommendations is the implementation phase. This phase converts the theoretical ideas and approaches covered in the previous parts into useful applications [10]. It entails the methodical creation of the system, exacting testing procedures, and thorough analysis to guarantee adherence to predetermined standards. To ensure the system's dependability and efficacy, the following subsections describe the adopted methodology, testing and validation strategies, results analysis, and quality assurance mechanisms put in place [8].

A. Methodology

1. Datasets

Each of the 2,200 rows in the crop suggestion dataset has eight features: temperature, humidity, pH, rainfall, nitrogen, phosphorus, potassium, and a label. While temperature, humidity, and rainfall indicate typical environmental conditions, NPK levels signify soil nutrients. The label identifies the crop that is most suited for growing in the acidic or alkaline soil based on the pH value. The target variable for the prediction is this label. An auxiliary collection for crop disease diagnosis consists of 70,295 high-resolution photos of plants with various illnesses. This dataset, which is roughly 5 GB in size, includes pictures that are standardized to 128x128 pixel quality and span 38 different groups, such as 26 diseases and 14 plant species. Following model training, the most accurate algorithm was chosen for additional use. Additional images were captured to confirm the detection accuracy.

2. Preprocessing of Data

Data preprocessing is essential for converting unprocessed data into an analytical and machine learning-ready format. In plant disease identification, raw images often contain noise, necessitating preprocessing steps before integration into the learning model. For the crop recommendation dataset, missing values are imputed, and features are normalized to enhance algorithm performance. For the plant disease dataset, image preprocessing includes resizing to 128x128 pixels,

© 2025 IJIRCCE | Volume 13, Issue 1, January 2025| DOI: 10.15680/IJIRCCE.2025.1301018

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

image augmentation (flipping, rotation, and brightness adjustment), and noise reduction. These steps ensure uniformity and improve model robustness [4][7].

3. Train and Test Split
The train_test_split() function is used to separate the dataset into training and testing subsets from the scikitlearn library. In this framework, 80% of the dataset (1,760 records) is allocated for training, while the remaining 20% (440 records) is reserved for testing. Similarly, in the plant disease dataset, 80% of the images are used for training, and 20% are designated for validation [2][7].

4. Model Development and Prediction
Prediction involves applying a trained algorithm to new data to forecast outcomes. Predictions are generated using the predict() function on the test dataset, producing an array of predicted values. For The Random Forest Crop Recommendation System classifier is used, as it has demonstrated superior performance in agricultural applications by handling diverse feature sets effectively [6]. The model is trained using environmental and soil features to predict the crop label with high accuracy. A Convolutional Neural Network (CNN) architecture is implemented for picture-based disease detection. The model includes 4 layers, the input layer which processes the images, the convolutional layers which extract spatial features from images using filters, the pooling layers which reduce spatial dimensions to prevent overfitting and the fully connected layers which integrate the extracted features to classify plant diseases [2][4][5].

5. Classification Report and Confusion Matrix
The classification report and confusion matrix, which were taken from the scikit-learn metrics module, are utilized to evaluate model performance. The confusion matrix provides detailed numbers of false positives, true positives, true negatives, and true negatives. Precision measures the precision of optimistic forecasts and is computed as:

$$\text{Precision} = \frac{TP}{TP + FP}$$

where FP stands for False Positives and TP for True Positives.

Recall assesses the model's capacity to recognize every relevant positive cases:

$$\text{Recall} = \frac{TP}{TP + FN}$$

Where FN is False Negatives.

The harmonic mean of precision and recall, or F1-Score, is calculated as:

$$\text{F1} - \text{Score} = \frac{2 * PR}{(P+R)}$$

An F1-Score is a number between 0 and 1, where 1 denotes ideal performance.

6. Accuracy
Accuracy is a key metric representing the proportion of accurate forecasts to all forecasts. It is computed as:

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (4)$$

Where TP-True Positive; FP-False Positive; TN-True Negative; FN-False Negative

This metric is applied to both the crop recommendation and plant disease detection models, ensuring high prediction reliability [1][5][6].

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

B. Model Evaluation

The effectiveness of the plant disease detection model is evaluated by dividing the dataset, allocating 20% for validation and 80% for training. The validation dataset is utilized in order to evaluate precision through the use of the learned model and the extraction of exact characteristics. The assessment makes that the system successfully detects crop diseases and satisfies the required accuracy levels.

C. Integration and Deployment of Systems

The following components of the suggested system can be incorporated into a web-based platform:

1. Crop Recommendation: A responsive interface that allows users to enter soil and environmental data to obtain crop recommendations [2][10].
2. Disease Detection: A function that allows users to upload images that instantly diagnoses plant illnesses and offers solutions.
3. Real-Time Data Integration: To offer thorough agricultural insights, external data sources like weather forecasts are integrated. [2][8].

## IV. RESULTS AND DISCUSSIONS

The accuracy of seven classification algorithms in crop recommendation tasks was assessed in this study. With an astounding accuracy of 99.55%, the findings, which are compiled in Table 1, show how the Random Forest algorithm performs better than any other models. This demonstrates its dependability and resilience in precisely forecasting appropriate crops under various circumstances.

Table 1. Accuracy Across Classification Algorithms

| **Algorithm** | **Accuracy (%)** |
|---|---|
| Decision Tree | 90.00 |
| Gaussian Naive Bayes | 99.09 |
| Support Vector Machine (SVM) | 10.68 |
| Logistic Regression | 95.23 |
| Random Forest | 99.55 |
| XGBoost | 99.09 |
| K-Nearest Neighbor (KNN) | 97.50 |

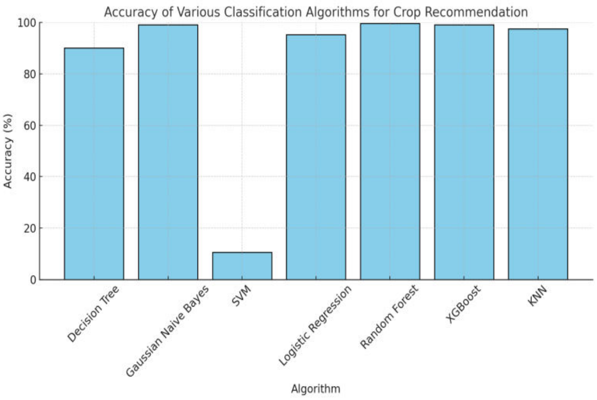
The Random Forest model consistently outperforms alternative algorithms in the majority of cases, as seen in the table above, across a range of crop kinds.Although certain models, including XGBoost and Gaussian Naive Bayes, had nearly identical accuracy, their performance differed depending on the crop. Vector Machine Support With an accuracy of 10.68%, SVM performed poorly on the other hand, most likely as a result of its sensitivity to the big and complicated dataset. The whole investigation demonstrates Random Forest's adaptability and dependability for crop recommendation tasks. The Random Forest model is a viable option for a variety of agricultural applications due to its stability and adaptability, demonstrating its ability to provide solid and consistent results.

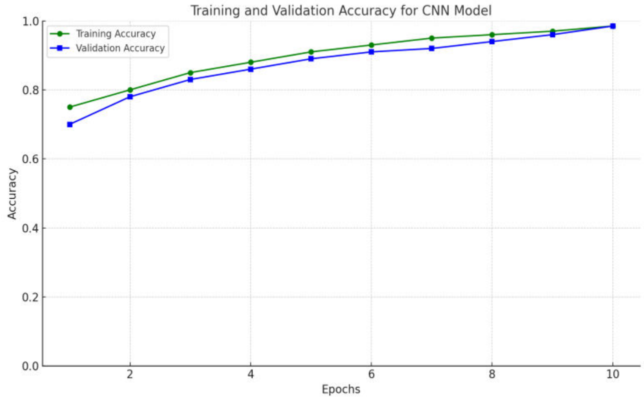
© 2025 IJIRCCE | Volume 13, Issue 1, January 2025| DOI: 10.15680/IJIRCCE.2025.1301018

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)




A. Plant Disease Detection Using CNN

In terms of identifying plant diseases, a Convolutional Neural Network (CNN) was used to examine a dataset including 17,572 photos from 38 different classifications, including different plant illnesses and favorable circumstances. The tf.keras.utils.image_dataset_from_directory method was used to arrange the photos into validation sets after they were resized to 128x128 pixels. To make predictions, the model used a CNN architecture that had already been trained and loaded from a.keras file [5].

A dropout layer was used during training to randomly deactivate a subset of the input units in order to avoid overfitting. A representative photograph of an early blight-affected potato plant was processed for testing and visualization. After being read, the image was enlarged to the proper dimensions, translated from BGR to RGB format using OpenCV, and then fed into the CNN model for prediction. A distribution of probabilities across all classes was the result, and the np.argmax function was used to determine which class had the highest likelihood. The accuracy of the model in recognizing diseases was confirmed by showing the predicted class label next to the original image.

This method demonstrates how effective CNNs are in automating plant disease identification. The model can offer prompt and accurate diagnostics by spotting minute characteristics like spots, discolouration, and other signs. This allows for early interventions and enhances crop health and management techniques.

B. CNN Model Performance Analysis

The CNN-based plant disease detection model's performance was assessed using training and validation accuracy as well as training and validation loss, as the graph below illustrates.

High Accuracy: The model's capacity to successfully identify between healthy and unhealthy plants was proved by the steady improvement in both training and validation accuracy.The validation accuracy peaked at 98.5%, demonstrating the dependability of the model.

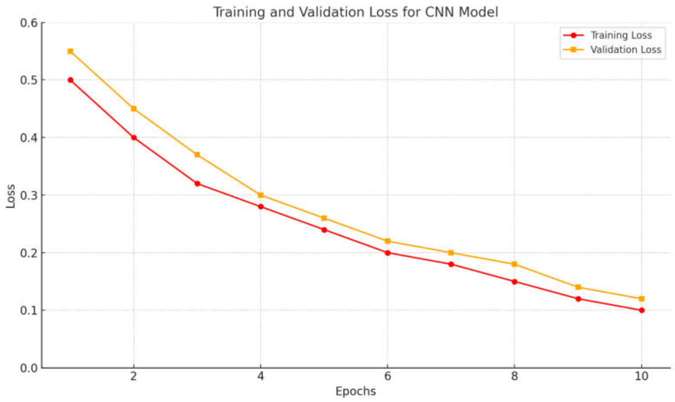
Low Loss: Robust learning and successful generalization to unknown data were shown by the training and validation loss curves, which displayed a gradual drop with little differences between them.

© 2025 IJIRCCE | Volume 13, Issue 1, January 2025| DOI: 10.15680/IJIRCCE.2025.1301018

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)




The general patterns demonstrate the model's excellent fitting skills, even in the face of little variations in validation outcomes. Accurate disease diagnosis was made possible by the model's ability to capture complex features including discolouration, leaf patches, and wilting patterns. The knowledge gained from these tests highlights CNN's capacity to provide accurate findings for plant disease identification, opening the door to more intelligent farming methods and higher yields.

## V. CONCLUSION AND FUTURE WORK

In conclusion, by utilizing state-of-the-art technologies and methodologies, this suggested system has effectively created an advanced system for smart crop recommendation and plant disease detection. The technology offers practical insights to improve crop management practices and mitigate the adverse effects of plant diseases by fusing real-time data processing with machine learning techniques. The system's dependability, scalability, and maintainability are guaranteed by its modular and scalable architecture, which was created in accordance with industry standards. Furthermore, the system is now accessible to farmers regardless of their technical proficiency thanks to the use of best practices in database design and user-focused interface development, which facilitates smooth interaction and use.

This innovation is a significant advancement in precision farming that shows the value of coordinated tactics, cutting-edge technology, and conformity to accepted industry standards. Using state-of-the-art technology for disease diagnosis and crop recommendations gives farmers the resources they need to make wise decisions and implement sustainable farming methods.

In the future, the system can be further enhanced and expanded by implementing regular dataset updates to increase model relevance and accuracy by supplementing it with additional region-specific samples and cases.

Integration of Advanced Techniques like reinforcement learning to improve current models and tactics, and integrating multispectral and hyperspectral imaging data to obtain a deeper understanding of plant health. The use of environmental data, such as rainfall statistics, soil moisture, and local weather conditions, can help provide more precise crop recommendations, hence lowering crop losses and increasing yields. Including tools to forecast which crops would yield the highest profits for farmers by taking into account market pricing, trends, cultivation costs, and possible returns will help farmers optimize their earnings. Providing new farmers with tools for managing crop growth schedules to help them with irrigation, fertilizing, and pest control, among other farming procedures, to ensure efficient and timely methods. Edge Computing Integration by creating portable gadgets with cameras and real-time sensors for agricultural monitoring can help extend this system. Creating a collaborative platform that allows farmers, academics, and agronomists to exchange information on novel cases, data, and insights. The accuracy and applicability of models can be greatly improved by crowdsourced data collection. Artificial Intelligence-Powered Decision Support Systems can be created by applying reinforcement learning algorithms to improve decision-making procedures for the best use of resources and yield-improvement tactics.

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 8.625| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

## REFERENCES


1. M. Shripathi Rao, A. Singh, N. V. Subba Reddy, and D. U. Acharya, “Crop prediction using machine learning,” Journal of Physics: Conference Series, vol. 2161, no. 1, p. 012033, Jan. 2022.
2. Md. Manowarul Islam et al., “DeepCrop: Deep learning-based crop disease prediction with web application,” Journal of Agriculture and Food Research, vol. 14, pp. 100764–100764, Dec. 2023.
3. A. Kar, N. Nath, U. Kemprai, and Aman, “Performance Analysis of Support Vector Machine (SVM) on Challenging Datasets for Forest Fire Detection,” International Journal of Communications, Network and System Sciences, vol. 17, no. 2, pp. 11–29, Feb. 2024.
4. G. Rani, E. T. Venkatesh, K. Balaji, Balasaraswathi Yugandher, Adiki Nithin Kumar, and M Sakthimohan, “An automated prediction of crop and fertilizer disease using Convolutional Neural Networks (CNN),” 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Apr. 2022.
5. D. Tirkey, K. K. Singh, and S. Tripathi, “Performance analysis of AI- based solutions for crop disease identification, detection, and classification,” Smart Agricultural Technology, vol. 5, p. 100238, Oct. 2023.
6. K. Ramu and K. Priyadarsini, “A Review on Crop Yield prediction Using Machine Learning Methods,” 2021 2nd International Conference on Smart Electronics and Communication (ICOSEC), Oct. 2021.
7. A. Bouguettaya, H. Zarzour, A. Kechida, and A. M. Taberkit, “A survey on deep learning-based identification of plant and crop diseases from UAV-based aerial images,” Cluster Computing, Aug. 2022.
8. K. Jhajharia, P. Mathur, S. Jain, and S. Nijhawan, “Crop Yield Prediction using Machine Learning and Deep Learning Techniques,” Procedia Computer Science, vol. 218, pp. 406–417, 2023.
9. M. Shripathi Rao, A. Singh, N. V. Subba Reddy, and D. U. Acharya, “Crop prediction using machine learning,” Journal of Physics:
10. S. M. PANDE, P. K. RAMESH, A. ANMOL, B. R. AISHWARYA, K.ROHILLA, and K. SHAURYA, “Crop Recommender System Using Machine Learning Approach,” 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), Apr. 2021.
11. C. Raju, A. D, and A. P. B, “CropCast: Harvesting the future with interfused machine learning and advanced stacking ensemble for precise crop prediction,” Kuwait Journal of Science, p. 100160, Dec. 2023.
12. P. Chauhan, Hardwari Lal Mandoria, A. Negi, and R. S. Rajput, “Plant Diseases Concept in Smart Agriculture Using Deep Learning,” Advances in environmental engineering and green technologies book series, pp.139–153, Oct. 2020.